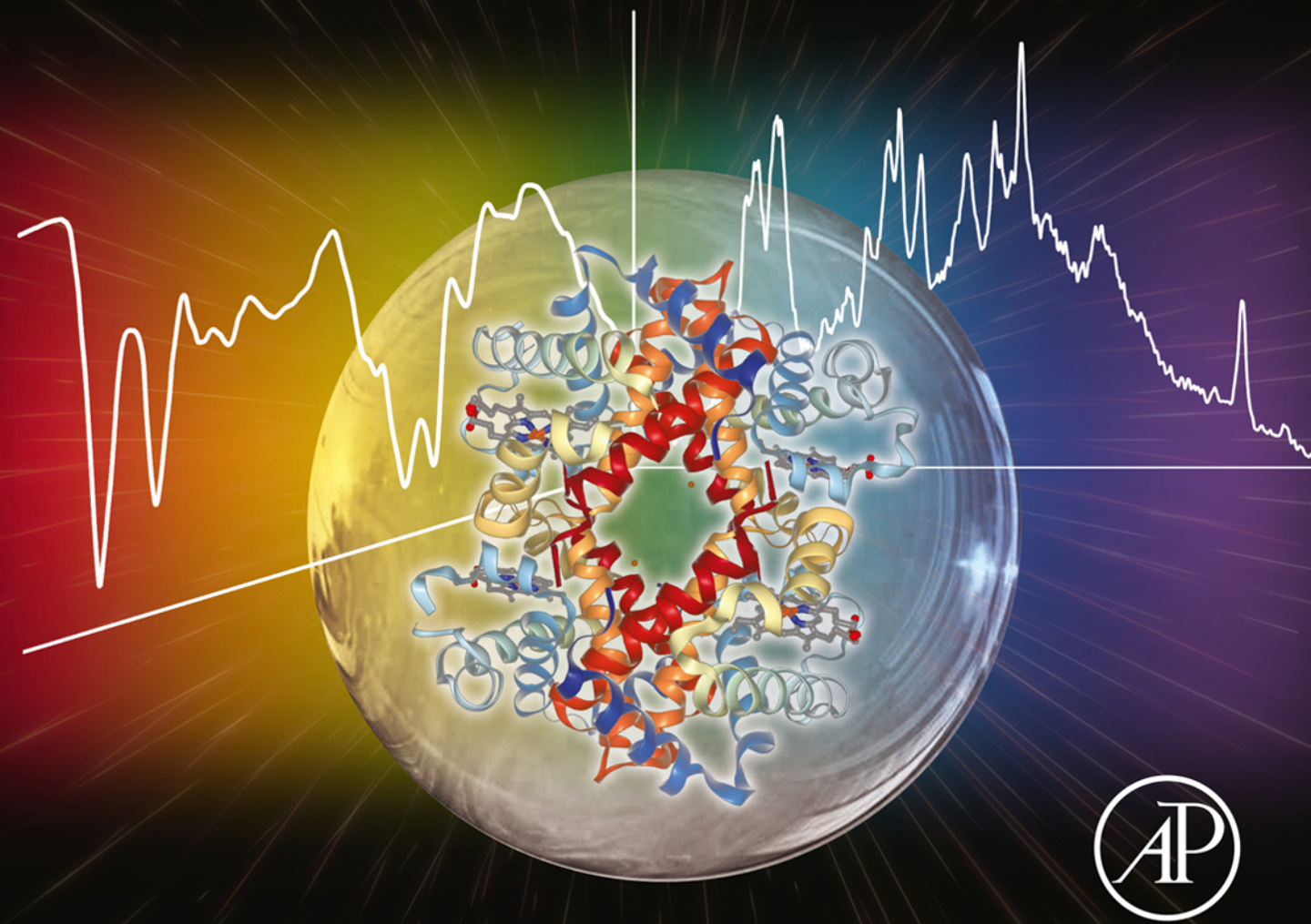
VIBRATIONAL SPECTROSCOPY IN PROTEIN RESEARCH

From Purified Proteins to Aggregates and Assemblies

Edited by
Yukihiro Ozaki, Malgorzata Baranska
Igor K. Lednev, and Bayden R. Wood



Vibrational Spectroscopy in Protein Research

This page intentionally left blank

Vibrational Spectroscopy in Protein Research

From Purified Proteins to Aggregates and Assemblies

Edited by

Yukihiro Ozaki

School of Science and Technology, Kwansei Gakuin University, Hyogo, Japan

Malgorzata Baranska

Faculty of Chemistry, Jagiellonian University, Krakow, Poland
Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

Igor K. Lednev

Department of Chemistry, University at Albany, State University of New York, Albany, NY, United States

Bayden R. Wood

Centre for Biospectroscopy, School of Chemistry, Monash University, Clayton, VIC, Australia

ACADEMIC PRESS

An imprint of Elsevier

Academic Press is an imprint of Elsevier
125 London Wall, London EC2Y 5AS, United Kingdom
525 B Street, Suite 1650, San Diego, CA 92101, United States
50 Hampshire Street, 5th Floor, Cambridge, MA 02139, United States
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, United Kingdom

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

ISBN: 978-0-12-818610-7

Publisher: Andre Gerhard Wolff
Acquisitions Editor: Peter B. Linsley
Editorial Project Manager: Sara Pianavilla
Production Project Manager: Punithavathy Govindaradjane
Cover Designer: Aliaksandra Sikirzhytskaya

Typeset by MPS Limited, Chennai, India

Contents

This page intentionally left blank

List of contributors

Ramon A. Alvarez-Puebla Department of Physical and Inorganic Chemistry and EMaS, Universitat Rovira i Virgili, Tarragona, Spain; ICREA, Barcelona, Spain

Matthew J. Baker WestCHEM, Department of Pure & Applied Chemistry, Technology and Innovation Centre, University of Strathclyde, Glasgow, United Kingdom; ClinSpec Dx Ltd., Technology and Innovation Centre, Glasgow, United Kingdom

Malgorzata Baranska Faculty of Chemistry, Jagiellonian University, Krakow, Poland; Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

James W. Beattie Department of Chemical Engineering and Department of Life Sciences, Imperial College London, London, United Kingdom

Ewelina Bik Faculty of Chemistry, Jagiellonian University, Krakow, Poland; Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

Franck Bonnier Faculty of Pharmacy, EA 6295 NMNS, University of Tours, Tours, France

Petr Bouř Institute of Organic Chemistry and Biochemistry, Academy of Sciences, Prague, Czech Republic

Clément Bruno Faculty of Pharmacy, EA 6295 NMNS, University of Tours, Tours, France; CHRU Tours, Laboratory of Biochemistry and Molecular Biology, Tours, France; UMR INSERM U 1253 - iBrain, University of Tours, Tours, France

Holly J. Butler WestCHEM, Department of Pure & Applied Chemistry, Technology and Innovation Centre, University of Strathclyde, Glasgow, United Kingdom; ClinSpec Dx Ltd., Technology and Innovation Centre, Glasgow, United Kingdom

Bernadette Byrne Department of Chemical Engineering and Department of Life Sciences, Imperial College London, London, United Kingdom

Hugh J. Byrne School of Physical and Clinical & Optometric Sciences, Technological University Dublin, Dublin, Ireland

James M. Cameron WestCHEM, Department of Pure & Applied Chemistry, Technology and Innovation Centre, University of Strathclyde, Glasgow, United Kingdom

Karolina Chrabaszcz Faculty of Chemistry, Jagiellonian University, Krakow, Poland; Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

Krzysztof Czamara Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

Aleksandra Dorosz Faculty of Chemistry, Jagiellonian University, Krakow, Poland; Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

Jakub Dybas Faculty of Chemistry, Jagiellonian University, Krakow, Poland; Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

Luca Guerrini Department of Physical and Inorganic Chemistry and EMaS, Universitat Rovira i Virgili, Tarragona, Spain

Joseph Handen Department of Chemistry, University at Albany, State University of New York, Albany, NY, United States

Jana Hudecová Institute of Organic Chemistry and Biochemistry, Academy of Sciences, Prague, Czech Republic; Department of Optics, Palacký University Olomouc, Olomouc, Czech Republic

Mika Ishigaki Raman Project Center for Medical and Biological Applications, Shimane University, Matsue, Japan; Faculty of Life and Environmental Sciences, Shimane University, Matsue, Japan

Keita Iwasaki The United Graduate School of Agricultural Sciences, Tottori University, Tottori, Japan

Young Mee Jung Department of Chemistry, and Institute for Molecular Science and Fusion Technology, Kangwon National University, Chunchon, Korea

Hideki Kandori Department of Life Science and Applied Chemistry and OptoBioTechnology Research Center, Nagoya Institute of Technology, Nagoya, Japan

Kota Katayama Department of Life Science and Applied Chemistry and OptoBioTechnology Research Center, Nagoya Institute of Technology, Nagoya, Japan

Sergei G. Kazarian Department of Chemical Engineering and Department of Life Sciences, Imperial College London, London, United Kingdom

Janina Kneipp Department of Chemistry, Humboldt-Universität zu Berlin, Berlin, Germany

Kamila Kochan Centre for Biospectroscopy, School of Chemistry, Monash University, Clayton, VIC, Australia

Igor K. Lednev Department of Chemistry, University at Albany, State University of New York, Albany, NY, United States

Bernhard Lendl Institute of Chemical Technologies and Analytics, Vienna University of Technology, Vienna, Austria

Kássio M.G. Lima Biological Chemistry and Chemometrics, Institute of Chemistry, Federal University of Rio Grande do Norte, Natal, Brazil

Ewelina Lipiec Department of Physics of Nanostructures and Nanotechnology, Institute of Physics, Jagiellonian University, Krakow, Poland

Katarzyna Majzner Faculty of Chemistry, Jagiellonian University, Krakow, Poland; Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

Kamilla Malek Faculty of Chemistry, Jagiellonian University, Krakow, Poland; Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

Francis L. Martin School of Pharmacy and Biomedical Sciences, University of Central Lancashire, Preston, United Kingdom; Biocel Ltd, Hull, United Kingdom

Katarzyna M. Marzec Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

Lisa M. Miller National Synchrotron Light Source II, Brookhaven National Laboratory, Upton, NY, United States

Yasuhisa Mizutani Department of Chemistry, Graduate School of Science, Osaka University, Toyonaka, Osaka, Japan

Camilo L.M. Morais School of Pharmacy and Biomedical Sciences, University of Central Lancashire, Preston, United Kingdom

Isao Noda Department of Materials Science and Engineering, University of Delaware, Newark, DE, United States

Hemanth Noothalapati Raman Project Center for Medical and Biological Applications, Shimane University, Matsue, Japan; Shimane University Administration Office, Shimane University, Matsue, Japan

Yukihiro Ozaki School of Science and Technology, Kwansei Gakuin University, Hyogo, Japan

Marta Z. Pacia Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

Zhijun Pan State Key Laboratory of Structural Chemistry, Fujian Institute of Research on the Structure of Matter, Chinese Academy of Sciences, Fuzhou, China

Drishya Rajan Parachalil School of Physical and Clinical & Optometric Sciences, Technological University Dublin, Dublin, Ireland; FOCAS Research Institute, Technological University Dublin, Dublin, Ireland

Yeonju Park Department of Chemistry, and Institute for Molecular Science and Fusion Technology, Kangwon National University, Chunchon, Korea

Marfran C.D. Santos Biological Chemistry and Chemometrics, Institute of Chemistry, Federal University of Rio Grande do Norte, Natal, Brazil

Sebastian Schlücker Department of Chemistry and Center for Nanointegration Duisburg-Essen (CENIDE), University of Duisburg-Essen, Essen, Germany

Andreas Schwaighofer Institute of Chemical Technologies and Analytics, Vienna University of Technology, Vienna, Austria

Cai Li Song Department of Chemical Engineering and Department of Life Sciences, Imperial College London, London, United Kingdom

Yuling Wang Department of Molecular Sciences, ARC Centre of Excellence for Nanoscale BioPhotonics, Macquarie University, Sydney, NSW, Australia

Ewelina Wiercigroch Faculty of Chemistry, Jagiellonian University, Krakow, Poland; Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

Adrianna Wislocka Faculty of Chemistry, Jagiellonian University, Krakow, Poland; Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

Bayden R. Wood Centre for Biospectroscopy, School of Chemistry, Monash University, Clayton, VIC, Australia

Tatsuyuki Yamamoto Raman Project Center for Medical and Biological Applications, Shimane University, Matsue, Japan; Faculty of Life and Environmental Science, Shimane University, Matsue, Japan

Yuying Zhang Medical School of Nankai University, Tianjin, China

Wei Zhuang State Key Laboratory of Structural Chemistry, Fujian Institute of Research on the Structure of Matter, Chinese Academy of Sciences, Fuzhou, China

This page intentionally left blank

Preface

Recently, a variety of state-of-the art vibrational spectroscopy techniques have been developed rapidly, for example, Raman imaging, Raman optical activity (ROA), surface-enhanced Raman scattering (SERS), nonlinear Raman spectroscopy, cascade laser IR spectroscopy, infrared (IR) imaging, near-infrared (NIR) imaging, and time-resolved vibrational spectroscopy. Moreover, several spectral analysis methods, such as quantum chemical calculations, chemometrics, and two-dimensional correlation spectroscopy, have also been advanced. These experimental and spectral analysis methods have all been extensively used for a variety of investigations of proteins.

The purpose of this book is to demonstrate the usefulness of these modern vibrational spectroscopic methods in protein research. This book places some emphasis on in vivo, in situ, and noninvasive studies of proteins. It also points to the imaging of biological systems including proteins; Raman, IR, and NIR imaging are introduced. Most of the chapters are concerned not only with simple purified proteins, but also with rather complicated protein systems, such as protein aggregates, fibrous proteins, membrane proteins, and protein assemblies. Moreover, the book can be connected with medical science to some extent.

The book outlines cutting-edge vibrational spectroscopy methods by the front-runners of this field. All the contributors are experts in both modern vibrational spectroscopy and protein research. A few new techniques, such as quantum cascade laser-based IR transmission spectroscopy and SERS with nanotag design, are introduced in this book.

This book aims at making a strong bridge between molecular spectroscopists and researchers in life sciences. Thus it is suitable for molecular spectroscopists who are interested in protein research and for protein scientists who are interested in molecular spectroscopy. This book is useful for scientists, engineers, and graduate students in various fields of protein research including chemistry, biological sciences, pharmaceutical science, agricultural science, and medicine. One can use this book as a text for a course, for example, at a graduate school.

We do hope this book can inspire readers to employ vibrational spectroscopy techniques for various investigations of proteins and/or to open new directions of protein research based on vibrational spectroscopy.

In closing, we would like to thank Dr. Aliaksandra Sikirzhytskaya for preparing the wonderful cover of this book and also to Mr. Peter Linsley and Ms. Sara Pianavilla of Elsevier for their continuous efforts in publishing this book.

Yukihiro Ozaki
Malgorzata Baranska
Igor K. Lednev
Bayden R. Wood
November 2019

CHAPTER 1

ATR-FTIR spectroscopy and spectroscopic imaging of proteins

Bernadette Byrne, James W. Beattie, Cai Li Song and Sergei G. Kazarian

Department of Chemical Engineering and Department of Life Sciences, Imperial College London, London, United Kingdom

1.1 Introduction

1.1.1 Study of protein behavior—protein in solution, film, and tissue

Fourier transform infrared (FTIR) spectroscopy is an excellent tool to obtain structural and chemical information of proteins. There are multiple infrared spectroscopic methods that can be applied to gain key information on secondary structure and the behavior of proteins. One such method is the use of attenuated total reflection (ATR)−FTIR spectroscopy. ATR-FTIR spectroscopy has proven to be useful in the investigation of biomedical samples, protein crystallization, surface interaction with proteins, and behavior of proteins that undergo structural changes due to isolation methods, such as therapeutic antibodies [1−4].

ATR-FTIR spectroscopy is a powerful label-free technique that allows the generation of a spectrum containing absorption bands specific to the chemical species present in a measured sample. The ATR mode reflects IR radiation via an internal reflective element (IRE) which is in direct contact with the sample. This reflection of the IR radiation generates an evanescent wave at the interface between the sample and the IRE which probes the sample. The key parameter for evanescent wave generation is that the critical angle for internal reflection must be exceeded. This evanescent wave is reflected causing attenuation of the IR radiation exiting the IRE [5]. The IRE element consists of infrared transparent material that has a higher refractive index than the sample being analyzed. IREs such as diamond, Ge, Si, and ZnSe are used in the study of proteins as they have a higher refractive index than that of protein samples.

The depth of penetration (d_p) by this evanescent wave is determined by multiple factors including the IRE refractive index (n_1), sample refractive index (n_2), wavelength (λ), and the angle of incidence (θ) as shown by the Harrick equation (Eq. 1.1) below [6−8].

Vibrational Spectroscopy in Protein Research.
DOI: https://doi.org/10.1016/B978-0-12-818610-7.00001-3

$$d_p = \frac{\lambda}{2\pi n_1 \left(\sin^2\theta - \left(n_2/n_1\right)^2\right)^{1/2}} \tag{1.1}$$

ATR-FTIR spectroscopy has a limited depth of penetration (c. 0.2−3 μm) due to the refractive index of the IRE crystal and the sample used. Therefore the spectrum generated is representative of the surface layer of the sample adjacent to the ATR crystal and not of the bulk phase [7,8]. It should be noted that the surface layer of the sample may not be representative of the bulk of a given sample [9].

1.1.2 Interaction of proteins with infrared—understanding amide bands

1.1.2.1 Interpreting secondary structures from amide bands

Proteins show clear interaction with infrared radiation through the generation of amide vibrational bands in measured spectra, the most prominent being the amide I, II, and III bands of protein spectra as illustrated by Fig. 1.1.

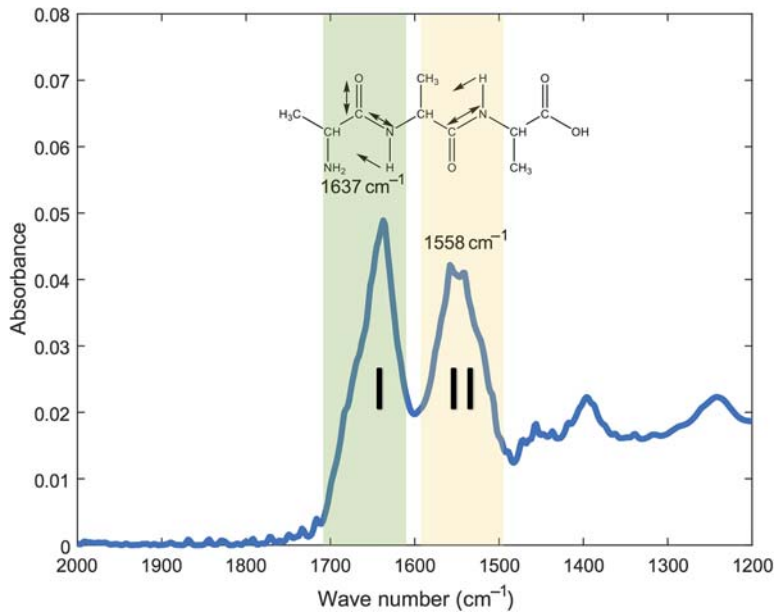


Figure 1.1

Example of a protein attenuated total reflection (ATR)—Fourier transform infrared (FTIR) spectrum highlighting the key amide I and II bands. The amide I band (1637 cm^{-1}) is highlighted in green and the amide II band (1558 cm^{-1}) is highlighted in yellow. The protein used to generate the spectrum is IgG4.

The amide I band is very useful for determining changes in the secondary structure of protein due to the fact that it corresponds to the vibrational modes of the chemical species that make up the peptide bonds of a protein. The amide I band typically absorbs at $\approx 1650\ cm^{-1}$. In Fig. 1.1, the amide I vibrational mode is observed at $1637\ cm^{-1}$, as a result of a C═O stretch weakly coupled to C−N stretching as well as the N−H in-plane bending mode as delineated in green [1,10]. The exact position of the amide I band is determined by hydrogen bonding as well as the protein backbone conformation [11].

The amide II band is similarly very useful for determination of secondary structural changes in proteins, as differences in secondary structure components, such as disordered regions and aggregated β-sheets, are reflected in this band. The amide II band, made up of N−H in-plane bending and C−N stretching [1,10,11], typically absorbs at $\approx 1550\ cm^{-1}$. In Fig. 1.1 this band is present at $1558\ cm^{-1}$.

1.1.2.2 Qualitative and quantitative analysis

Qualitative and quantitative interpretations should both be used to interpret a measured infrared spectrum. Qualitative methods include assigning vibrational modes present in a spectrum using amide bands as previously shown in Fig. 1.1. This method allows for changes in protein structure, such as changes that occur when a protein aggregates, to be described. By comparing peak height of the spectral band and its position to a difference spectrum at a given wave number (i.e., amide I) certain structural features including formation of intermolecular β-sheet aggregates can be detected [1].

ATR-FTIR spectroscopic measurements can be interpreted quantitatively to determine the concentration of analytes. Protein concentrations can be determined by integration of the amide II band and through generation of a calibration curve of the integrated absorbance. For example, the protein A ligand concentration in immunoaffinity resin adjacent to an IRE was determined in this manner, as shown in Fig. 1.2A. This quantification method provides a viable alternative to using UV spectroscopy to quantify protein [12]. This method of quantification also allows for application of ATR-FTIR spectroscopy as a means of determining binding capacity of the stationary phase (of protein A resin) as denoted in Fig. 1.2B [12].

An additional method of quantifying ATR-FTIR spectra is to use a calibration curve as the basis of a partial least squared (PLS) regression method. The PLS method used in Fig. 1.2B utilizes the range of $1400-1800\ cm^{-1}$ to quantify protein concentration of an immunoaffinity resin accurately with a confidence interval of 95% [13].

1.1.2.3 Challenges—interference of water spectral bands

It is well documented that protein analysis by FTIR spectroscopy is limited by the strong absorption of IR radiation by water, particularly for measurements in transmission. ATR-FTIR spectroscopy probes only a few micrometers of a sample that is directly in

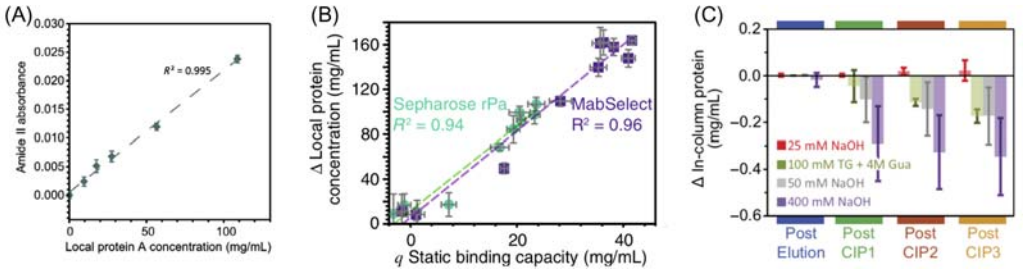


Figure 1.2

(A) Standard addition curve used for protein A local concentration quantification. (B) Difference in local protein concentration measured by attenuated total reflection (ATR)—Fourier transform infrared (FTIR) spectroscopy as a function of the static binding capacity (q). (C) In-column protein concentration difference calculated from ATR-FTIR-based partial least squared (PLS) regression (*red*) after elution and cleaning in place (CIP) for different conditions. Source: *Reproduced with permission from Springer Nature (M. Boulet-Audet, B. Byrne, S.G. Kazarian, Cleaning-in-place of immunoaffinity resins monitored by in situ ATR-FTIR spectroscopy, Anal. Bioanal. Chem. 407 (2015) 7111—7122; M. Boulet-Audet, S.G. Kazarian, B. Byrne, In-column ATR-FTIR spectroscopy to monitor affinity chromatography purification of monoclonal antibodies, Sci. Rep. 6 (2016) 30526).*

contact with the ATR crystal, removing the issue of complete absorption of the IR radiation by the aqueous phase of a sample. This also provides an additional advantage as usually all spectral bands in ATR-FTIR spectroscopy are on-scale, the exception to this is when accessories with many internal reflections are used.

However, water remains a challenge in spectral analysis even after removal of water saturation, since the amide I band often overlaps with the water bending mode ($\approx 1650\ \text{cm}^{-1}$). This can cause issues such as oversubtraction and thus the loss of key structural features of a desired protein analyte. A solution to this issue is to use deuterated water to shift the bending band away from the amide I band. However this causes the generated spectrum to be unrepresentative of the native protein structure because of isotope exchange; structural changes under these conditions are indicated by a shift in the amide II band from $\approx 1550\ \text{cm}^{-1}$ to $\approx 1450\ \text{cm}^{-1}$ [10]. To ensure correct subtraction of the background buffer spectrum from the sample spectrum specific criteria must be met. The subtraction coefficient should be determined by using a least squares method on the combination band of water ($2150\ \text{cm}^{-1}$) in protein samples to remove variations of human bias in the subtraction of water over multiple samples. In addition, the ratio of amide I to amide II absorbance should not be higher than 1.5 in a measured spectrum of a protein sample [14].

1.1.2.4 Comparison between transmission and ATR spectroscopic analysis of proteins

Transmission spectra are generated by passing infrared light through a sample typically held between an IR transparent window such as CaF_2. A significant disadvantage of the

transmission method when measuring solutions of proteins is that proteins are in an aqueous phase. As previously mentioned water strongly absorbs IR radiation, thus only a very small path length ($\approx 8\ \mu m$) can be used to minimize the risk of completely absorbing all the IR radiation being passed through the sample before it reaches the detector [15]. Similarly, longer path lengths through the solution of protein result in spectral bands being off-scale, making the spectroscopic analysis very challenging or even impossible.

Employing ATR rather than transmission methods is preferable when measuring aqueous samples due to the interaction of the reflected IR radiation with only the surface layer of a sample ($0.2-3\ \mu m$) which is in direct contact with the ATR element, ensuring that water does not saturate the spectrum. ATR methodologies require minimal sample preparation when compared to transmission mode as samples can simply be placed on the ATR element for measurement. The evanescent wave probing of the surface layer of the sample and the technique's ability to be combined with other methods to alter protein structure such as changing pH, temperature, and surface properties make this technique ideal for characterization of proteins [1,4].

However, ATR methodologies are less sensitive than transmission methodologies because of the smaller path length used, resulting in a lower signal to noise ratio, making it harder to discern spectral noise from actual vibrational modes present. As a minimum, spectral bands should be three times stronger in absorbance than that of spectrum noise [16]; this can limit the data that is acceptable to be interpreted for a given spectrum [15].

1.1.3 The significance of study of protein crystallization and aggregation with new vibrational spectroscopic methods

A particularly useful advance to the analysis of proteins using IR spectroscopy has been the integration of a focal plane array (FPA) to FTIR spectroscopic measurements. The use of a FPA allows thousands of spectra to be simultaneously measured in a timescale of seconds without sample preparation methods such as the addition of protein stains [1,4]. This spectroscopic imaging method, an inherently high-throughput approach, allows for the drawbacks of low-throughput methods such as traditional point to point mapping of samples to be overcome and for real-time monitoring of structural changes in proteins to be carried out [2,3,12].

The FPA is a multichannel detector that creates images detailing high absorbance and low absorbance of corresponding vibrational modes, for example, those that are used to determine protein structure such as amide I and II bands [16]. This generation of a chemically specific "Heat map" allows multiple components in a sample to be analyzed simultaneously [1,12,17].

The ability of the FPA to differentiate substances within a multicomponent system is shown in Fig. 1.3. The plotting distribution of integrated absorbance of chemically specific bands (i.e., amide II) allows chemical images to be generated as a nondestructive method of assessing a particular protein crystallization condition and to help identify whether crystals formed for X-ray crystallography are protein or salt [2,4]

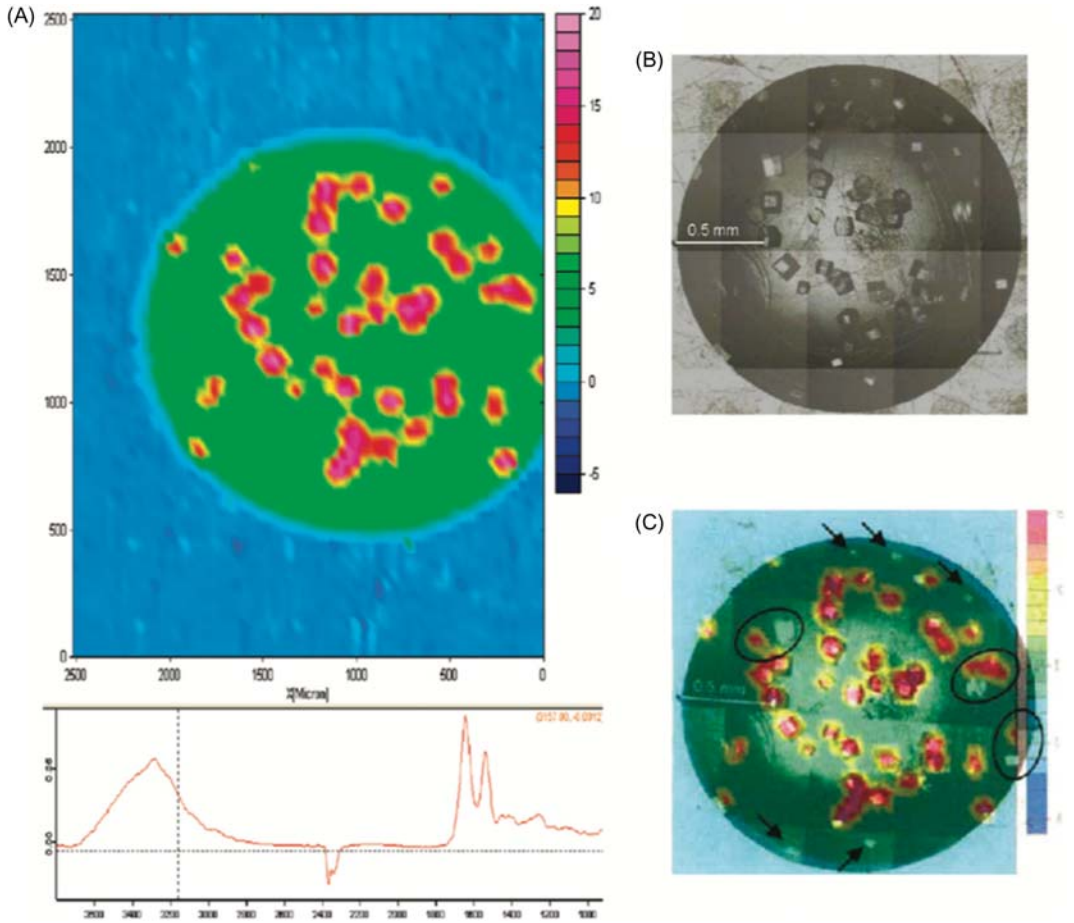


Figure 1.3

Attenuated total reflection (ATR)—Fourier transform infrared (FTIR) image (the imaged area is approximately 2.5×3.6 mm^2) (A) of the protein crystal formed on the measuring surface 20 h after introduction of the crystallizing agent. Red represents high protein concentration. The optical microscopy image (B) shows the protein crystals formed coincide well with the location of the red spots on the ATR-FTIR image. (C) Result of overlaying images (A) and (B). Five out of 50 crystals (highlighted with *arrows*) were not captured in the ATR-FTIR image, and three crystals (*circled*) were misallocated. Source*: Reproduced with permission from K.L.A. Chan, et al., Attenuated total reflection-FT-IR spectroscopic imaging of protein crystallization, Anal. Chem. 81 (10) (2009) 3769–3775. Copyright 2009 American Chemical Society.*

1.2 ATR-FTIR spectroscopic imaging of proteins

1.2.1 Macro-ATR spectroscopic imaging

1.2.1.1 High-throughput measurements: protein crystallization growth, aggregation, study of protein adsorption by functionalizing ATR crystal surfaces

X-ray crystallography remains the most common method of protein structure determination. At the time of writing this technique is responsible for ~90% of all the structures deposited in the protein databank (www.rcsb.pdb.org). However, determination of structures using X-ray crystallography requires the formation of protein crystals, a process dependent on extensive and time-consuming screening of crystallization conditions requiring large amounts of protein [18]. Any crystals obtained have to be confirmed as protein and tested for ability to diffract X-rays.

The most widely used crystallization technique is vapor diffusion using either a hanging drop or sitting drop setup involving equilibration of a droplet of protein and precipitant solution with a reservoir containing a higher concentration of precipitant [19]. During equilibration the protein concentration in the droplet increases and, if conditions are correct, nucleation and ultimately crystallization occur. Other techniques, including microbatch under oil [20] and lipidic cubic phase [21], have been used successfully. Given the comparatively stochastic nature of the crystallization process, structure determination of proteins using this approach remains a very inefficient process. Methods to improve control of the crystallization process would be of great benefit to the community.

A macro-ATR-FTIR spectroscopic imaging approach [2] has been used to study protein crystallization under oil with either single or multiple crystallization droplets applied directly to a ZnSe ATR accessory. Using this setup crystals ($>40\ \mu m$ in size) of lysozyme and thaumatin formed within 1 day and were clearly detectable as ATR-FTIR chemical images which corresponded to crystals visible under a light microscope (Fig. 1.3B). Higher spatial resolution was possible when using a diamond ATR accessory [22] allowing detection of very small crystals ($12\ \mu m$). Using this approach an integral membrane protein was submitted to crystallization trials and although crystals visible under a light microscope were obtained, ATR-FTIR imaging revealed that these were nonproteinaceous.

Currently there are a number of methods to determine if a crystal is protein. Testing the crystals in an X-ray beam can be performed, however, this requires crystal manipulation and is wasteful of beam time. Protein crystals typically fluoresce strongly under a UV microscope and such an approach is widely used, however, it is not completely reliable in terms of differentiation of protein and nonprotein crystals [23]. Dye-based fluorescence methods have also been reported [24]. Thus ATR-FTIR chemical imaging, which is label-free, nondestructive, and able to very reliably differentiate between protein and nonprotein crystals, represents a viable alternative to these approaches. The Chan et al. study [2]

described above focused on crystallization under oil which, while used by researchers in the field, is not one of the most common crystallization techniques.

Glassford and colleagues [3] further applied micro-ATR-FTIR chemical imaging, a technique with higher spatial resolution than macro-ATR-FTIR, to study the crystallization of a number of proteins—in this case, the much more widely used hanging drop approach with the setup shown in Fig. 1.4A with drops of protein solution placed directly on the removable germanium crystal and the reservoir placed below. Using this setup it was possible to both image lysozyme crystals and observe them as they grew (Fig. 1.4B).

The small area measured in micro-ATR-FTIR imaging means that it might not be possible to always capture individual nucleations and follow crystal growth. However, because of the high spatial resolution of micro-ATR-FTIR imaging, it is possible to detect crystals of thaumatin and lobster α-crustacyanin as small as 6 μm, much smaller than those detectable using macro-ATR-FTIR spectroscopic imaging [2].

Clearly many factors can affect protein crystallization including the physical properties of the vessel that is being used. Another study utilized ATR-FTIR imaging with a Si ATR element to explore the effect that surface gradient property has on protein crystallization [4].

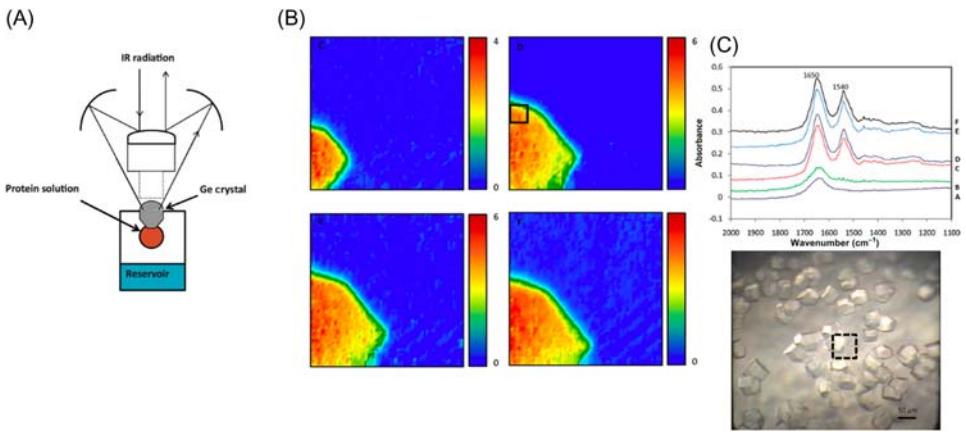


Figure 1.4

(A) Micro-attenuated total reflection (ATR)—Fourier transform infrared (FTIR) imaging setup for detection of protein crystals in a hanging drop. (B) Micro-ATR images of a lysozyme crystal obtained after 80, 110, and 140 min, and 20 h [3] indicated by letters C, D, E and F respectively. (C) Spectra for these times as well as 0 and 40 min indicated by the letters A and B respectively. An additional light microscopic image shows typical lysozyme crystals with the approximate imaging area indicated by the dashed box. Source*: Reproduced from S.E. Glassford, et al., Micro ATR FTIR imaging of hanging drop protein crystallisation, Vibrat. Spectrosc. 63 (2012) 492–498 with permission from Elsevier.*

A gradient of octyltrichlorosilane (OTS), a hydrophobic molecule, was generated on the surface of the ATR element which had been coated with a thin layer of gold (Fig. 1.5A). Gradient formation was monitored by ATR-FTIR spectroscopic imaging (Fig. 1.5B).

Using this setup it was possible to grow lysozyme crystals on the Si ATR element and ATR-FTIR imaging revealed that there was a clear preference for crystal growth on the region of the ATR element with higher concentrations of octyltrichlorosilane and thus on the more hydrophobic regions of the surface (Fig. 1.5C). Although it should be noted that fewer crystals grew on the least hydrophobic areas of the ATR element, those crystals that did grow were of similar dimensions to those that grew at higher density on the more hydrophobic regions of the ATR element. This study only explored crystallization of one protein on this hydrophobic gradient. Clearly it will be interesting to also assess the effects of this surface gradient on crystallization of other proteins including, for example, integral membrane proteins. Such gradient surface properties also have the potential for facilitating adhesion and study of live cells.

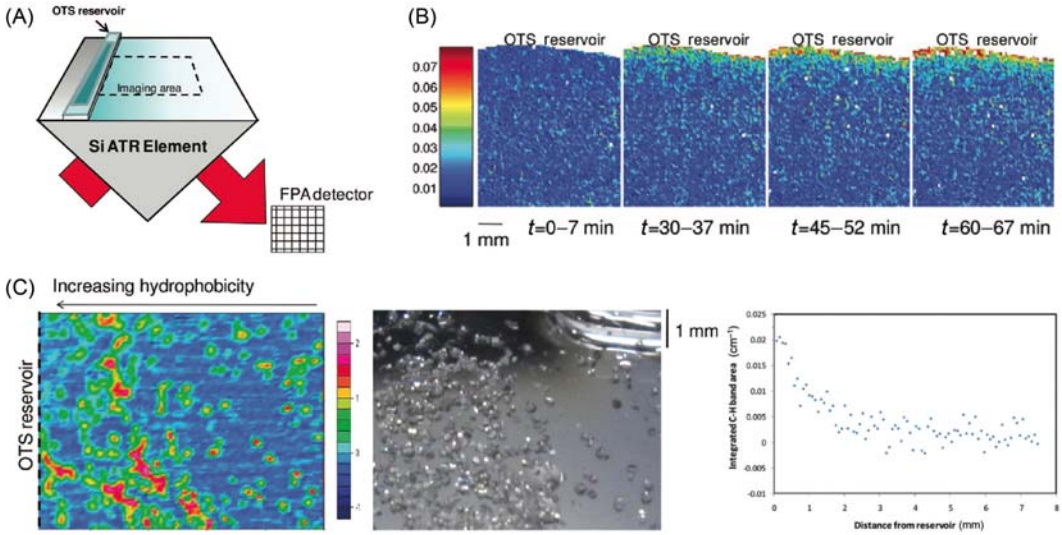


Figure 1.5

(A) Overall experimental setup used for the study of gradient surface property on protein crystallization. FPA = focal plane array detector used to measure thousands of spectra in a short period of time. (B) Generation of the OTS gradient on the surface of the attenuated total reflection (ATR) element as assessed by ATR–Fourier transform infrared (FTIR) imaging with images taken at different time points as indicated. The OTS reservoir is at the top of the images. (C) ATR-FTIR imaging of the lysozyme protein crystals which grew on the ATR element with the corresponding light microscopic image. Source: *Reproduced with permission from S. Glassford, et al., Chemical imaging of protein adsorption and crystallization on a wettability gradient surface, Langmuir, 28 (6) (2012) 3174–3179. Copyright 2012 American Chemical Society.*

1.2.1.2 Eliminating anomalous dispersion with varying angle-macro-ATR

Anomalous dispersion, which poses a challenge for quantitative and qualitative analysis when present in a spectrum, is often observed in the ATR-FTIR measurement when the refractive index of the sample is close to the wavelength of an absorbance band. This dispersion artifact manifests itself as a shift in the strong absorbing spectral bands and could exhibit a derivative-like baseline shape near the bands [25].

The influence of anomalous dispersion on the shifting of the amide bands was reported by Boulet-Audet et al. in a study on protein film conducted on an ATR accessory. It should be noted that the band shift was significantly reduced when a germanium ATR element was used instead of a diamond because of its higher refractive index [26]. Since a shift in the amide I band has serious implications for interpretation of the secondary structure of the protein it is necessary to correct for this. The researchers accounted for the band distortion resulting from dispersion by quantification through determination of the optical constant using ATR [26]. However, Kazarian and coworkers also found that the spectral distortion could be minimized simply by increasing the angle of incidence of the IR beam using specially made apertures [27,28]. This is particularly useful in cases where a diamond ATR element is preferable, for example, where chemical inertness and high resilience to pressure of the ATR element are required.

The use of apertures to physically alter the angle of incidence of the IR beam for ATR-FTIR imaging with a diamond Golden Gate accessory has successfully shown that distortion of spectral bands can be effectively minimized by increasing the angle of incidence [27,28]. The apertures are designed to have various positions that can be fitted onto the condenser lens of the diamond imaging optics, allowing light to pass through at a specific angle of incidence [28].

Comparison of the lysozyme spectrum obtained using a diamond ATR at different angles of incidence reveals that at 37 degrees, the smallest angle of incidence, the amide I peak is at 1636 cm^{-1} compared to 1645 cm^{-1} at 45 degrees (Fig. 1.6) [15]. Similarly, the amide II peak also shifted from 1515 to 1535 cm^{-1} with increasing angle of incidence. Both amide bands shifted toward higher wave number as the angle of incidence increased (Fig. 1.6).

It is not possible to completely remove all the spectral distortion due to the limitation in the range of allowed angles of incidence possible with the arrangement and alignment of the Golden Gate IRE. However, as the angle of incidence increases, absorbance of the amide bands is observed. This is because of the shallower depth of penetration of the evanescent wave at a greater angle of incidence of the IR beam. Due to the improved performance offered by variable angle ATR-FTIR measurement, it has been employed extensively for investigation of absorbance change in the z-direction, also known as depth-profiling [29–31], which has the potential to allow study of aspects of protein behavior including protein aggregation.

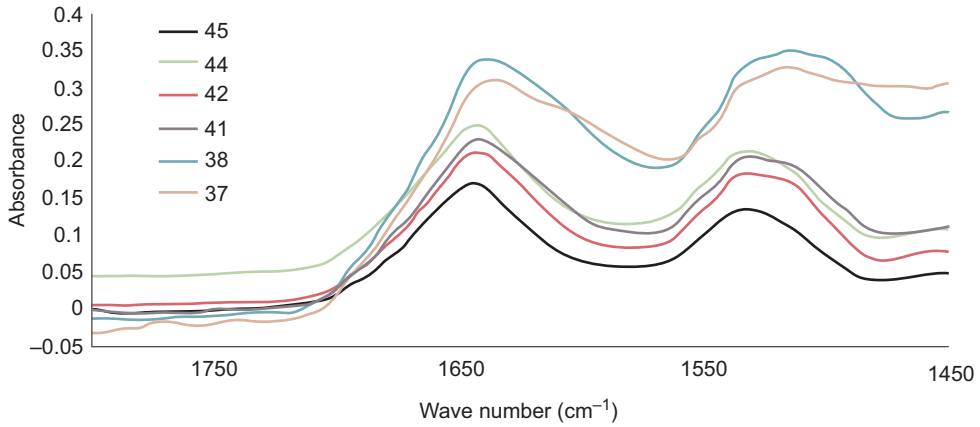


Figure 1.6
Attenuated total reflection (ATR)—Fourier transform infrared (FTIR) spectra of lysozyme film at various angles of incidence. Source: *Reproduced from S.E. Glassford, Applications of ATR-FTIR spectroscopic imaging to proteins, in: Chemical Engineering, Imperial College London, London, 2013, p. 158.*

1.2.1.3 High-throughput analysis of aggregation of a monoclonal antibody by macro-ATR-FTIR spectroscopic imaging

Protein aggregation can have major impact on the use of biotherapeutics including monoclonal antibodies (mAbs). Aggregation of a therapeutic mAb is associated with reduced efficacy and can lead to highly undesirable immunological responses when administered [32,33]. It is thus important to have an effective means of assessing the propensity of a protein sample to aggregate. There are a number of methods for studying protein aggregation including size-exclusion chromatography (SEC), dynamic light scattering, and circular dichroism [34], however, many of the available methods are unable to observe changes in aggregation in protein samples in situ, or are limited in their ability to quantify aggregation or limited in the range of conditions which can be tested. Protein aggregation of purified protein in vitro can be effectively analyzed using macro-ATR-FTIR spectroscopic imaging. Boulet-Audet et al. utilized a high-throughput approach to study the aggregation of mAbs in a range of different pH and ionic strength conditions, including the conditions experienced by mAbs during the isolation process [1]. In this case PDMS wells were cast allowing 12 samples to be analyzed simultaneously. The wells were mounted on to a ZnSe ATR crystal, which was used in preference to a germanium crystal due to the greater absorbance of the bands produced. Data analysis was performed in protein samples in a range of buffer conditions while the protein was being heated. Heating is commonly used as a stress to induce aggregation in a timescale more suited to laboratory experiments. The formation, upon heating, of aggregates and/or precipitates which are heavier than the monodispersed protein in the bulk solution results in a larger amount of protein being in contact with the ATR crystal (Fig. 1.7). Thus the relative stabilities of the protein in the

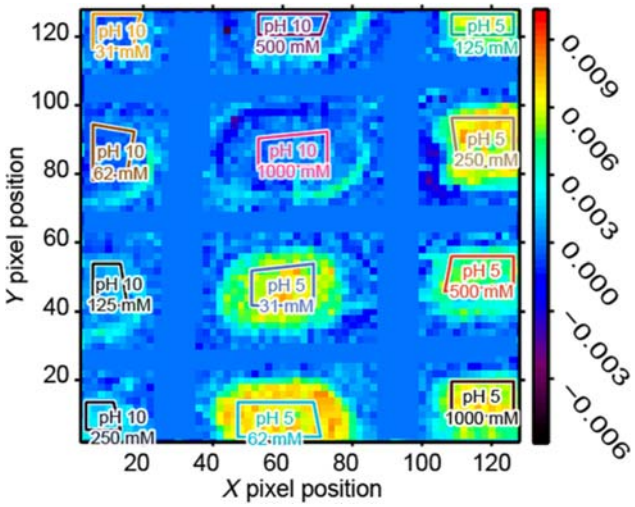


Figure 1.7

ATR-FTIR spectroscopic images generated using the amide II band area after 2000 s. The unit of the scale bar is integrated absorbance in cm^{-1}. The polygons represent the integrated region of interest of each well. The field of view covered approximately 9.8 mm along the *Y* axis and 7.0 mm along the *X* axis. Source: *Reproduced with permission from M. Boulet-Audet, et al. High-throughput thermal stability analysis of a monoclonal antibody by attenuated total reflection FT-IR spectroscopic imaging. Anal. Chem. 86 (19) (2014) 9786-9793. Copyright 2014 American Chemical Society.*

different conditions can be readily compared. However, the real power in ATR-FTIR spectroscopy is that in addition to protein stability analysis, the technique also reveals key changes in the secondary structure of the mAb that lead to aggregation in a high-throughput manner over a period of about 30 minutes [1] (Fig. 1.8).

1.2.1.4 Protein purification: cleaning-in-place for immunoaffinity resin and in-column ATR-FTIR spectroscopy

Another field in which ATR-FTIR spectroscopy has obtained novel insights is protein purification. mAbs, proteins widely employed for treatment of a range of cancers and chronic diseases such as rheumatoid arthritis, constitute a fast-growing area of biopharmaceuticals [35]. However, the high cost associated with their production has a major impact on the production scale, profitability, and, ultimately, patient availability of mAbs [36]. Antibody purification is a key part in the mAb production pipeline and involves an ultrafiltration step to concentrate the tissue culture media containing the mAbs and a number of chromatography steps. The most effective purification step employs affinity chromatography, utilizing the very specific interaction between mAbs and protein A [37] at pH 7, with the mAbs eluting in buffer at pH 3. This protein A affinity chromatography step

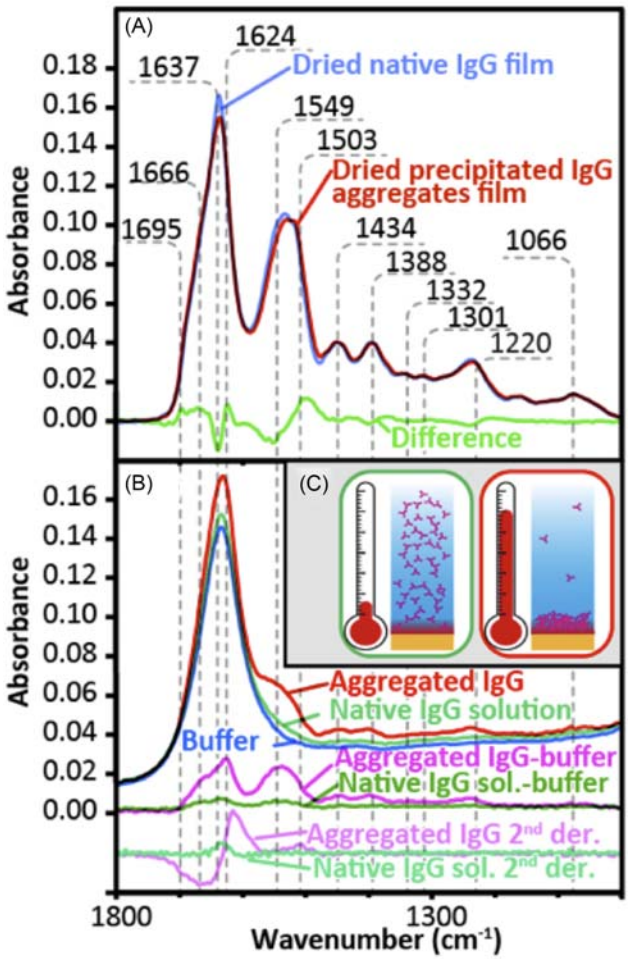


Figure 1.8

(A) Attenuated total reflection (ATR)—Fourier transform infrared (FTIR) spectra of a dried film of native IgG from a demineralized solution (*blue*), a dried film from thermally induced aggregated IgG solution (*red*), and the difference between them (*green*). The dried native film was normalized for comparison using the average absorbance between 1300 and 1350 cm^{-1}. (B) ATR-FTIR spectra of buffer only (*blue*), 1 mg/mL native IgG solution (*light green*), precipitated aggregates induced by heating of a 1 mg/mL IgG solution (*red*), difference between the native IgG solution and the buffer (*dark green*), its second derivative (*teal*), as well as the difference between the aggregated IgG and the buffer (*fuchsia*) and its second derivative (*pink*). (C) Schematics illustrating the native IgG (*green box*) and the increased local concentration in the volume probed by ATR-FTIR spectroscopy caused by heat-induced protein aggregates precipitation (*red box*).
Source*: Reproduced with permission from M. Boulet-Audet, et al. High-throughput thermal stability analysis of a monoclonal antibody by attenuated total reflection FT-IR spectroscopic imaging. Anal. Chem. 86 (19) (2014) 9786—9793. Copyright 2014 American Chemical Society.*

is a major contributor to the cost of mAb production, since the resin used accounts for more than 50% of the total production expense [36,38]. A particular problem is the relatively low life span of the resin, thought to be due to nonspecific binding of contaminants and loss of the protein A ligand. Such contaminants can block pores or hinder the access to surface ligands. To prevent contaminant buildup, washing the resin with alkaline solutions is often adopted as part of the resin regeneration protocol [cleaning in place (CIP)]; these protocols increase resin life span, however, harsh regeneration conditions also contribute to degradation or leaching of the ligands over time. Changes in the binding capacity of the resin can be assessed in a number of ways but most of these are indirect. A mechanistic approach consists of measurement of Langmuir adsorption isotherms [38], with other methods relying on sodium dodecyl sulfate polyacrylamide gel electrophoresis and analytical SEC in an attempt to analyze eluate samples for the presence of leaked protein A and host cell protein content [39].

A better understanding of the issues associated with reduction in binding capacity through more direct ATR-FTIR spectroscopic assessment of on-column changes during both protein isolation steps and the CIP procedures used to remove bound contaminants is still needed.

Since ATR is a surface-sensitive technique that relies on close contact between sample (e.g., resin agarose beads) and internal reflection element (IRE, e.g., diamond), the beads were pressed onto the IRE under a controlled load. A plunger was used to exert pressure and a load cell was employed to precisely control the load applied. An optimal load maintains effective contact between the beads and the IRE without damaging the beads or affecting flow across the IRE [12]. Fig. 1.9A and B show schematics of the experimental setup. Using this setup it was possible to detect the protein A ligand on the resin beads and mAb binding, as well as changes in protein conformation and proteolysis caused by addition of a high concentration of NaOH (Fig. 1.9C) [12].

The resin is usually washed with alkaline solutions according to one of a number of established CIP protocols to prevent buildup of contaminants and thus a decrease in binding capacity. The ATR-FTIR spectroscopic approach described above allowed comparative analysis of a range of different CIP protocols to assess both their relative effectiveness at removing contaminants and their effects on the protein A ligand [12].

More recent studies explored mAb purification under dynamic flow during protein A affinity chromatography by direct monitoring using in-column ATR-FTIR spectroscopy whilst also monitoring eluates by standard UV spectroscopy [13]. This utilized a modified version of the setup shown in Fig. 1.9A and B, allowing controlled and dynamic alteration of the buffer surrounding the beads on the IRE, allowing mAb binding, elution, and CIP cycles to be continuously monitored [13]. ATR-FTIR spectra revealed that protein A leaches from the column during CIP and that even after CIP, contaminants remain bound to the column reducing the binding capacity of the resin [13]. The types of contaminants

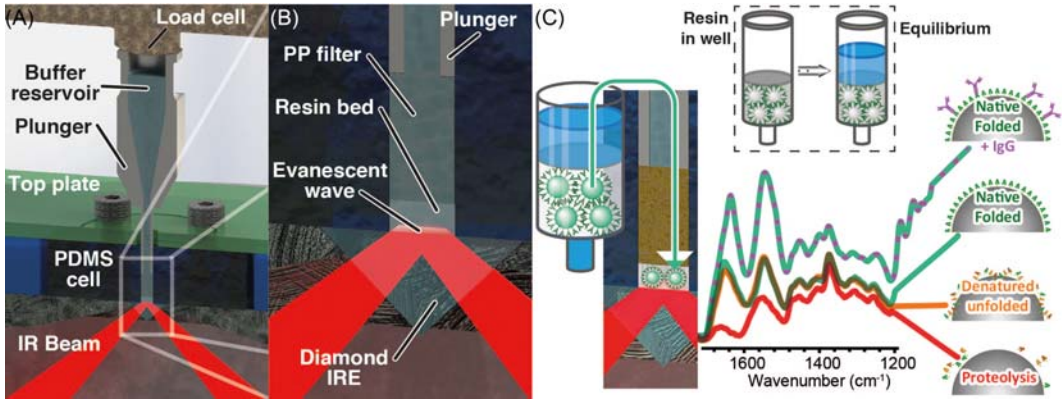


Figure 1.9

Experimental setup utilized for the analysis of protein A resin with ATR-FTIR spectroscopy and imaging (A). Zoomed in view of the setup (B); a PDMS cell is clamped to the ATR accessory using an acrylic top plate and to exert pressure on the resin beads a plunger is utilized. To measure the force applied, a load cell is placed on top of the plunger. (C) Resin in equilibrium is injected on the IRE surface and pressed under controlled load; ATR-FTIR spectra of resin beads with/without mAb (IgG) are measured as a function of denaturant concentration (NaOH). Source*: Reproduced with permission from Springer Nature (M. Boulet-Audet, B. Byrne, S.G. Kazarian, Cleaning-in-place of immunoaffinity resins monitored by in situ ATR-FTIR spectroscopy, Anal. Bioanal. Chem. 407 (2015) 7111–7122. M. Boulet-Audet, S.G. Kazarian, B. Byrne, In-column ATR-FTIR spectroscopy to monitor affinity chromatography purification of monoclonal antibodies, Sci. Rep. 6 (2016) 30526).*

bound to protein A columns have been a key unknown for a very long time and this has limited the development of novel protocols and approaches to reduce contaminant binding in industrial settings. One major advantage to ATR-FTIR spectroscopy is that since each species in a sample has a distinct chemical fingerprint it is possible to discern what types of molecules are in a particular sample. The analysis of the resin beads indicated that no lipid or DNA was detectable, thus it is likely that most if not all the contaminant material is proteinaceous in nature [13].

Further optimization of the purification and CIP protocols are possible using this type of setup, since even small improvements in the life span of the protein A columns has potential to significantly reduce the cost of industrial scale production of therapeutic mAbs.

1.2.2 Micro-FTIR spectroscopic imaging

1.2.2.1 Association with disease: time-resolved imaging of protein aggregation in living cells

FTIR microspectroscopy has been used to study protein aggregates in tissues and cells in vivo, in vitro, and in situ to better understand a wide range of protein-folding diseases.

Protein aggregates are usually detected based on the presence of a specific marker band, such as the amide I band or amide II band, as a result of the formation of intermolecular β-sheet structures visible in the FTIR spectrum [40].

A study was conducted by Mitri et al., whereby the response of live mammary breast adenocarcinoma cell lines to severe heat-shock was monitored and recorded using FTIR microspectroscopy. As a result of the stress induced on the live cells, persistent intracellular accumulation of extended β-sheet protein aggregates was detected [41]. Thus the capability of the IR microscopy for real-time measurement of in-depth variation of the cellular biochemical profile was confirmed. This IR microspectroscopic approach has also been applied to explore the kinetics of the formation of inclusion bodies within growing *Escherichia coli* cells under various expression conditions [42]. In this case, the spectroscopic analysis suggested that the recombinant protein is more prone to aggregate at 37°C. Interestingly, it is also possible to obtain complementary information on cell processes that accompany protein aggregation, including, for example, the effects on cell membranes from further analysis of the IR spectra obtained, such as the effects of recombinant protein misfolding and aggregation on bacterial membranes [43].

A further study by Ami et al. used an FTIR microscope to screen cardiac tissues for the presence of characteristic amyloid deposits [44]. These results demonstrated that tissue analysis by FTIR spectroscopy not only aids in diagnosis, but could also detect the presence of fibrillar aggregates in clinical specimens. Futhermore, this approach can also provide meaningful spectra allowing analysis of aggregate core properties, including the strength of the hydrogen bonds, which affect the compactness of aggregate structure and thus may affect their functions.

The challenge of examining small aggregates can be overcome by the use of a synchrotron-based IR microscope. One of the earliest uses of this approach investigated the secondary structure of Aβ protein directly within the brain tissue samples of a deceased subject suffering from Alzheimer's [45]. They found that the spectra obtained in vivo and in situ are able to associate protein behavior with the surrounding environment. Since then, synchrotron-based FTIR study in Alzheimer's disease has also revealed that the density of dense-core amyloid plaques is approximately 1.5 times higher than the surrounding brain tissue (Fig. 1.10) [46].

More recently, time-lapse infrared imaging has been applied to aggregate formation in a cell culture model of amyotrophic lateral sclerosis (ALS). ALS is a neurodegenerative disease that causes death of neurons controlling voluntary actions and is associated with mutations in the antioxidant protein copper−zinc superoxide dismutase (SOD1) [47]. With the ability of FTIR spectroscopy to distinguish between parallel and antiparallel β-sheet structure, real-time imaging of SOD1 aggregation in the cell culture model provided insight into structural intermediates, timescale, and mechanisms of cell toxicity triggered by aggregation.

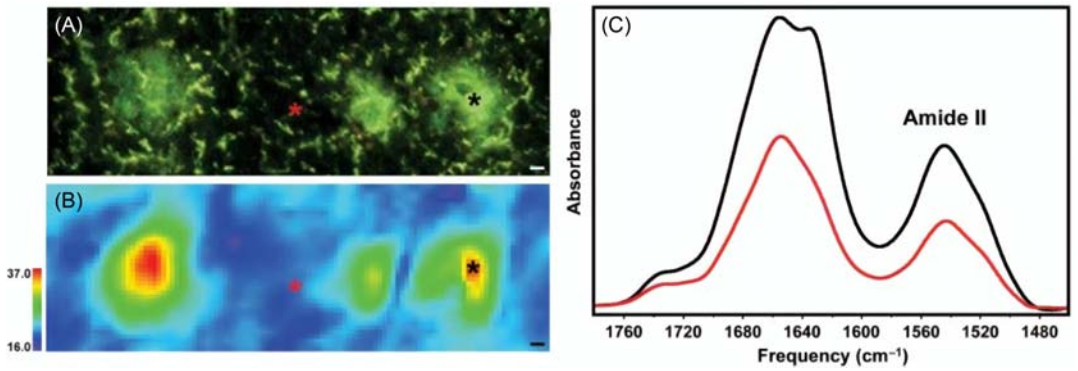


Figure 1.10

(A) Thioflavin S-stained PSAPP mouse brain tissue showing three plaques. (B) Infrared image of the same tissue showing the distribution of protein measured by the Amide II band. (C) Infrared spectra collected from the areas marked with asterisks in (A) and (B), showing the relative amount of protein in the center of a plaque (*black*) and the surrounding tissue (*red*). All scale bars are 5 μm [46]. Source*: Reproduced with permission from Elsevier.*

1.3 Further applications

1.3.1 Monitoring low-concentration protein conformational change with QCL spectroscopy; potential of micro-ATR-FTIR imaging for analysis of tissues

This chapter has highlighted ATR-FTIR spectroscopic imaging as a powerful and versatile technique for the study of protein crystallization and aggregation using isolated protein samples, as well as protein aggregation in in vivo tissues and cell samples and protein behavior on a chromatographic column. Future developments in this field will only enhance the potential of infrared imaging in enabling the further understanding of protein behavior. One of the limitations of ATR-FTIR spectroscopy and spectroscopic imaging is relatively low sensitivity of detection because of the relatively short path lengths used, particularly for imaging where a single reflection IRE is employed. This limitation can be overcome through the use of much more powerful infrared sources, such as a quantum cascade laser (QCL), rather than the thermal sources used in conventional spectrometers. QCLs, which have been applied to study tissues, are tunable laser sources which output discrete frequencies of infrared light brighter than a synchrotron but small enough for benchtop instruments [48].

With conventional ATR-FTIR spectroscopy protein of low concentration may be undetectable; however, QCL-based infrared spectroscopy has been demonstrated to measure protein at a concentration as low as 0.25 mg/mL in transmission mode [49]. The observed results of the novel setup agree with those obtained from conventional infrared

spectroscopic measurement. The robust measurement of protein at low concentrations is made possible with QCL because such sources produce mid-infrared light several orders of magnitude brighter than that from a silicon carbide heating element or synchrotron. The most notable recent studies on protein samples have utilized QCL infrared spectroscopy to capture dynamic changes in the formation of β-sheet aggregates at varying pH values and protein concentrations [50], and the screening of bovine milk [51]. In addition, a time-resolved study of enzymatic reactions was accomplished with a novel dual-comb QCL system [52]. In the future we expect the adoption of QCLs will allow the application of ATR-FTIR spectroscopy and spectroscopic imaging to an increasingly wide range of protein studies.

As more work progresses in this field, improved in-depth understanding of protein structure is inevitable. Not only the secondary structure of proteins can be explored, but it is envisioned that nano-FTIR spectroscopy will provide insights into the quaternary structure of proteins. This type of spectroscopy circumvents the diffraction-limited resolution of FTIR microscopy by focusing infrared light onto a metalized AFM tip, similar to tip-enhanced Raman scattering spectroscopy. Infrared images of nanoscale spatial resolution as low as 20 nm can be recorded. Amenabar et al. [53] applied nano-FTIR to study 3-nm-thin individual insulin fibrils, a widely used model system for neurodegenerative disorders, including Alzheimer's and Parkinson's disease. Their results not only demonstrated that nano-FTIR has a high sensitivity to differentiate individual protein complexes, but also revealed that the presence of α-helices in the shell of insulin fibrils is likely to contribute to aggregation of the fibrils [53]. Protein microarrays are pushing the boundaries of high-throughput analysis of proteins. Coupled with IR microscopic imaging, the sensitivity of the microarrays was tested with lysozyme, albumin, and hemoglobin solutions in a total of 45 spots, each with an approximate diameter of 100 μm, recorded simultaneously over a few minutes. Despite the minute amount of protein in each spot, high-quality spectra were obtained from single monolayers of the different protein with FTIR microspectroscopy. These findings demonstrate that it is possible for IR imaging to be employed for the detection of protein binding to other proteins, substrates, etc. [54]. In future, there are opportunities for still further development of the technology in terms of the methodology used for capture and analysis of spectra, as well as improvements in instrumentation which will increase both sensitivity and the breadth of investigations that can be carried out on proteins. Additionally, future developments in standard protocol and chemometric methods may render spectral analysis of proteins quicker and easier.

Finally, a capability, developed in the authors' laboratory [30,31,55], of ATR-FTIR spectroscopic imaging with variable angle of incidence to obtain a depth profile of samples may prove to be useful for assessment of processes at the surface layers of the resins in stationary chromatographic columns. This depth profiling capability with the

sub-micrometer spatial resolution (the spatial resolution in z-direction when depth profiling is not limited by diffraction of light, unlike the resolution in the x and y directions) reveals chemical information by studying the variation in a sample's composition as a function of depth in a nondestructive way.

1.4 Conclusions

The research presented in this chapter summarizes the applicability of ATR-FTIR spectroscopic imaging and ATR-FTIR spectroscopy to a range of different protein systems and presents new opportunities for applications of ATR-FTIR imaging to challenges in this field. Overall, spectroscopic imaging capability could have benefits for research in structural proteomics and contribute to enhancement of crystallization conditions and optimization of protein isolation protocols. In particular, this powerful analytical and research tool has significant potential for the study of a range of different proteins and protein complexes in addition to important biopharmaceutical and bioprocessing targets. This chapter presented novel approaches, such as macro-ATR-FTIR imaging, for high-throughput protein analysis of both protein crystallization and aggregation; this approach is also well-suited to study the effects of wettability gradient surfaces on protein adsorption and crystallization. Micro-ATR-FTIR imaging can be used to study hanging drop protein crystallization with high spatial resolution.

A significant breakthrough has been the application of ATR-FTIR spectroscopy to the study of loss in protein A resin binding capacity. This approach allowed monitoring of both ligand conformation and the amount of adsorbed protein on the beads and as a result demonstrated ligand denaturation and proteolysis caused by CIP protocols and irreversible protein contaminant binding, as described in this chapter.

The importance of the understanding of resin aging in stationary chromatography columns was recently highlighted in a discussion in *Trends in Biotechnology* [56]. It has been suggested that it is necessary to monitor the events that result in efficiency loss in chromatography columns. At present there is no established procedure for this as an industry standard [56]. Therefore the advanced FTIR spectroscopic approaches [57], discussed in this chapter, to monitor column aging during operation may contribute to the understanding of column aging and help to establish procedures in industry pipelines.

Acknowledgments

The authors wish to thank the BBRSC and the Bioprocessing Research Industry Club (BRIC) for funding this research (BB/K0111030/1 and BB/R019533/1). We also thank BBSRC for BBSRC Targeted Priority Studentships in the area of Bioprocessing and for iCASE Studentships.

References

[1] M. Boulet-Audet, B. Byrne, S.G. Kazarian, High-throughput thermal stability analysis of a monoclonal antibody by attenuated total reflection FT-IR spectroscopic imaging, Anal. Chem. 86 (19) (2014) 9786–9793.

[2] K.L.A. Chan, et al., Attenuated total reflection-FT-IR spectroscopic imaging of protein crystallization, Anal. Chem. 81 (10) (2009) 3769–3775.

[3] S.E. Glassford, et al., Micro ATR FTIR imaging of hanging drop protein crystallisation, Vibrat. Spectrosc. 63 (2012) 492–498.

[4] S. Glassford, et al., Chemical imaging of protein adsorption and crystallization on a wettability gradient surface, Langmuir 28 (6) (2012) 3174–3179.

[5] S.E. Glassford, B. Byrne, S.G. Kazarian, Recent applications of ATR FTIR spectroscopy and imaging to proteins, Biochim. Biophys. Acta 1834 (12) (2013) 2849–2858.

[6] N.J. Harrick, Variable angle attachment for internal reflection spectroscopy, Anal. Chem. 37 (11) (1965) 1445.

[7] N.J. Harrick, K.H. Beckmann, Internal reflection spectroscopy, in: P.F. Kane, G.B. Larrabee (Eds.), Characterization of Solid Surfaces, Springer US, Boston, MA, 1974, pp. 215–245.

[8] K. Ohta, R. Iwamoto, Experimental proof of the relation between thickness of the probed surface-layer and absorbance in FT-IR ATR spectroscopy, Appl. Spectrosc. 39 (3) (1985) 418–425.

[9] S.G. Kazarian, K.L. Chan, Micro- and macro-attenuated total reflection Fourier transform infrared spectroscopic imaging, Appl. Spectrosc. 64 (5) (2010) 135A–152A.

[10] B.H. Stuart, Infrared Spectroscopy: Fundamentals and Applications, Wiley, 2004.

[11] P.I. Haris, F. Severcan, FTIR spectroscopic characterization of protein structure in aqueous and non-aqueous media, J. Mol. Catal. B-Enzymatic 7 (1–4) (1999) 207–221.

[12] M. Boulet-Audet, B. Byrne, S.G. Kazarian, Cleaning-in-place of immunoaffinity resins monitored by in situ ATR-FTIR spectroscopy, Anal. Bioanal. Chem. 407 (23) (2015) 7111–7122.

[13] M. Boulet-Audet, S.G. Kazarian, B. Byrne, In-column ATR-FTIR spectroscopy to monitor affinity chromatography purification of monoclonal antibodies, Sci. Rep. 6 (2016) 30526.

[14] F. Dousseau, M. Therrien, M. Pezolet, On the spectral subtraction of water from the FT-IR spectra of aqueous-solutions of proteins, Appl. Spectrosc. 43 (3) (1989) 538–542.

[15] S.E. Glassford, Applications of ATR-FTIR spectroscopic imaging to proteins, PhD Thesis, Chemical Engineering, Imperial College London, London, 2013, p. 158.

[16] K.L.A. Chan, S.G. Kazarian, Detection of trace materials with Fourier transform infrared spectroscopyusing a multi-channel detector, Analyst 131 (2006) 126–131.

[17] K.L.A. Chan, S.G. Kazarian, Chemical imaging of the stratum corneum under controlled humidity with the attenuated total reflection Fourier transform infrared spectroscopy method, J. Biomed. Opt. 12 (2007) 044010.

[18] A. McPherson, B. Cudney, Optimization of crystallization conditions for biological macromolecules, Acta Crystallogr. F Struct. Biol. Commun. 70 (Pt 11) (2014) 1445–1467.

[19] I. Russo Krauss, et al., An overview of biological macromolecule crystallization, Int. J. Mol. Sci. 14 (6) (2013) 11643–11691.

[20] N.E. Chayen, The role of oil in macromolecular crystallization, Structure 5 (10) (1997) 1269–1274.

[21] M. Caffrey, A comprehensive review of the lipid cubic phase or in meso method for crystallizing membrane and soluble proteins and complexes, Acta Crystallogr. F. Struct. Biol. Commun. 71 (Pt 1) (2015) 3–18.

[22] K.L.A. Chan, S.V. Hammond, S.G. Kazarian, Applications of attenuated total reflection infrared spectroscopic imaging to pharmaceutical formulations, Anal. Chem. 75 (9) (2003) 2140–2146.

[23] S. Desbois, S.A. Seabrook, J. Newman, Some practical guidelines for UV imaging in the protein crystallization laboratory, Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun. 69 (Pt 2) (2013) 201–208.

[24] M.L. Pusey, Developing a fluorescence-based approach to screening for macromolecule crystallization conditions, Cryst. Growth Des. 11 (4) (2011) 1135–1142.
[25] L.A. Averett, P.R. Griffiths, K. Nishikida, Effective path length in attenuated total reflection spectroscopy, Anal. Chem. 80 (8) (2008) 3045–3049.
[26] M. Boulet-Audet, et al., Quantitative determination of band distortions in diamond attenuated total reflectance infrared spectra, J. Phys. Chem. B 114 (24) (2010) 8255–8261.
[27] S.G. Kazarian, K.L. Chan, ATR-FTIR spectroscopic imaging: recent advances and applications to biological systems, Analyst 138 (7) (2013) 1940–1951.
[28] F.H. Tay, S.G. Kazarian, Study of petroleum heat-exchanger deposits with ATR-FTIR spectroscopic imaging, Energy Fuels 23 (8) (2009) 4059–4067.
[29] K.L. Chan, S.G. Kazarian, Attenuated total reflection Fourier transform infrared imaging with variable angles of incidence: a three-dimensional profiling of heterogeneous materials, Appl. Spectrosc. 61 (1) (2007) 48–54.
[30] T. Frosch, et al., Nondestructive three-dimensional analysis of layered polymer structures with chemical imaging, Langmuir 26 (24) (2010) 19027–19032.
[31] C.L. Song, S.G. Kazarian, Three-dimensional depth profiling of prostate tissue by micro ATR-FTIR spectroscopic imaging with variable angle of incidence, Analyst 144 (2019) 2954–2964.
[32] M. Sauerborn, et al., Immunological mechanism underlying the immune response to recombinant human protein therapeutics, Trends Pharmacol. Sci. 31 (2) (2010) 53–59.
[33] D.S. Pisal, M.P. Kosloski, S.V. Balu-Iyer, Delivery of therapeutic proteins, J. Pharm. Sci. 99 (6) (2010) 2557–2575.
[34] J.S. Philo, Is any measurement method optimal for all aggregate sizes and types? AAPS J. 8 (3) (2006) E564–E571.
[35] J.M. Reichert, Which are the antibodies to watch in 2013? MAbs 5 (1) (2013) 1–4.
[36] B. Kelley, Very large scale monoclonal antibody purification: the case for conventional unit operations, Biotechnol. Prog. 23 (5) (2007) 995–1008.
[37] S. Flatman, et al., Process analytics for purification of monoclonal antibodies, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 848 (1) (2007) 79–87.
[38] C. Jiang, et al., A mechanistic study of protein A chromatography resin lifetime, J. Chromatogr. A 1216 (31) (2009) 5849–5855.
[39] R. Hahn, et al., Comparison of protein A affinity sorbents III. Life time study, J. Chromatogr. A 1102 (1-2) (2006) 224–231.
[40] D. Ami, et al., Structural analysis of protein inclusion bodies by Fourier transform infrared microspectroscopy, Biochim. Biophys. Acta 1764 (4) (2006) 793–799.
[41] E. Mitri, et al., Time-resolved FT-IR microspectroscopy of protein aggregation induced by heat-shock in live cells, Anal. Chem. 87 (7) (2015) 3670–3677.
[42] D. Ami, et al., Kinetics of inclusion body formation studied in intact cells by FT-IR spectroscopy, FEBS Lett. 579 (16) (2005) 3433–3436.
[43] D. Ami, et al., Why and how protein aggregation has to be studied in vivo, Microb. Cell Fact. 12 (2013) 17.
[44] D. Ami, et al., In situ characterization of protein aggregates in human tissues affected by light chain amyloidosis: a FTIR microspectroscopy study, Sci. Rep. 6 (2016) 29096.
[45] L.P. Choo, et al., In situ characterization of beta-amyloid in Alzheimer's diseased tissue by synchrotron Fourier transform infrared microspectroscopy, Biophys. J. 71 (4) (1996) 1672–1679.
[46] A.C. Leskovjan, A. Lanzirotti, L.M. Miller, Amyloid plaques in PSAPP mice bind less metal than plaques in human Alzheimer's disease, Neuroimage 47 (4) (2009) 1215–1220.
[47] L.M. Miller, M.W. Bourassa, R.J. Smith, FTIR spectroscopic imaging of protein aggregation in living cells, Biochim. Biophys. Acta 1828 (10) (2013) 2339–2346.
[48] J.A. Kimber, S.G. Kazarian, Spectroscopic imaging of biomaterials and biological systems with FTIR microscopy or with quantum cascade lasers, Anal. Bioanal. Chem. 409 (25) (2017) 5813–5820.

[49] A. Schwaighofer, et al., External cavity-quantum cascade laser infrared spectroscopy for secondary structure analysis of proteins at low concentrations, Sci. Rep. 6 (2016) 33556.
[50] M.R. Alcaraz, et al., EC-QCL mid-IR transmission spectroscopy for monitoring dynamic changes of protein secondary structure in aqueous solution on the example of beta-aggregation in alcohol-denaturated alpha-chymotrypsin, Anal. Bioanal. Chem. 408 (15) (2016) 3933–3941.
[51] A. Schwaighofer, et al., Recent advancements of EC-QCL based mid-IR transmission spectroscopy of proteins and application to analysis of bovine milk, Biomed. Spectrosc. Imaging 7 (1-2) (2018) 35–45.
[52] J.L. Klocke, et al., Single-shot sub-microsecond mid-infrared spectroscopy on protein reactions with quantum cascade laser frequency combs, Anal. Chem. 90 (17) (2018) 10494–10500.
[53] I. Amenabar, et al., Structural analysis and mapping of individual protein complexes by infrared nanospectroscopy, Nat. Commun. 4 (2013) 2890.
[54] J. De Meutter, et al., Analysis of protein microarrays by FTIR imaging, Biomed. Spectrosc. Imaging 5 (2) (2016) 145–154.
[55] T.P. Wrobel, et al., Micro-attenuated total reflection Fourier transform infrared (micro ATR FT-IR) spectroscopic imaging with variable angles of incidence, Appl. Spectrosc. 69 (10) (2015) 1170–1174.
[56] M.C. Nweke, A.S. Rathore, D.G. Bracewell, Lifetime and aging of chromatography resins during biopharmaceutical manufacture, Trends Biotechnol. 36 (10) (2018) 992–994.
[57] H. Tiernan, B. Byrne, S.G. Kazarian, Insight into heterogeneous distribution of protein aggregates at surface layer using ATR-FTIR spectroscopic imaging. Anal. Chem. (2020) http://dx.doi.org/10.1021/acs.analchem.0c00250.

CHAPTER 2

Light-induced difference Fourier-transform infrared spectroscopy of photoreceptive proteins

Hideki Kandori and Kota Katayama

Department of Life Science and Applied Chemistry and OptoBioTechnology Research Center, Nagoya Institute of Technology, Nagoya, Japan

2.1 Introduction

Biological systems utilize light as the source of energy and cellular signaling. These light-induced functions are governed by photoreceptive proteins, which contain an embedded chromophore molecule that senses light. Chromophore binding occurs with or without covalent bond formation. The covalent mode of chromophore binding can be seen in rhodopsins, phytochromes, and photoactive yellow proteins (PYPs) [1–11]. While retinal molecule is bound to a Lys residue through the Schiff base linkage in rhodopsins, the S-H group of Cys forms a covalent bond with the bilin chromophore and p-coumalic acid in phytochrome and PYPs, respectively. The noncovalent mode of chromophore binding can be seen in flavoproteins and photosynthetic reaction centers [12–21]. Upon light absorption by the chromophore, specific photochemical reactions such as isomerization and electron transfer take place in covalent and noncovalent modes, respectively, in the protein environment. These reactions are common for both modes of chromophore binding to photoreceptive proteins [1–21]. In the case of flavoproteins, a light-induced electron transfer reaction further leads to adduct formation with Cys and keto–enol tautomerization in LOV and BLUF domains, respectively. Altered chromophore–protein interaction induced by light finally leads to functional protein entities. It is however not easy to understand these processes, because the protein structural changes that lead to their function have to be analyzed in atomic detail.

Methods of determining structural biology require positional acquisition of atoms (or components) of biomolecules. Fig. 2.1 illustrates the crystal structures of three photoreceptive proteins, bacteriorhodopsin (BR), bovine rhodopsin (Rh), and (6-4) photolyase (PHR), from *Drosophila melanogaster*. These atomic structures serve as the molecular basis of our understanding of the mechanism of a light-driven proton pump,

Vibrational Spectroscopy in Protein Research.
DOI: https://doi.org/10.1016/B978-0-12-818610-7.00002-5

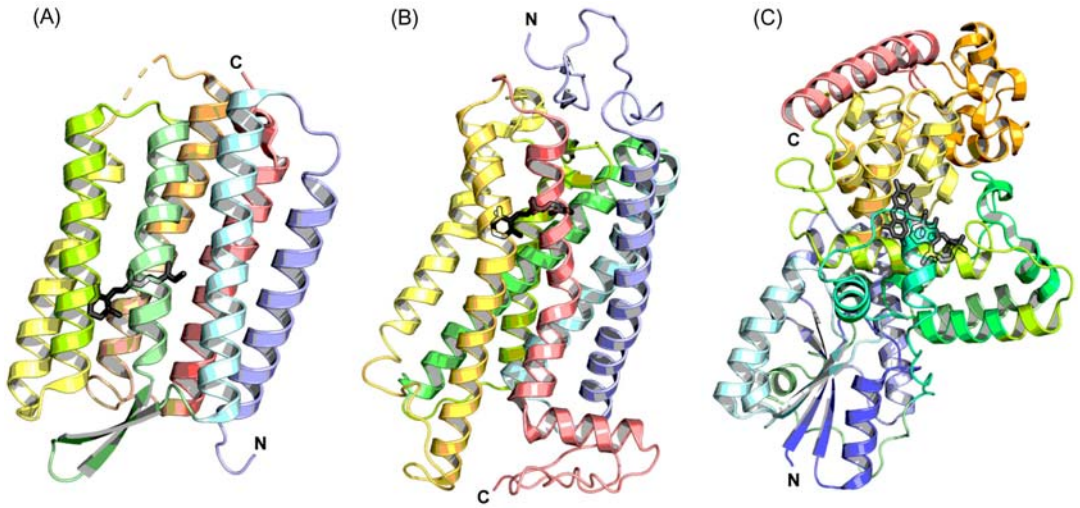


Figure 2.1
Structures of a light-driven proton pump bacteriorhodopsin (BR) [22] (A), bovine rhodopsin (Rh) [23] (B), and (6-4) photolyase (PHR) from *Drosophila melanogaster* [24] (C), which is a typical protein of microbial rhodopsin, animal rhodopsin, and flavoprotein, respectively. BR and bovine Rh are transmembrane (TM) proteins, while (6-4) PHR is a soluble protein. Each chromophore molecule, all-*trans* retinal (A), 11-*cis* fravin adenine dinucleotide (FAD) (C), is illustrated by a grey stick drawing.

a light-induced transduction in vision, and reparation of damaged DNA by light, respectively. Other structural techniques such as NMR and electron microscopy are also being widely utilized.

Vibrational spectroscopy such as Raman and IR has not been regarded as an experimental tool of structural biology, as it cannot determine atomic positions of biomolecules. In addition, in the case of IR spectroscopy, strong absorption of water results in difficulties during measurement. However, if the water content of the sample is significantly reduced without losing its function, IR spectroscopy can be a powerful and extremely sensitive technique for functional studies of photoreceptive proteins [25–30]. For instance, reducing the water content in the hydrated film method, we have shown that alterations in the hydrogen-bonding network can be monitored for protein-bound water molecules, which is extremely important in order to understand the structure–function relationship in photoreceptive proteins.

This method is based on a commercial Fourier-transform infrared (FTIR) machine, and photointermediates are trapped at low temperatures or time-resolved by rapid-scan or step-scan techniques. Difference spectra derived by subtracting the photointermediate spectrum to its resting state spectrum (light minus dark spectrum) are the standard data type in light-induced difference FTIR spectroscopy. We have been working on light-induced difference

FTIR spectroscopy of photoreceptive proteins for more than 25 years. This chapter will summarize the results of light-induced difference FTIR spectroscopy studies on microbial rhodopsins such as BR, animal rhodopsins such as Rh, and flavoproteins such as (6-4) PHR.

2.2 Methods: light-induced difference Fourier-transform infrared spectroscopy

2.2.1 Sample preparation

For light-induced difference FTIR spectroscopy, we have established three kinds of sample preparation methods: (1) hydrated film, (2) redissolved sample, and (3) concentrated solution [31]. Fig. 2.2 illustrates and summarizes each sample preparation method, together with absorption spectra obtained in the IR region. We first dry the sample solution on an IR window, followed by the addition of buffer next to a film for a hydrated sample or onto the film for redissolved sample preparations. On the other hand, preparation of concentrated solution does not involve sample drying. Consequently, water content in the sample is lowest in the hydrated film method followed by the redissolved sample, and lastly the concentrated solution preparation methods (Fig. 2.2).

It has been known that many proteins, either soluble or transmembrane (TM), are functional in hydrated films [25–31]. We essentially use the hydrated film method for the FTIR measurements of microbial and animal rhodopsins. Purified samples in detergent are reconstituted into lipids, which are dried into a film with an aspirator. The hydrated film method is also used for studying LOV and BLUF domains. For most microbial rhodopsins, animal rhodopsins, and flavoproteins, protein molecules are randomly oriented on the IR window. In contrast, BR molecules display an orientational preference on the IR window due to their propensity to form two-dimensional crystals also known as purple membrane. Polarized FTIR spectroscopy of the oriented BR film can determine particular angles of the dipole moment of the vibration in question, in addition to its frequency.

For low-temperature FTIR measurements, a small aliquot ($\sim$1 μL) of either H_2O or its stable isotope (D_2O, $H_2{}^{18}O$, or $D_2{}^{18}O$) is added next to the film on the IR window, which is then sealed with the second IR window and a rubber ring and cooled (Fig. 2.2A). Thus, in this case, hydration stops when water gets frozen. In contrast, if the measurements are performed at temperatures >260K, hydration is continuous and difficult to control. Therefore a dry film is typically hydrated by putting $\sim$1 μL of 20% glycerol/water (v/v) next to the film on the IR window to avoid continuous hydration.

If a higher water content is needed to facilitate enzyme turnover, the sample is redissolved. Here we describe our established redissolved sample method for (6-4) PHR sample [32]. First, we put 2 μL of the sample solution containing 1 mM (6-4) PHR on an IR window

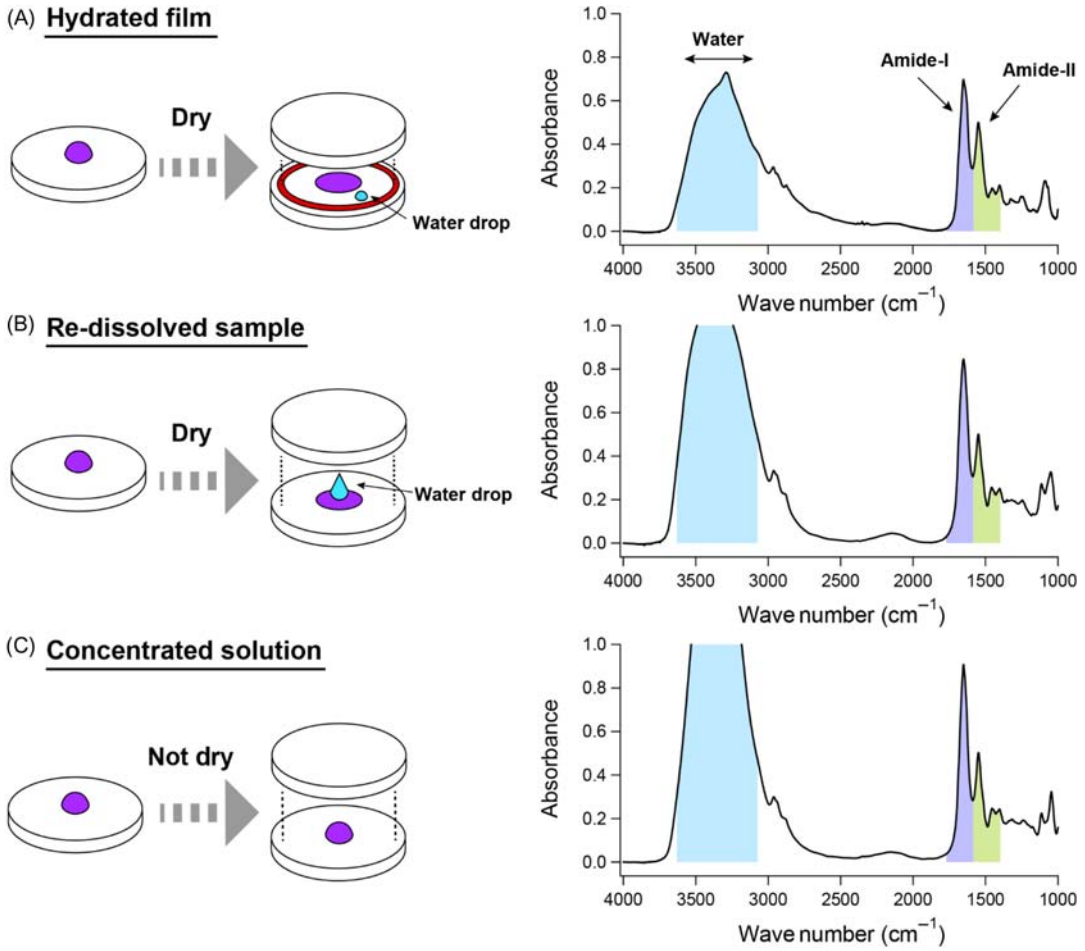


Figure 2.2

Summary of sample preparations of photoreceptive proteins for light-induced difference FTIR spectroscopy, together with absorption spectra in IR region. To prepare hydrated film (A) or redissolved sample (B), we first dry the sample solution on an IR window (BaF_2), followed by addition of buffer next to or onto the film, respectively. On the other hand, preparation of concentrated solution (C) does not dry the sample. Water content in the sample is hydrated film < redissolved sample < concentrated solution, as is clearly shown in the 3700−3000 cm^{-1} region (*light blue*) in each spectrum. Amide-I and -II correspond to C = O stretch and C−N stretch/N−H bending of peptide backbone, while the former also contains O−H bending of water. Each spectrum originates from hydrated film of LOV domain (A), redissolved sample of (6-4) PHR and concentrated solution of CRY-DASH (C), where the peaks of amide-II are normalized.

(BaF_2), and it is dried (Fig. 2.2B). For measurements of (6-4) PHR in complex with the (6-4) photoproduct, 2 μL or 1 μL of 2 mM (6-4) photoproduct in aqueous solution is gently mixed with the sample solution of (6-4) PHR on the IR window, and then dried. In this way

we are able to control the substrate/enzyme ratio. We then put 0.4 μL of buffer directly onto the dried film, and sandwich the film with another IR window. The (6-4) PHR sample in D_2O is prepared by diluting the (6-4) PHR with the same buffer prepared in D_2O followed by concentrating it by using an Amicon YM-30 filtration device (Millipore) for a total of three repetitions. Similarly, the (6-4) photoproduct is dissolved in D_2O, and the same procedure was carried out except D_2O was used as the buffer for the preparation of redissolved samples.

In PHRs the reduced form of flavin adenine dinucleotide (FAD) is the active form for DNA repair, whereas isolated PHRs in the resting state contain oxidized FAD molecules. PHRs inactivated by FAD oxidation can be reactivated by light-induced reduction [32]. Under the redissolved sample conditions, the oxidized FAD in the resting state of (6-4) PHR is reduced by continuous illumination with light above 390 nm, which results in destabilization of the semiquinone form of FAD. With this process, we can obtain reduced-minus-oxidized difference FTIR spectra of redissolved samples, enabling us to identify characteristic peaks for the oxidized (before illumination) and reduced (after illumination) forms of (6-4) PHR.

We found that *Synechocystis* sp. CRY-DASH exhibits a different photoreaction from a redissolved sample, suggesting protein unfolding due to sample drying. Therefore we typically perform FTIR spectroscopy on concentrated protein solution without drying [33]. This fact also implies that CRY-DASH is less tolerant to drying than (6-4) PHR. CRY-DASH was concentrated $\sim$20-fold using a Microcon YM-30 unit (Millipore). After 1.8 μL of concentrated CRY-DASH and 0.2 μL of 1 M DTT were mixed, the solution was put on a BaF_2 window (diameter of 18 mm), sandwiched with another BaF_2 window directly, and sealed with parafilm (Fig. 2.2C). H/D exchange was conducted by diluting the CRY-DASH sample with the buffer prepared in D_2O and concentrating three times with an Amicon YM-30 filter.

In Fig. 2.2 we compare the absorption spectra of hydrated film of LOV domain (A), redissolved sample of (6-4) PHR (B), and concentrated solution of CRY-DASH (C), where the peaks of amide-II are normalized. Water exhibits O−H stretching and bending vibrations at 3700−2800 and 1700−1600 cm^{-1}, respectively. As is clearly seen, absorbance at 3500−3200 cm^{-1} was $>$1.0 for the redissolved sample and concentrated solution, where we were not able to obtain reliable data in the light-induced difference FTIR spectroscopy. On the other hand, accurate difference FTIR spectra can be obtained for hydrated films.

2.2.2 Experimental measurements

Stationary light-induced difference FTIR spectra of photoreceptive proteins are measured using commercial FTIR spectrophotometers, such as FTS-40 and FTS-7000 (Bio-Rad) [34].

Low-temperature spectra are measured using a cryostat (Oxford DN, Oxford) and a temperature controller (ITC 4, Oxford) with liquid nitrogen as a coolant. Fig. 2.3A shows a typical experimental setup to measure light-induced difference FTIR spectra at >77K. As the illumination light source, we use a tungsten-halogen lamp (1 kW) or a Xenon lamp (300 mW). The latter is more suitable for the measurements on flavin-binding photoreceptors due to its stronger UV-blue region. A liquid N_2−cooled mercury cadmium telluride detector is normally used. Time-resolved step-scan experiments are performed in an optics compartment of an evacuated Bruker IFS 66 v/S FTIR spectrometer placed over a vibration isolation table [35]. Synchronization of data acquisition and laser triggering is performed with a delay pulse generator (DG535; Stanford Research Systems). Single-beam spectra recorded before laser excitation are averaged and used to transform single-beam spectra to absorbance spectra.

Fig. 2.3B illustrates the optical alignment of the FTIR spectrophotometer. Illumination light is introduced either parallelly or perpendicularly to the IR beam. The FTIR spectrometer has a movable mirror to take out the IR beam after the interferometer, whose position was changed in order to accommodate an illumination. In this case of parallel illumination, a shutter after the sample is closed during illumination, and the difference of the after-minus-before illumination is measured. For the perpendicular illumination, the sample tilted by 45 degrees is illuminated from the side. In this case, the difference of the during-minus-before illumination can be measured, which is advantageous to capture short-lived intermediates. Polarized FTIR spectroscopy is applied to the oriented BR films by placing an IR polarizer in front of the detector [34].

Fig. 2.4A shows absorption spectra of hydrated films of monkey green pigment (MG) in D_2O hydration. The spectra before (black) and after (red) light irradiation coincide very well with each other, but the difference of the after-minus-before illumination provides the reproducible spectral features (Fig. 2.4B). This is the light-induced difference FTIR spectrum.

2.3 Microbial rhodopsins

2.3.1 Bacteriorhodopsin

BR is a light-driven proton pump from *Halobacterium salinarum* and contains all-*trans* retinal as its chromophore. The retinal binds covalently to K216 residue through the protonated Schiff base linkage. Absorption of light triggers a cyclic reaction that comprises a series of intermediates, designated as the K-, L-, M-, N-, and O-states, leading to unidirectional proton transport across the membrane. These intermediate states were first identified by visible absorption spectroscopy [36], while resonance Raman spectroscopy revealed the structure of its retinal chromophore [37]. Upon photoexcitation of BR, the

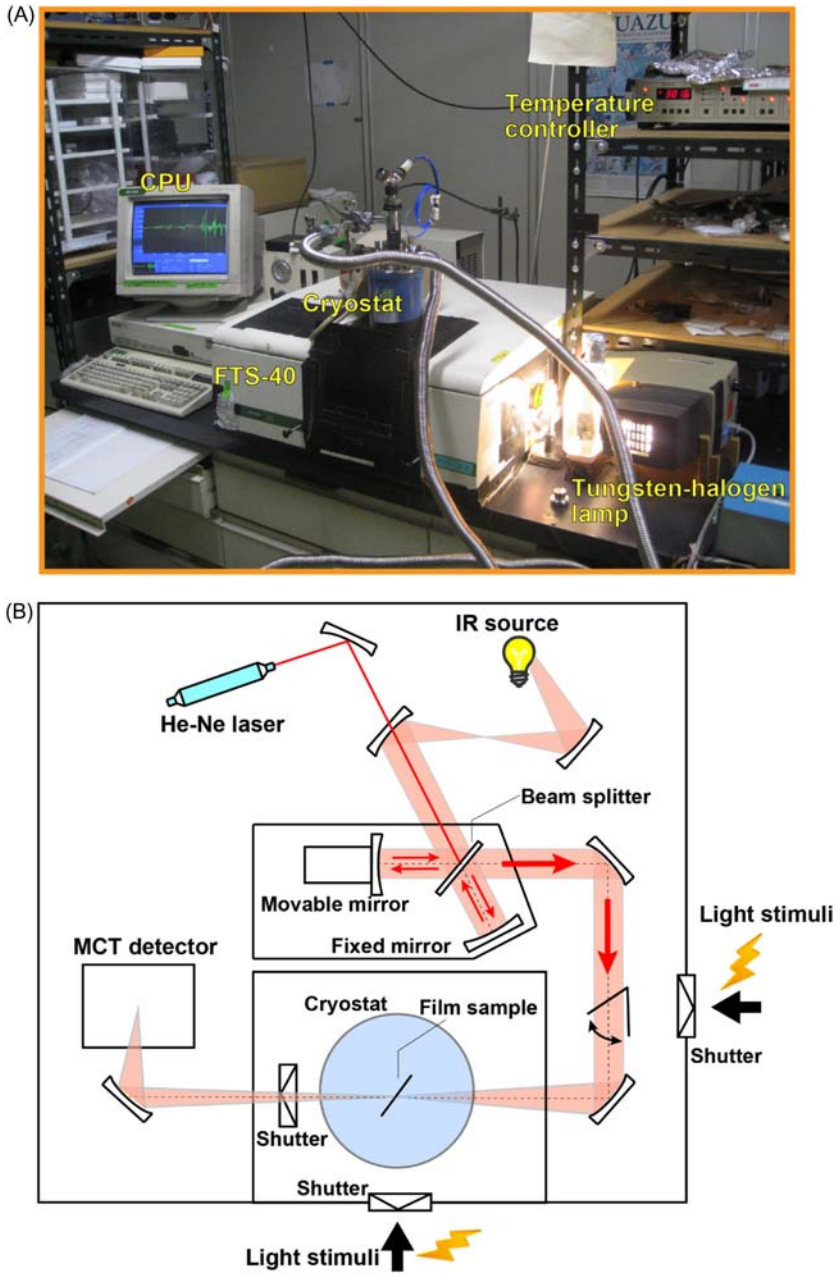


Figure 2.3

(A) Typical experimental setup to measure light-induced difference FTIR spectra. We use a FTS-40 (Bio-Rad) spectrophotometer, which is equipped with a cryostat (Oxford DN, Oxford) and a temperature controller (ITC 4, Oxford) with liquid nitrogen as a coolant. As the illumination light source, we use a tungsten-halogen lamp (1 kW) or a Xenon lamp (300 mW), and several optical filters are used. Open/close of shutters is controlled by the switch box. (B) Optical alignment of the FTIR spectrophotometer. Illumination light is introduced either parallel or perpendicular to the IR beam, and for the latter, the difference of the during-minus-before illumination can be measured in addition to the after-minus-before illumination, which is advantageous to capture short-lived intermediate states.

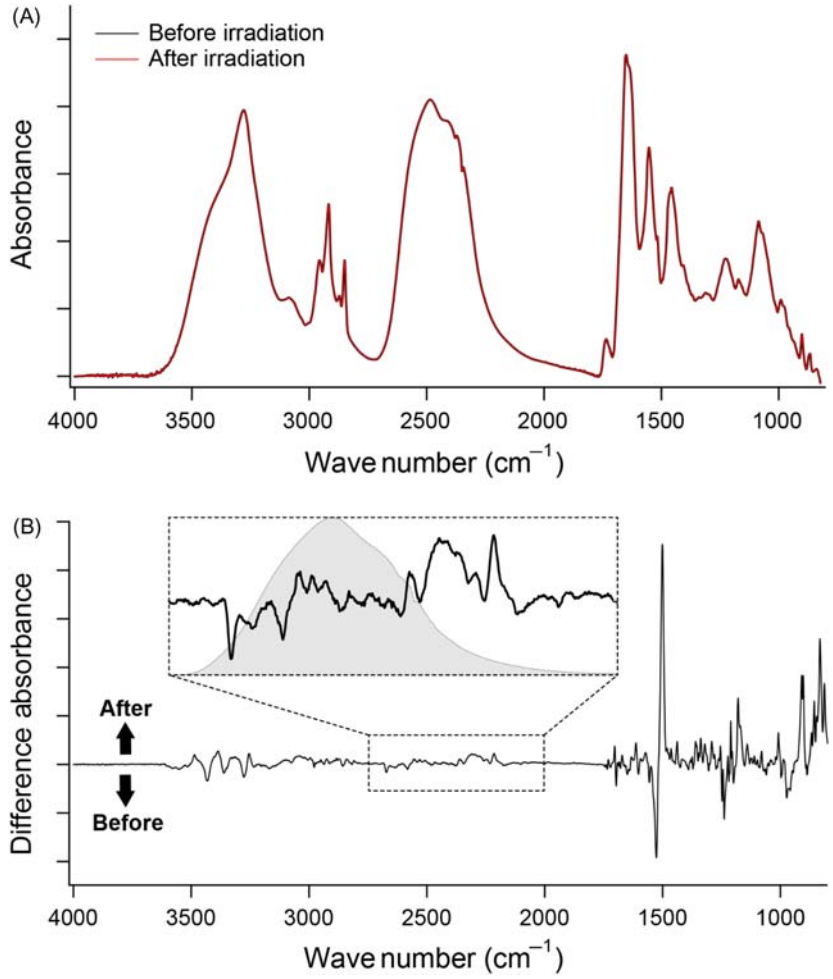


Figure 2.4

(A) Absorption spectra of hydrated films of monkey green pigment (MG), being hydrated with D_2O. Black and red curves represent the spectra before and after light irradiation at 77K. One division of the *y*-axis corresponds to 0.1 absorbance unit. (B) Light-induced difference FTIR spectra (after-minus-before illumination) of MG. The hydrated film was illuminated by a 501 light, which was supplied with a combination of a halogen-tungsten lamp (1 kW) and an interference filter. One division of the *y*-axis corresponds to 0.001 absorbance unit. Inset shows the expanded spectrum in the X-D stretch region for protein-bound water molecules.

all-*trans* to 13-*cis* isomerization occurs followed by a thermal reisomerization of the 13-*cis* to the all-*trans* form after formation of the O intermediate. The Schiff base is deprotonated in the M intermediate and reprotonated in the N intermediate, indicating that the Schiff base proton is conducted to the extracellular side upon M formation, and the cytoplasmic side supplies a proton to the Schiff base upon M decay.

Light-induced difference FTIR spectroscopy revealed how protons are captured by specific amino acids during the pumping process. The C = O stretching vibrations of protonated carboxylic acids in the 1800−1700 cm^{-1} region provided fruitful information on its proton pathway, even though the deprotonated C−O stretching vibrations (antisymmetric mode at about 1550−1610 cm^{-1} and symmetric mode at about 1300−1420 cm^{-1}) were obscured by many other vibrations in these frequency regions. It was concluded that the proton acceptor from the Schiff base is D85, and the proton donor to the Schiff base is D96.

In addition, internal water molecules were presumed to play important roles in BR function. They could stabilize the ion-pair structure inside protein, and possibly assist the transport of protons through the hydrophobic membrane-spanning region. The first experimental observation of internal water molecules of BR was by light-induced difference FTIR spectroscopy in 1992 [38], which was several years before BR crystal structural determination. One of the authors joined the group just after the first publication, and started light-induced difference FTIR spectroscopy of BR [39−46], which was summarized in a review article [25]. A water molecule has two O−H groups, and their frequencies are distributed in a wide 3700−2700 cm^{-1} region (Fig. 2.2), dependent on their coupling and hydrogen-bonding strength [47]. Gaseous water exhibits asymmetric and symmetric stretching modes at 3756 and 3657 cm^{-1}, respectively [48], and the stretching frequency is lowered in the condensed phase such as in liquid form and when bound to protein as its hydrogen bonding becomes stronger. O−H stretching vibrations of water molecules were identified by comparing hydrations of H_2O and $H_2{}^{18}O$, whose locations were analyzed by measuring various BR mutants [25]. However, the observed frequency region of water molecules in the initial FTIR studies was limited to $>3450\ cm^{-1}$, because (1) the spectral accuracy is less because of intense absorption by water, (2) many protein bands other than water O−H stretches overlap at this region, and (3) strong hydrogen-bonded water possesses broad O−H stretching bands. These facts have limited our observations of clear isotope shifts of water. In particular, an isotope shift between O−H and ^{18}O−H is about 10 cm^{-1}, and such a small shift could be easily obscured in complex spectral features in the 3450−3000 cm^{-1} region.

We have extended the frequency region of water stretching vibrations that monitor strong hydrogen bonds by optimizing the spectral accuracy in low-temperature light-induced difference FTIR spectroscopy [34]. Extremely accurate detection of light-induced difference FTIR spectra of BR in the whole mid-IR region provided useful information not only on water vibrations, but also on other vibrations such as the functional group of amino acid side chain [49−56]. One of the critical steps was to hydrate BR samples with D_2O. BR pumps deuterium in a D_2O environment, so that the O−H groups of internal water molecules that participate in proton transport should be exchanged for O−D groups and D_2O-insensitve stretching vibrations such as amide-A vibrations (N−H stretch) of peptide backbone are separated in frequency. In addition, an expected isotope shift between O−D

and $^{18}O-D$ of about 17 cm^{-1} is greater than that between O−H and $^{18}O-H$ (about 10 cm^{-1}). We thus anticipated the observation of the isotope shift even for the bridged water stretches, and this indeed worked. The K minus BR difference spectra at 77K (Fig. 2.5A) show the isotope shift of water not only for weak hydrogen bonds (2700−2500 cm^{-1}), but also for strong hydrogen bonds (<2400 cm^{-1}) [59,60]. There are at least six negative bands of water O−D stretches at 2690, 2636, 2599, 2323, 2292, and 2171 cm^{-1}, located at both edges in broad absorption spectra of D_2O (gray spectra in Fig. 2.5A). As the structural changes around the retinal chromophore upon formation of the K intermediate are limited, these water molecules are presumably located in the Schiff base region. Of particular note is the O−D stretch at 2171 cm^{-1}, which is of much lower frequency than those of the fully hydrated tetrahedral water molecules, suggesting that the hydrogen bond acceptor of this water molecule is negatively charged, such as the oxygen atom of D85 and D212 (Fig. 2.5B).

The origin of water O−D stretches in the K-minus-BR spectra was investigated by the use of various BR mutants [58,61]. The water bands were not affected by mutations at the cytoplasmic side, such as T46V, D96N, and D115N, implying that water molecules in the cytoplasmic side do not change their hydrogen bonding pattern in the BR resting state to K transition. In contrast, significant modifications of the water bands were observed for mutations in the Schiff base region and at the extracellular side, such as R82Q, D85N, T89A, Y185F, D212N, R82Q/D212N, and E204Q. These results suggested that the six O−D stretches of BR originate from three water molecules, water 401, 402, and 406, involved in the pentagonal cluster (Fig. 2.5B). Two stretching modes of each water molecule are highly separate (300−470 cm^{-1} for O−D stretches, and 500−770 cm^{-1} for O−H stretches), which is consistent with the quantum mechanics/molecular mechanics (QM/MM) calculation [62]. The small amplitudes of vibrational coupling are presumably due to strong association of the water molecules to negative charges of D85 and D212. It should be noted that the hydrogen-bonding interaction of water402 is stronger with D85 than with D212, though the distance between water402 and D85 (2.6 Å) is similar to that between water402 and D212 (2.8 Å) (Fig. 2.5B) [22]. This suggests that the geometry of the hydrogen bond of water402 is ideal with D85, but not with D212. Since the bridged water interacts more strongly with D85, D85 is presumably a stronger counterion than D212, which is consistent with our knowledge that D85 and D212 are the primary and secondary counterions, respectively.

Notably and unlike other BR mutants, D85N and D212N that lack strong hydrogen-bonded waters do not pump protons. Therefore the proton pump activity seems to be correlated with the presence of internal water molecules under strong hydrogen-bonded conditions. BR lacks proton-pumping activity if D85 or D212 are mutated to neutral residues, but binding of Cl^- has different functional effects in BR mutants at these positions. Binding of Cl^- to D85T or D85S converts BR into a chloride-ion pump, while binding of Cl^- to D212N

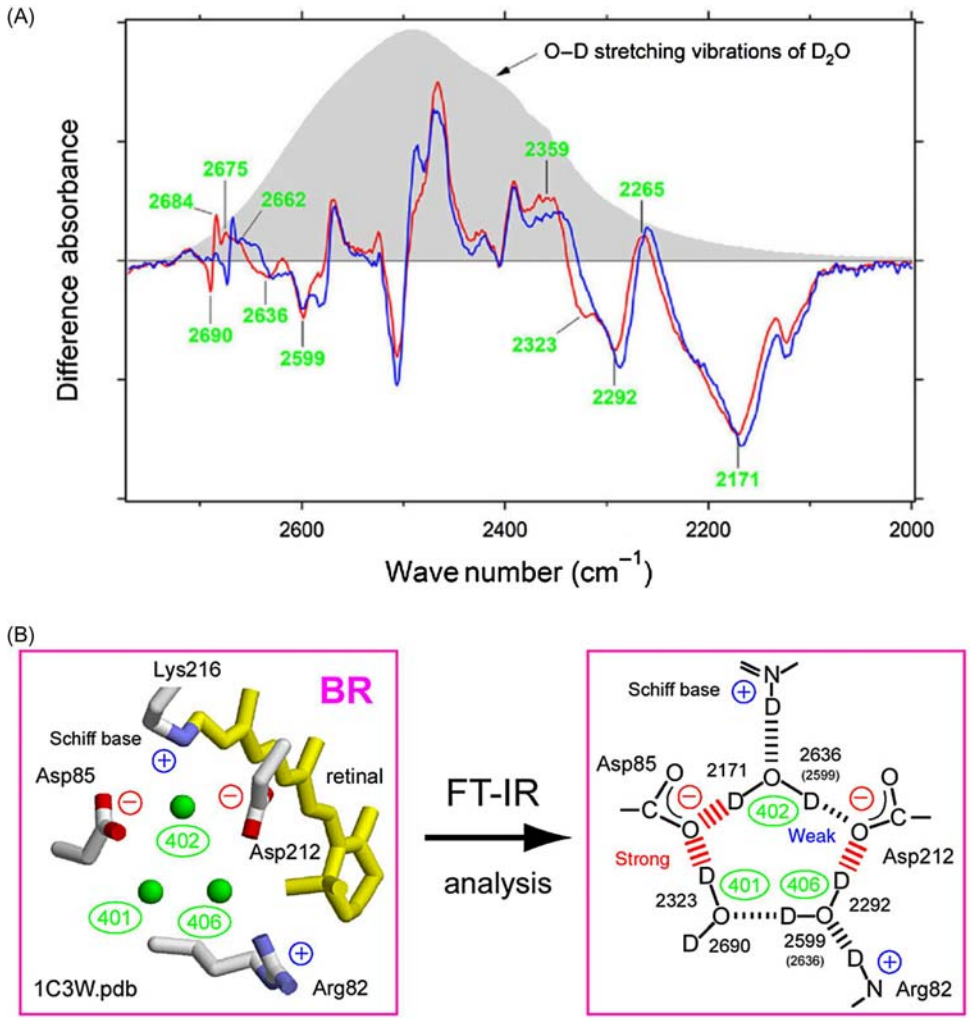


Figure 2.5

(A) The K minus BR difference FTIR spectra measured at 77K in the 2760–2000 cm^{-1} region. The spectra are compared between hydration with D_2O (*red lines*) and $D_2{}^{18}O$ (*blue lines*). The gray curve represents O-D stretching vibrations of D_2O. Green-labeled frequencies correspond to those identified as O-D stretching vibrations of water. (B) Schematic drawing of the deuterated pentagonal cluster structure in the Schiff base of BR. The numbers are the O-D stretching frequencies of water molecules in cm^{-1} assigned in the FTIR study. Source: *(A) This figure is modified from H. Kandori, Hydration switch model for the proton transfer in the Schiff base region of bacteriorhodopsin, Biochim. Biophys. Acta 1658 (2004) 72–79 [57]. (B) This figure is modified from M. Shibata, H. Kandori, FTIR studies of internal water molecules in the Schiff base region of bacteriorhodopsin, Biochemistry 44 (2005) 7406–7413 [58].*

restores the proton-pumping activity at low pH. FTIR spectroscopy revealed the absence of strong hydrogen-bonded water molecules in D85S [63]. In contrast, although D212N BR lacks strong hydrogen-bonded water molecules, they do appear at 2350−2310 cm^{-1} upon halide binding [64]. Therefore these results are consistent with our hypothesis that strong hydrogen-bonded water molecules are a prerequisite for the proton pumping activity of rhodopsins. This hypothesis is further examined by testing the water bands of other rhodopsins below.

Low-temperature studies of internal water molecules in BR are based on the assumption that positions and properties of water molecules are similar at room- and low-temperatures. It is generally accepted that BR intermediates trapped at low temperatures should display a protein structure and a retinal conformation essentially equivalent to their room-temperature counterparts. But can such an argument be applied to internal water molecules? Even though the time-averaged BR protein structures and retinal conformations for different intermediates are truly conserved at low temperature, internal water molecules and the associated hydrogen bonding network could be significantly affected. Our time-resolved room-temperature FTIR difference spectra for the BR photocycle showed that the room-temperature M-minus-BR spectrum was almost identical to that at 230K (Fig. 2.6), even in the frequency region of water stretching vibrations [35,65].

Time-resolved FTIR spectroscopy has proven to be extremely useful in studying the proton release group (PRG) in the extracellular domain of BR. A proton is released from BR by lowering the p*K*a value of the PRG from 9 to 5.8 in the M intermediate. Although two glutamic acids at the extracellular surface, E204 and E194, were good candidates for the PRG, a deprotonation signal of carboxylic acid was not clearly detectable upon formation of the M intermediate [66], suggesting that the PRG is not formed by these glutamic acids. Previous time-resolved FTIR studies reported a broad negative IR continuum band at 2000−1800 cm^{-1} in the M-minus-BR difference spectrum, and proposed that the continuum band originates from a protonated water cluster [67]. Our time-resolved FTIR study of BR indeed observed an isotope effect of ^{18}O water on the continuum signal, indicating that the continuum band contains vibrations of water molecules. Thus it is now established that the PRG in BR is a protonated water cluster, although the structure of the water cluster and mechanism of its high p*K*a are still under investigation. Interestingly, the continuum band was not observed at low temperatures [65].

2.3.2 Other microbial rhodopsins

For the 30 years from the early 1970s, microbial rhodopsins were epitomized by haloarchaeal proteins such as BR and its close relatives, light-driven inward chloride pump halorhodopsin (HR), sensory rhodopsin I and II (SRI and SRII) [1]. Later, several thousand related photoactive proteins with similar or different functions were identified in

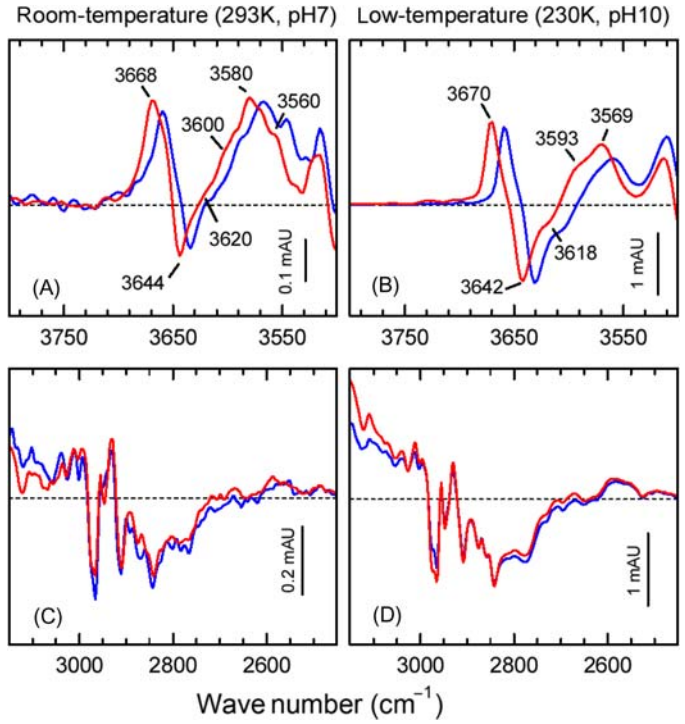


Figure 2.6

Comparison of water O-H stretching vibration regions of M minus BR difference FTIR spectra measured at between the room (A, C) and low (B, D) temperature. Bands from water O−H vibrations were assigned and labeled by their isotopic downshift when $H_2{}^{16}O$ (*red lines*) was substituted by $H_2{}^{18}O$ (*blue lines*). (A, B) Weakly hydrogen-bonded O−H stretching region. (C, D) Strongly H-bonded O−H stretching region. Source*: These figures are modified from V.A. Lorenz-Fonfria, Y. Furutani, H. Kandori, Active internal waters in the bacteriorhodopsin photocycle. A comparative study of the L and M intermediates at room and cryogenic temperatures by infrared spectroscopy, Biochemistry 47 (2008) 4071−4081 [35] (Copyright: American Chemical Society).*

Archaea, *Eubacteria*, and *Eukaryota*. Proteorhodopsins (PRs), discovered in the year 2000 [68], are the most abundant microbial rhodopsins in marine bacteria, and function as light-driven proton pumps like BR. Channelrhodopsins (ChRs), another group of microbial rhodopsins, were discovered in 2002−03 in green algae [69,70], where they function as light-gated cation channels within the algal eye to depolarize the plasma membrane upon light absorption. Discovery of ChR led to the emergence of a new field, optogenetics, in which light-gated ion channels and light-driven ion pumps were being used to depolarize and hyperpolarize selected neuronal cells to understand the circuitry of the human brain [71−74]. Discovery of new functions of microbial rhodopsins has accelerated recently, including the discovery of the light-driven sodium pump, light-gated anion channel, light-driven inward proton pump, and light-activated enzyme [1]. Structures of these rhodopsins have been determined except for enzyme rhodopsins,

which are structurally similar despite their functional diversity. This suggests the importance of structural changes for each function.

Through several studies conducted on BR, it has been well established that strong hydrogen-bonded water molecules play a crucial role in its function, for which highly accurate low-temperature FTIR spectroscopy was used. The analysis was extended to various rhodopsins including animal rhodopsins to elucidate whether they contain strong hydrogen-bonded water or not in their resting state (O−D stretch at $<2400\ cm^{-1}$). As a result, strong hydrogen-bonded water molecules were discovered in all light-driven proton pumps, including eubacterial pumps such as PR [75] and *Gloeobacter* rhodopsin (GR) [76], and eukaryotic pumps such as *Leptosphaeria* rhodopsin (LR) [77]. In 2012 we presented a correlation theory between the frequency of water O−D stretching vibrations and protein function [78]. The rhodopsins exhibiting proton-pumping activity possess water O−D stretching vibrations at $<2400\ cm^{-1}$, but those of chloride pumps such as HR from *H. salinarum* (*s*HR) [79] and *Natromonas pharaonis* (*p*HR) [80] are at $>2400\ cm^{-1}$. Notably, *p*HR pumps protons in the presence of azide, and we observed strong hydrogen-bonded water molecules in the azide-bound *p*HR [81]. On the other hand, chloride-bound *p*HR has no strong hydrogen-bonded waters. An absence of strong hydrogen-bonded water was also found in sensory proteins such as *Anabaena* sensory rhodopsin (ASR) [82,83], *Neurospora* rhodopsin (NR) [84], and visual sensors in animal such as bovine rhodopsin [85], squid rhodopsin [86], and monkey green and red sensitive pigments [87]. SRI from *Salinibacter ruber* (*Sr*SRI) [88] and SRII from *N. pharaonis* (*p*SRII) [89] possess water O−D stretching vibrations at $<2400\ cm^{-1}$. However, SRI and SRII function as sensors and not pumps. Interestingly, SRI and SRII function as proton pumps in the absence of their transducer proteins, and therefore agree with the correlation theory.

While strong hydrogen-bonded water molecules were observed only for proton-pumping rhodopsins until the year 2012 [78], emergence of new rhodopsins complicates this correlation. While the rhodopsins exhibiting proton-pumping activity possess water O−D stretching vibrations at $<2400\ cm^{-1}$, such strong hydrogen-bonded water molecules are also observed for other rhodopsins such as light-driven sodium pumps KR2 [90] and NdR2 [91], an eubacterial chloride pump *Fluvimarina* rhodopsin (FR) [91], inward proton pump PoXeR [92], and ChR (C1C2 chimera) [93] (Fig. 2.7). These results clearly show the wide presence of strong hydrogen-bonded water molecules in microbial rhodospins. As the Schiff base is protonated and its positive charge is stabilized by counterion(s), it is likely that such a water molecule hydrates a negatively charged counterion.

Protein-bound water molecules were taken into consideration in interpreting the functional swapping of ion pumps. In 1995, we successfully converted BR into a chloride pump by a single amino acid replacement (D85T BR) [94]. This finding suggests that archaeal proton and chloride pumps share a common mechanism and the functionality is determined at the

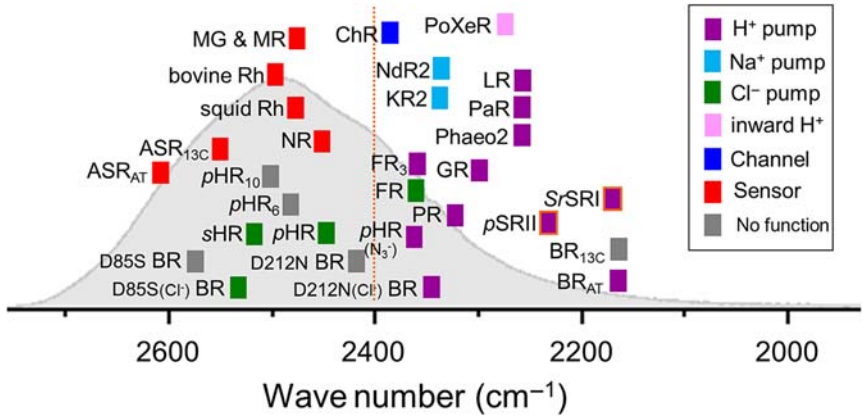


Figure 2.7

In the light-induced difference FTIR spectra of various rhodopsins at 77K, the lowest frequency among the observed water O−D stretching vibrations is plotted for the unphotolyzed state. Purple squares represent the results of the proton pump, while light-blue, green, pink, dark-blue, and red squares are those of the sodium pump, chloride pump, inward proton pump, cation channel, and sensor, respectively. The proteins of dark gray squares have no function. Abbreviations are as follows: BR_{AT} (BR containing all-*trans* retinal), BR_{13C} (BR containing 13-*cis* retinal), D85S BR, D85S(Cl^-) BR, D212N BR, D212N(Cl^-) BR, *Sr*SRI (*Salinibacter ruber* sensory rhodopsin I), *p*SRII (*Natronomonas pharaonis* sensory rhodopsin II), LR (*Leptosphaeria* rhodopsin), PaR (*Pantoea ananatis* rhodopsin), Phaeo2 (*Phaeosphaeria* rhodopsin 2), GR (*Gloeobacter* rhodopsin), PR (proteorhodopsin), ASR_{AT} (*Anabaena* sensory rhodopsin containing all-*trans* retinal), ASR_{13C} (*Anabaena* sensory rhodopsin containing 13-*cis* retinal), NR (*Neurospora* rhodopsin), bovine Rh (bovine rhodopsin), squid Rh (squid rhodopsin), MG and MR (monkey green- and red-sensitive visual pigments), *s*HR (*H. salinarum*HR), *p*HR (*Natronomonas pharaonis* HR), *p*HR(N_3^-) (*Natronomonas pharaonis* HR that binds azide), pHR_6 (sextet mutant of *Natronomonas pharaonis* HR), pHR_{10} (10-fold mutant of *Natronomonas pharaonis* HR), ChR (C1C2 chimeric channelrhodopsin), KR2 (*Krokinobacter eikastus* rhodopsin 2), NdR2 (*Nonlabens dokdonesis* rhodopsin 2), FR (*Fulvimarina* rhodopsin), FR_3 (triple mutant of *Fulvimarina* rhodopsin), and PoXeR (Xenorhodopsin from). SRI and SRII are sensors, but it is known that they pump protons in the absence of transducer proteins (*purple squares* with *orange frames*). Light gray spectrum shows O−D stretching vibration of D_2O, and proton-pumping rhodopsins possess strongly hydrogen-bonded waters (O−D stretch at $<2400\ cm^{-1}$). Source: *This figure is modified from Y. Nomura, S. Ito, M. Teranishi, H. Ono, K. Inoue, H. Kandori, Low-temperature FTIR spectroscopy provides evidence for protein-bound water molecules in eubacterial light-driven ion pumps, Phys. Chem. Chem. Phys. 20 (2018) 3165−3171 [91] by adding the data of PoXeR.*

active center (position 85 in BR). However, neither the respective reverse swap mutation (T→D) nor combinations of other point mutations could convert HR into a proton pump [78], indicating that functional conversion between archaeal proton and chloride pumps is asymmetric in nature. This unsuccessful conversion of *p*HR into a proton pump could be attributed to the lack of strong hydrogen-bonded water molecules. Similarly, we have succeeded in converting eubacterial Cl^- and Na^+ pumps into proton pumps, and the Na^+ pump into a Cl^- pump, but their reverse conversions were unsuccessful [95]. Regarding the

asymmetric functional conversion between the two pumps, successful functional conversions are achieved exclusively when mutagenesis reverses the evolutionary amino acid sequence changes. This suggests that the essential structural mechanism of an ancestral function is retained even after the gain of a new function, and FTIR spectroscopy monitored the presence of strong hydrogen-bonded water molecules for all eubacterial H^+, Na^+, and Cl^- pumps, including a functionally converted mutant [91]. Thus a strong hydrogen-bonded water molecule is maintained for these new functions during evolution, and the water-containing hydrogen-bonding network constitutes one of the essential structural mechanisms in light-driven ion pumps. In rhodopsins photon energy is first stored by retinal photoisomerization, which is used for subsequent structural changes for each function. In addition to the distorted chromophore, the hydrogen bond of protein-bound water molecules carries energy for the proton pump function, which can be monitored by FTIR spectroscopy.

2.4 Animal rhodopsins

Visual perception in animals occurs through the absorption of electromagnetic radiation from 400 to 780 nm by photoreceptor visual pigments in the retina. These visual proteins include rhodopsin and cone opsins→color pigments harboring the 11-*cis*-retinal chromophore. The absorption of visible light causes photoisomerization of the 11-*cis*-retinal chromophore to an all-*trans* configuration [1−4], which is followed by conformational changes in the opsin protein moiety that allows binding of a G-protein (transducin) to the cytoplasmic surface of the receptors.

2.4.1 Bovine rhodopsin

Rhodopsin belongs to the super family of G-protein-coupled receptors (GPCRs) and has been a crucial player in the field of GPCR structure−function biology. The determination of the atomic resolution crystal structure of rhodopsin produced major progress in our understanding of the molecular properties of rhodopsin and GPCRs in general [23,96,97]. The unique seven TM architecture of rhodopsin features TM3, a cytoplasmic helix (H8) following TM7, and a disulfide bridge linking TM3 with the extracellular loop 2 connecting TM4 and TM5. The chromophore-binding pocket of rhodopsin harbors a retinal chromophore stabilized by several polar and nonpolar amino acid residues that cumulatively affect the light absorbance properties of rhodopsin. This chromophore−protein specific interaction allows preferential absorption of a selected range of wavelengths, so-called opsin shift. In addition, the retinal distortion also plays a critical role in rhodopsin light sensitivity and spectral tuning.

Crystal structure of photoactivated rhodopsin, Meta-II state, indicates a relaxed conformation of retinal chromophore with translocation of the β-ionone ring and large helical movements in TM5 and TM6 resulting in a large fissure for G-protein binding and activation [98,99]. However, the crystal structures of inactive and active forms of rhodopsin are not sufficient to provide direct information on the dynamic light-induced changes. Light-induced difference FTIR spectroscopy has proven to be an extremely useful technique to unveil various molecular events that take place during the formation of Batho, BSI (transient blue-shifted intermediate), and Meta-I intermediate states [1,100,101]. Most intermediate states after the chromophore *cis*−*trans* isomerization can be stabilized by low temperature or by specific pH. FTIR spectroscopy can also monitor the structural changes at the protein−protein and protein−water level when Meta II interacts with transducin [102,103].

Chromophore configurations in all rhodopsin intermediate states have been investigated by monitoring the vibrational bands in the fingerprint, hydrogen-out-of-plane (HOOP), and protonated Schiff base C = NH (and C = ND) stretching vibration region [104,105]. In contrast, light-induced changes of protonation states or hydrogen-bond strengths are interpreted from the C = O stretch of carboxylic acids involved in proton transfer reactions [106]. For example, three carboxylic C = O stretches are affected during the photoactivation to the Meta-II state. These groups were successfully assigned as D83 [107,108], E122 [107] at 1768 (−)/1748 (+)/1728 (−) cm^{-1}, and E113 [109] at 1713 cm^{-1}, based on the disappearance of specific difference bands in site-directed mutants. The residues D83 and E122 are protonated in both inactive and active states, whereas E113 as the primary counterion becomes protonated.

Some internal water molecules also participate in the photoactivation process of rhodopsin by stabilizing both inactive and active conformations through specific hydrogen bonding networks. Similar to microbial rhodopsins, the assignment of water vibrations was first performed in the O−H stretch region at 3700−3500 cm^{-1} that revealed weak hydrogen-bonded water molecules [110,111]. Then, accurate measurements using $D_2O/D_2{}^{18}O$ hydration enabled assignment of water vibration in the whole mid-IR region. Consequently, the study revealed that bovine rhodopsin does not contain strong hydrogen-bonded water molecules during the photoactivation (Fig. 2.8) [112]. The crystal structure of rhodopsin indicated that the ion pair state of the Schiff base and Glu113 has no bridging water molecule, which was consistent with FTIR observations [112].

Analysis of the X-H stretch led to an unexpected finding related to dark noise of rhodopsin. Efficient retinal photoisomerization contributes to a single photon response in our vision, but discrete electrical signals in the dark (dark events) limit the potential capability of the rod visual system to discern single photons from thermal noise [113]. Although both light and dark events originate from rhodopsin activation, the experimental activation barrier of

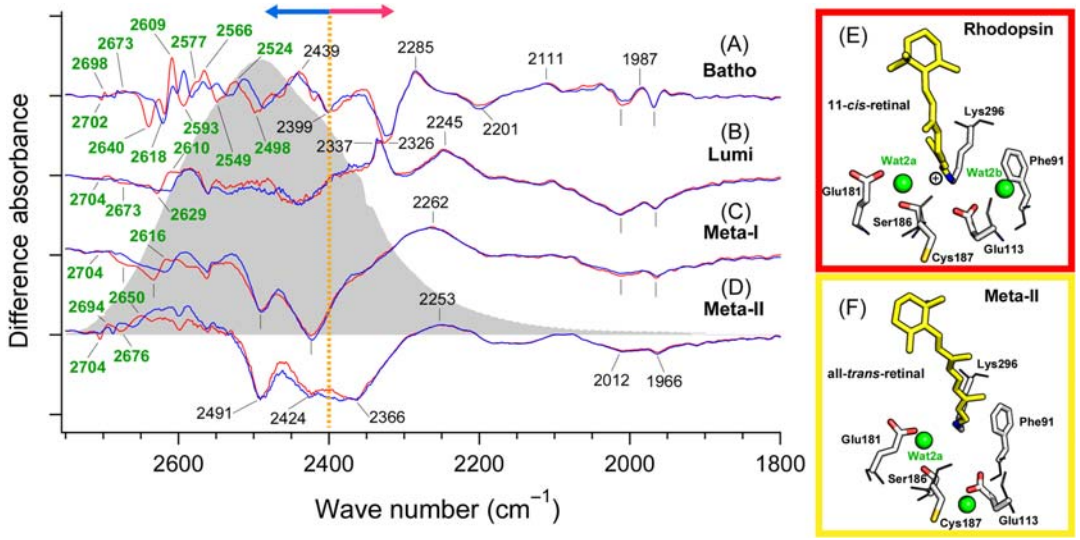


Figure 2.8

Light-induced difference FTIR spectra of bovine rhodopsin in the 2750−1800 cm^{-1} region, which correspond to the Batho-minus-rhodopsin spectra at 77K (A), Lumi-minus-rhodopsin spectra at 200K (B), Meta I-minus-rhodopsin spectra at 240K (C), and Meta II-minus-rhodopsin spectra (D). Red and blue lines represent the spectra in D_2O and $D_2{}^{18}O$, respectively. Light gray spectrum shows O−D stretching vibration of D_2O, and the frequencies at >2400 and <2400 cm^{-1} correspond to weak and strong hydrogen bonds, respectively. Green labeled frequencies correspond to those identified as water stretching vibrations. One division of the *y*-axis corresponds to 0.0017 absorbance units. Structures of rhodopsin (E) [93] and Meta-II (F) [98] are also shown. Source: *(A−D) The data are obtained from Y. Furutani, Y. Shichida, H. Kandori, Structural changes of water molecules during the photoactivation process in bovine rhodopsin, Biochemistry 42 (2003) 9619−9625 [85].*

the dark events is only half (80−110 kJ/mol) of the barrier for light-dependent activation of rhodopsin [114]. This suggests the presence of specific pathway of thermal isomerization, to which extensive studies have been performed experimentally [115] and theoretically [116]. Low-temperature FTIR spectroscopy measurements showed an H/D unexchangeable X-H stretch upon retinal isomerization [117], which was later identified as the O−H group of T118 [118]. Our group found that the O−H group slowly deuterates at 40°C−70°C, suggesting that protein structural fluctuations drive the transient opening of the retinal-binding pocket. This two-step model that involves transient opening of the retinal-binding pocket and retinal thermal isomerization accurately reproduced the statistical properties of the dark events [119].

In addition to the intrinsic probes such as C = O stretches of protein backbone and carboxylic acids, and O−H or N−H stretches of hydrophilic amino acids and water molecules, unique vibrational signatures such as SH, CN, or N_3 could be also used as

indirect probes to analyze the intramolecular changes within a protein, especially rhodopsin. It is highly advantageous to use indirect probes as their vibrational frequencies are free from other vibrational absorptions typically present in proteins. Site-directed and nonnatural amino acid mutagenesis using *p*-azido-L-phenylalanine residues or site-directed mutagenesis to Cys residues have been utilized to investigate the conformational changes in the cytoplasmic surface of rhodopsin upon G protein interaction [120,121]. The results clearly showed early conformational changes in TM5 and TM6 during the formation of Meta-I state, which facilitates subsequent transition to Meta-II.

2.4.2 Primate color visual pigments

Opsin shift that results from a specific chromophore conformation and its interaction with specific amino acid residues, anions, and internal water molecules is more apparent for color pigments. Three cone opsins contribute to the normal human trichromatic vision. These color pigments have maximum absorbances (λ_{max}) at 420, 530, and 560 nm, for blue, green, and red pigments, respectively [122,123]. All color pigments share some level of primary sequence similarity and bind to a common chromophore, 11-*cis*-retinal. Several theoretical studies have suggested retinal polyene twisting away from its normally planar configuration to be the major mechanism of spectral tuning in color pigments. The retinal twisting is induced mainly by the electrostatic environment of the chromophore-binding pocket [124–126]. With no crystal structures available to date, preliminary structural studies on human green and red pigments have been reported using resonance Raman spectroscopy, where the observed vibrational bands were very similar [127]. Notably, resonance Raman spectroscopy only provides insights into vibrational signals from the chromophore and not from the protein moiety. IR spectroscopy has been able to provide vibrational signals from both the protein and the chromophore as well as from protein-bound water molecules.

We previously reported difference FTIR spectra of the chicken red pigment that was prepared from >2000 chicken retinae [128], but identification of vibrational bands of protein is difficult for native proteins. Thus we have optimized the expression of monkey color pigments in HEK293 cell lines for structural analysis using FTIR spectroscopy [129,130] (Fig. 2.9). Formation of the Batho-state is clear from the downshifted ethylenic C = C stretches of the chromophore at 1576 (−)/1560 (+) cm^{-1}, 1561 (−)/1536 (+) cm^{-1}, 1534 (−)/1509 (+) cm^{-1}, and 1526 (−)/1500 (+) cm^{-1} for monkey blue (MB), monkey rhodopsin (MRh), monkey green (MG), and monkey red (MR), respectively. This corresponds to the red-shift in the visible region, and provides direct evidence for a linear correlation between λ_{max} and the negative frequency of the 11-*cis* configuration [129,130]. HOOP wagging modes in the frequency region of 1000–800 cm^{-1} is one well-characterized factor that determines the chromophore distortion of pigments. In MG and

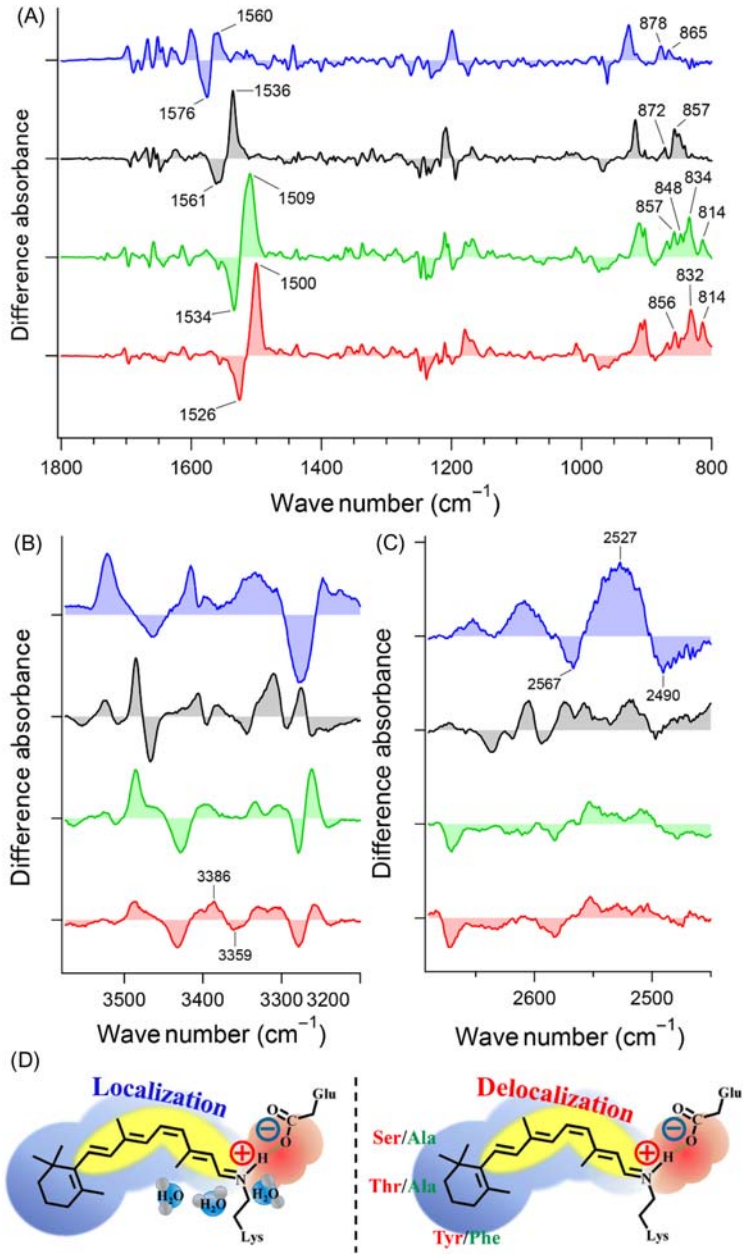


Figure 2.9

Light-induced difference FTIR spectra of monkey visual pigments, monkey blue (MB), monkey rhodopsin (MRh), monkey green (MG), and monkey red (MR) (from top to bottom), in the 1800–800 cm^{-1} region (A), in the 3575–3200 cm^{-1} region (B), and in the 2690–2450 cm^{-1} region (C). The difference FTIR spectra were measured at 77K. (D) Schematic drawing of the spectral tuning mechanism in MB (left) and MG/MR (right).

MR, the spectral features were very similar, whereas the IR band intensity was significantly decreased in MB, indicating that the larger spectral blue shift in λ_{max} of MB results from a different feature of chromophore distortion (Fig. 2.9A).

Electrostatic interaction between the PSB and its counterion could be deduced from the characteristic frequency of the N−D stretching vibration region of the PSB at 2800−1800 cm^{-1} measured in D_2O conditions. The N−D stretch of MG has been identified by using C_{15}−D synthetic retinal, from which N−D stretch was isotope-effected by vibrational coupling between N−D/C_{15}−D [131]. Notably, the frequencies of N−D stretch between MG and MR were identical, indicative of the same hydrogen-bonding strength between the PSB and counterion in both pigments. Therefore we conclude that the 30-nm difference between MG and MR originates exclusively from the polarity around the β-ionone ring and polyene chain. Amino acid residues A164, F261, and A269 in MG, and S164, Y261, and T269 in MR together with protein-bound water molecules around the β-ionone ring and polyene chain, but not at the Schiff base region might be responsible for the spectral tuning (Fig. 2.9D).

IR absorption by functional groups of amino acid and water molecules could be also further studied where the frequency shift of O−H/N−H or O−D/N−D stretching vibrations will be detectable at 3700−3000 cm^{-1} or 2800−2000 cm^{-1} regions. This will be applicable in the presence of an electrostatic environment between hydrophilic amino acids and water molecules in the chromophore-binding pocket. The observation of an MR-specific H/D unexchangeable O−H stretch of hydrophilic amino acid [129] and MB-specific water O−D stretches [130] might contribute to the π-electron localization/delocalization along the chromophore polyene chain resulting in a spectral red-shift in λ_{max} in MR and a spectral blue-shift in MB. We have recently discovered unique water IR signals suggesting the presence of a water cluster in the retinal binding pocket of MB (Fig. 2.9C), which was influenced by a mutation at position E113 and W265 in Rh. Therefore we infer that these water molecules are responsible for increasing polarity toward the retinal Schiff base, leading to the localization of the positive charge into the Schiff base, and consequently resulting in a blue shift in λ_{max} (Fig. 2.9D).

2.5 Flavoproteins

Flavin-binding photoreceptors are classified into three groups: (1) LOV domains of phototropin, (2) BLUF domains, and (3) cryptochromes (CRYs) [12−17,28,30,31]. Photolyases (PHRs) are DNA repair enzymes that belong to the CRY family. The chromophore of the LOV domain is flavin mononucleotide (FMN), while that of the BLUF domain and CRY/PHR family is a fravin adenine dinucleotide (FAD) molecule. Flavoproteins have been also used as optogenetic tools [13,132−134]. Most of the flavin-binding proteins in nature are redox enzymes [135]. This suggests that product formation by photochemical reaction of flavin is dangerous for these redox enzymes, which might have specific mechanisms to avoid such photochemical reactions. The photoexcited flavin state is

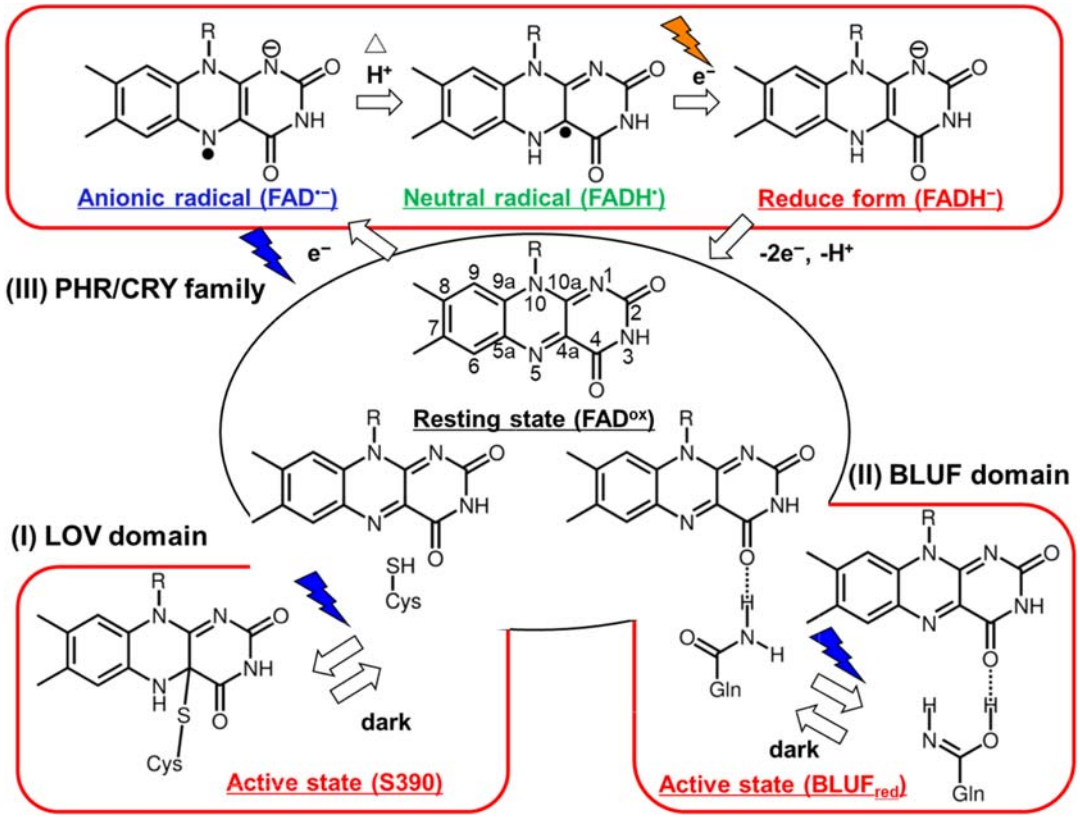


Figure 2.10
Chromophore molecule in flavoproteins. Resting state is the oxidized form of FMN (LOV domain) and FAD (BLUF domain and CRY/PHR family). The activation process of the LOV domain is an adduct formation between FMN and a nearby cysteine, while that of the BLUF domain is keto–enol tautomerism of Gln near FAD. The activation process of the CRY/PHR family is photoreduction of FAD.

relaxed efficiently by radiative or nonradiative transitions. Even when electron transfer takes place in flavin after light illumination, back electron transfer must occur rapidly. In contrast, flavin-binding photoreceptors utilize key amino acids in the chromophore-binding pocket to conduct chemical reactions that lead to chromophore relaxation. The photochemistry of these flavin-binding photoreceptors is known to be different from each other, and their mechanisms were studied by light-induced difference FTIR spectroscopy (Fig. 2.10).

2.5.1 LOV domain

Phototropin (Phot), a blue-light photoreceptor in plants, consists of two FMN-binding domains and a serine/threonine (Ser/Thr) kinase domain [136]. The photochemical reaction

of FMN yields kinase activation through changes in domain−domain interactions. The two FMN-binding domains (c. 100 residues) are called LOV (light, oxygen, and voltage sensing) domains, LOV1 and LOV2, due to their high homology to bacterial light-sensor PYP, oxygen-sensor FixL, and voltage-sensor HERG of a channel protein. Their protein fold belongs to the PAS (Per-Arnt-Sim) superfamily.

The primary photoreaction in the LOV domain is an adduct formation between FMN and a nearby cysteine (Fig. 2.10) [136]. After light absorption by FMN, intersystem crossing leads to the formation of a triplet-excited state that absorbs at 660 nm (L660) in a nanosecond timescale. The adduct formation is accompanied by the appearance of the S390 intermediate in a microsecond timescale. One of key questions about the LOV photoreaction mechanism revolves around its Cys reaction center. Although deprotonation of Cys was previously suggested by pH titration experiments on the FMN fluorescence [137], difference FTIR spectroscopy clearly monitored the S-H stretch of the reactive Cys in the dark, revealing protonation before light absorption (Fig. 2.10) [138]. The stretching frequency of cysteine S−H is in the 2580−2525 cm^{-1} region, where the other vibrations are absent. We observed a negative S−H vibration at $\sim$2570 cm^{-1} [138]. Since this LOV domain possesses only one Cys at position 966, this observation clearly indicated that C966 was indeed protonated. While the transfer reaction of a proton, a hydrogen, or an electron had been proposed from Cys in its triplet excited state, we detected the S−H stretch at 2537 cm^{-1} in the triplet excited state of neo1-LOV2 [139]. This excludes the possibility of proton and hydrogen transfer reactions. These frequencies imply that the S−H group is free from hydrogen bonding in the unphotolyzed state, while it forms a strong hydrogen bond in the triplet excited state. We infer the hydrogen-bonding acceptor to be the N5 atom of FMN, and such strong interaction presumably drives adduct formation in a microsecond timescale. In addition to the identification of the reactive cysteine, we elucidated light-induced structural changes in a hydrogen-bonding environment between FMN and the protein backbone of LOV domains [140−147]. Overall, FTIR studies conducted by several research groups [138−152] have provided great insights into the molecular mechanism of light-induced spectral changes in LOV domains leading to our understanding of their function.

2.5.2 BLUF domain

BLUF (sensor of blue-light using FAD) domain is a blue-light sensor that binds to a FAD chromophore and regulates a variety of functions in bacteria and eukaryotes [17]. For example, AppA from *Rhodobacter sphaeroides* and PAC (photoactivated adenylyl cyclase) from *Euglena gracilis* function as an antirepressor of photosynthesis gene expression, and a sensor for step-up photophobic response, respectively, where BLUF domains initiate light-induced signaling. Illumination of BLUF domains (λ_{max} = 450 nm) results in the formation

of a 10-nm red-shifted active intermediate, followed by dark regeneration of its original state with a very slow timescale. Interestingly, the chemical structure of FAD remains unchanged between the unphotolyzed and intermediate states of BLUF. Ultrafast spectroscopy showed that electron-coupled proton transfer takes place from the conserved Tyr/Gln in BLUF domain to the FAD chromophore, followed by the reverse transfer from FAD in the formation of the red-shifted intermediate [153].

In BLUF domains, Tyr/Gln are highly conserved near the FAD chromophore, suggesting that the activation of BLUF is accompanied by changes in the hydrogen-bonding network (Fig. 2.10). Light-induced difference FTIR spectroscopy reported strengthened hydrogen bonding of the C4 = O group of FAD in the active intermediate [154–158]. By applying accurate spectral detection for the X–H stretching region, we observed broad positive peaks at 2800–2500 cm^{-1} for BLUF domains [159]. Using isotope-labeling we have identified positive signals from the Fermi-resonance vibrations of Tyr, where the O–H stretch is coupled with overtone vibrations of phenol ring. Mutational analysis showed that this Tyr residue is present within the active center. Phenolic O–H stretching vibration appears at $\sim$3600 cm^{-1} in the absence of hydrogen bonds. The frequency is lowered as the hydrogen-bonding of the O–H groups becomes stronger. For example, the O–H stretch appears at $\sim$3200 cm^{-1} under strong intramolecular hydrogen bonding to a carbonyl group. However, the frequency observed here was 400 − 600 cm^{-1}, which was lower and indicates an unusually strong hydrogen bond. This provides direct evidence for the light-induced switching in the hydrogen-bonding network, where the conserved Tyr (Y21 for AppA) forms an unusually strong hydrogen bond with FAD in the intermediate.

Then what about the hydrogen-bonding network in the resting and intermediate states? An interesting proposal is that the conserved Gln (Q63 for AppA) changes its chemical structure from keto to enol form (Fig. 2.10). This proposal was first made by Stelling et al. using ultrafast infrared spectroscopy [158], and other FTIR and theoretical studies showed that the enol isomer of the Gln side chain was formed in the stable red-shifted intermediate state of BlrB and BlrP1-BLUF [160]. According to their study, the ^{15}N-Gln labeling of BlrB showed that keto–enol tautomerism of the Gln side chain occurred at the intermediate state. Using FTIR spectroscopy and QM/MM simulations, we recently found that Q63 in the AppA-BLUF domain shows keto–enol tautomerism upon the formation of the red-shifted intermediate [161], which is consistent with the previous report. However, unique to the AppA-BLUF domain, W104 does not flip out in the intermediate state. The O–H group of the side chain of Y21 forms a moderately strong hydrogen bond with the oxygen atom of the side chain of Q63. Thus the enol form of Q63 should be stabilized by the quite strong hydrogen bonding with the O–H group of Y21. It is thus likely that keto–enol tautomerism is the general mechanism for the BLUF domain to utilize light for signaling.

2.5.3 Photolyase/cryptochrome

The CRY/PHR family proteins are found in all three kingdoms of life, from bacteria to archaea to animals, including humans [12,162]. CRYs and PHRs share FAD chromophore bound to the PHR-like domain architecture (Fig. 2.1C). Despite structural homology, their physiological functions are quite distinct. PHRs repair ultraviolet (UV) light-induced photoproducts in DNA using near-UV and blue light with its cofactor. There are two types of PHRs in nature, CPD and (6-4) PHR, which repair cyclobutane pyrimidine dimer (CPD) and (6-4) photoproducts, respectively. In higher organisms PHR-like gene products, termed CRY to distinguish them from the DNA repair PHRs, function in a light-dependent manner to control growth and regulate flowering time in plants, and maintain circadian rhythms in animals.

In the CRY/PHR family of proteins the most stable redox state in vitro is their oxidized form (Fig. 2.10). Photoreduction is the activation process of CRY/PHR family proteins. The neutral radical state is probably the activation state of CRY for signaling, where light-induced electron and proton transfer to flavin takes place. Light-induced difference FTIR spectroscopy of CRY reported protein structural changes accompanying its activation. FTIR studies have also been instrumental in delineating the intramolecular proton transfer reaction that takes place during the activation process [33,163–165].

Light-induced difference FTIR spectroscopy was similarly used to study the activation processes of CPD and (6-4) PHRs, respectively [166–170]. The results show that the substrate-binding does not significantly affect protein structural changes by the photoreduction of FAD. In addition, FTIR study provided important information on why the radical state is stabilized in protein. The spread in the absorbance of the stable $FADH^{\bullet}$ radical (300–700 nm) enables PHR to efficiently form $FADH^{-}$, making it functional for DNA repair. FTIR spectroscopy of CPD PHR detected a strong hydrogen bond, from FAD N_5–H to the carbonyl group of the N378, that is modulated by the redox state of FAD [170]. The observed characteristic frequency shifts were reproduced in QM/MM studies of the flavin binding site, which were then employed to elucidate redox tuning governed by N378.

As PHR is an enzyme, it is possible to probe the enzymatic reaction by light-induced FTIR spectroscopy. In fact the first study clearly reported the difference spectra between substrate (CPD) and product (repaired DNA) of CPD PHR by measuring such enzymatic reactions with an excess amount of substrate [171]. In our recent study on (6-4) PHR, we measured the FTIR spectra under different molar ratios between enzyme and substrate [32,172–174]. When the molar ratio of (6-4) PHR and (6-4) photoproduct is 1:1, light-induced difference FTIR spectra correspond to [repaired DNA + unbound enzyme] minus [enzyme–substrate complex]. In contrast, with a 1:2 molar ratio, (6-4) photoproducts are half bound in the

dark. In this case, at the early stage of illumination, the repaired DNA is released into solution, while a new (6-4) photoproduct will bind to the empty PHR. The difference FTIR spectra correspond to the structural changes of DNA only, because PHR binds (6-4) photoproduct before and after illumination. On the other hand, at the late stage of illumination, a new (6-4) photoproduct cannot bind to the empty PHR after the repaired DNA is released into solution. Thus we expect no illumination time dependence of the spectra for the 1:1 mixture. On the other hand, we expect illumination time dependence for the 1:2 mixture, and it was indeed the case [32]. From these results, we were able to obtain the substrate binding signal to the enzyme. Highly accurate measurement of unlabeled and $^{15}N_3$,-labeled (6-4) photoproduct allowed to identify vibrational modes of (6-4) photoproduct upon repair, being sensitive enough to distinguish a single nitrogen atom (^{15}N versus ^{14}N) among $>$6700 atoms in the enzyme−substrate complex [172].

Despite similar structures, the functions of PHR are highly selective, which is evidenced by the reparation of CPD by CPD PHR, but not (6-4) photoproduct, or vice versa. Therefore we attempted to functionally swap CPD and (6-4) PHRs by site-directed mutagenesis. We found that a triple mutant of (6-4) PHR is able to repair the CPD photoproduct, although the repair efficiency is one order of magnitude lower than the wild-type CPD PHR [175]. Difference FTIR spectra for repairs demonstrate the lack of secondary structure alterations in the mutant, suggesting that the triple mutant gains substrate binding ability, but it does not release the repaired DNA. Interestingly, (6-4) photoproduct is not repaired by mutating CPD PHR. The observed asymmetric conversion suggests a more complex repair mechanism for (6-4) repair, which was also supported by QM/MM calculations [175,176]. These results suggest that CPD PHR might be the evolutionary origin of the PHR family of proteins.

2.6 Concluding remarks

In this chapter we have reviewed recent highlights of the use of light-induced difference FTIR spectroscopy on photoreceptive proteins. Although vibrational spectroscopy cannot determine the atomic position, its high sensitivity is extremely advantageous to elucidate the structure−function relationship in photoreceptive proteins. In particular, this method is useful in order to gain structural information on the protein and its interaction with the chromophore as well as with bound water molecules. Strong absorption of water in the IR region has been a bottleneck in our understanding of photoreceptive proteins. The ability to use FTIR in conjunction with controlled sample hydration have enabled data acquisition at the single water molecule level. This has led to a wealth of structural information on microbial rhodopsins, animal rhodopsins, and flavoproteins. With their emerging role in the field of optogenetics, obtaining a detailed mechanistic understanding of photoreceptive proteins will be of immense interest.

Acknowledgment

We thank Dr. Daichi Yamada for his assistance in figure preparation. We sincerely thank the many collaborators in the references.

References

[1] O.P. Ernst, D.T. Lodowski, M. Elstner, P. Hegemann, L.S. Brown, H. Kandori, Microbial and animal rhodopsin: structures, functions, and molecular mechanisms, Chem. Rev. 114 (2014) 126–163.
[2] S. Gozem, H.L. Luk, I. Schapiro, M. Olivucci, Theory and simulation of the ultrafast double-bond isomerization of biological chromophores, Chem. Rev. 117 (2017) 13502–13565.
[3] Y. Shichida, T. Matsuyama, Evolution of opsins and phototransduction, Philos. Trans. R. Soc. Lond. B Biol. Sci. 364 (2009) 2881–2895.
[4] K. Palczewski, Chemistry and biology of vision, J. Biol. Chem. 287 (2012) 1612–1619.
[5] K. Inoue, Y. Kato, H. Kandori, Light-driven ion-translocating rhodopsins in marine bacteria, Trends Microbiol. 23 (2015) 91–98.
[6] E.G. Govorunova, O.A. Sineshchekov, H. Li, J.L. Spudich, Microbial rhodopsins: diversity, mechanisms, and optogenetic applications, Annu. Rev. Biochem. 86 (2017) 845–872.
[7] N.C. Rockwell, J.C. Lagarias, Phytochrome structure and signaling mechanisms, Annu. Rev. Plant Biol. 57 (2006) 837–858.
[8] A. Nagatani, Phytochrome: structural basis for its functions, Curr. Opin. Plant Biol. 13 (2010) 565–570.
[9] A.T. Ulijasz, R.D. Vierstra, Phytochrome structure and photochemistry: recent advances toward a complete molecular picture, Curr. Opin. Plant Biol. 14 (2011) 498–506.
[10] Y. Imamoto, M. Kataoka, Structure and photoreaction of photoactive yellow protein, a structural prototype of the PAS domain superfamily, Photochem. Photobiol. 83 (2007) 40–49.
[11] T.E. Meyer, J.A. Kyndt, S. Memmi, T. Moser, B. Colón-Acevedo, B. Devreese, et al., The growing family of photoactive yellow proteins and their presumed functional roles, Photochem. Photobiol. Sci. 11 (2012) 1495–1514.
[12] A. Sancar, Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors, Chem. Rev. 103 (2003) 2203–2237.
[13] A. Losi, W. Gärtner, The evolution of flavin-binding photoreceptors: an ancient chromophore serving trendy blue-light sensors, Annu. Rev. Plant Biol. 63 (2012) 49–72.
[14] J.M. Christie, J. Gawthorne, G. Young, N.J. Fraser, A.J. Roe, LOV to BLUF: flavoprotein contributions to the optogenetic toolkit, Mol. Plant. 5 (2012) 533–544.
[15] D. Zhong, Electron transfer mechanisms of DNA repair by photolyase, Annu. Rev. Phys. Chem. 66 (2015) 691–715.
[16] L.O. Essen, S. Franz, A. Banerjee, Structural and evolutionary aspects of algal blue light receptors of the cryptochrome and aureochrome type, J. Plant Physiol. 217 (2017) 27–37.
[17] T. Kottke, A. Xie, D.S. Larsen, W.D. Hoff, Photoreceptors take charge: emerging principles for light sensing, Annu. Rev. Biophys. 47 (2018) 291–313.
[18] N. Nelson, Photosystems and global effects of oxygenic photosynthesis, Biochim. Biophys. Acta 1807 (2011) 856–863.
[19] J. Yano, V. Yachandra, Mn4Ca cluster in photosynthesis: where and how water is oxidized to dioxygen, Chem. Rev. 114 (2014) 4175–4205.
[20] J.R. Shen, The structure of photosystem II and the mechanism of water oxidation in photosynthesis, Annu. Rev. Plant Biol. 66 (2015) 23–48.
[21] T. Mirkovic, E.E. Ostroumov, J.M. Anna, et al., Light absorption and energy transfer in the antenna complexes of photosynthetic organisms, Chem. Rev. 117 (2017) 249–293.

[22] H. Luecke, B. Schobert, H.-T. Richter, J.P. Cartailler, J.K. Lanyi, Structure of bacteriorhodopsin at 1.55 Å resolution, J. Mol. Biol. 291 (1999) 899–911.
[23] K. Palczewski, T. Kumasaka, T. Hori, C.A. Behnke, H. Motoshima, B.A. Fox, et al., Crystal structure of rhodopsin: a G protein-coupled receptor, Science 289 (2000) 739–745.
[24] M.J. Maul, T.R. Barends, A.F. Glas, M.J. Cryle, T. Domratcheva, S. Schneider, et al., Crystal structure and mechanism of a DNA (6 − 4) photolyase, Angew. Chem. Int. Ed. Engl. 47 (2008) 10076–10080.
[25] H. Kandori, Role of internal water molecules in bacteriorhodopsin, Biochim. Biophys. Acta 1460 (2000) 177–191.
[26] C. Zscherp, A. Barth, Reaction-induced infrared difference spectroscopy for the study of protein reaction mechanisms, Biochemistry 40 (2001) 1875–1883.
[27] C. Kötting, K. Gerwert, Proteins in action monitored by time-resolved FTIR spectroscopy, Chemphyschem. 6 (2005) 881–888.
[28] T. Kottke, P. Hegemann, B. Dick, J. Heberle, The photochemistry of the light-, oxygen-, and voltage-sensitive domains in the algal blue light receptor phot, Biopolymers 82 (2006) 373–378.
[29] T. Noguchi, Fourier transform infrared difference and time-resolved infrared detection of the electron and proton transfer dynamics in photosynthetic water oxidation, Biochim. Biophys. Acta 1847 (2015) 35–45.
[30] T. Kottke, V.A. Lórenz-Fonfría, J. Heberle, The grateful infrared: sequential protein structural changes resolved by infrared difference spectroscopy, J. Phys. Chem. B 121 (2017) 335–350.
[31] H. Kandori, D. Yamada, FTIR spectroscopy of flavin-binding photoreceptors, Methods Mol. Biol. 2014 (1146) (2014) 361–376.
[32] Y. Zhang, T. Iwata, J. Yamamoto, K. Hitomi, S. Iwai, T. Todo, et al., FTIR study of light-dependent activation and DNA repair processes of (6-4) photolyase, Biochemistry 50 (2011) 3591–3598.
[33] T. Iwata, Y. Zhang, K. Hitomi, E.D. Getzoff, H. Kandori, Key dynamics of conserved asparagine in a cryptochrome/photolyase family protein by Fourier transform infrared spectroscopy, Biochemistry 49 (2010) 8882–8891.
[34] H. Kandori, N. Kinoshita, Y. Shichida, A. Maeda, Protein structural changes in bacteriorhodopsin upon photoisomerization as revealed by polarized FTIR spectroscopy, J. Phys. Chem. B 102 (1998) 7899–7905.
[35] V.A. Lorenz-Fonfria, Y. Furutani, H. Kandori, Active internal waters in the bacteriorhodopsin photocycle. A comparative study of the L and M intermediates at room and cryogenic temperatures by infrared spectroscopy, Biochemistry 47 (2008) 4071–4081.
[36] J.K. Lanyi, G. Váró, The photocycles of bacteriorhodopsin, Israel J. Chem. 35 (1995) 365–385.
[37] S.O. Smith, J. Lugtenburg, R.S. Mathies, Determination of retinal chromophore structure in bacteriorhodopsin with resonance Raman spectroscopy, J. Membrane Biol. 85 (1985) 95–109.
[38] A. Maeda, J. Sasaki, Y. Shichida, Yoshizawa, Water structural changes in the bacteriorhodopsin photocycle: analysis by Fourier transform infrared spectroscopy, Biochemistry 31 (1992) 462–467.
[39] H. Kandori, Y. Yamazaki, J. Sasaki, R. Needleman, J.K. Lanyi, A. Maeda, Water-mediated proton transfer in proteins: an FTIR study of bacteriorhodopsin, J. Am. Chem. Soc. 117 (1995) 2118–2119.
[40] Y. Yamazaki, M. Hatanaka, H. Kandori, J. Sasaki, W.F. Jan Karstens, J. Raap, et al., Water structural changes at the proton uptake site (the Thr46-Asp96 domain) in the L intermediate of bacteriorhodopsin, Biochemistry 34 (1995) 7088–7093.
[41] L.S. Brown, G. Varo, M. Hatanaka, J. Sasaki, H. Kandori, A. Maeda, et al., The complex extracellular domain regulates the transient decrease in the pKa of the retinal Schiff base during the bacteriorhodopsin photocycle, Biochemistry 34 (1995) 12903–12911.
[42] Y. Yamazaki, S. Tuzi, H. Saito, H. Kandori, R. Needleman, J.K. Lanyi, et al., Hydrogen-bonds of water and C = O groups coordinate long-range structural changes in the L photointermediate of bacteriorhodopsin, Biochemistry 35 (1996) 4063–4068.
[43] M. Hatanaka, J. Sasaki, H. Kandori, T.G. Ebrey, R. Needleman, J.K. Lanyi, et al., Effects of arginine-82 on the interactions of internal water molecules in bacteriorhodopsin, Biochemistry 35 (1996) 6308–6312.

[44] M. Hatanaka, R. Kashima, H. Kandori, N. Friedman, M. Sheves, R. Needleman, et al., Trp86 -> Phe replacement in bacteriorhodopsin affects a water molecule near Asp85 and light adaptation, Biochemistry 36 (1997) 5493−5498.
[45] M. Hatanaka, H. Kandori, A. Maeda, Location and orientation of functional water molecules in bacteriorhodopsin as revealed by polarized FT-IR spectroscopy, Biophys. J. 73 (1997) 1001−1006.
[46] Y. Yamazaki, H. Kandori, R. Needleman, J.K. Lanyi, A. Maeda, Interaction of the protonated Schiff base with the peptide backbone of valine 49 and the intervening water molecule in the N photointermediate of bacteriorhodopsin, Biochemistry 37 (1998) 1559−1564.
[47] D. Eisenburg, Kauzmann, The Structure and Properties of Water, Oxford Press, London, 1969.
[48] D.N. Glew, N.S. Rath, H_2O, HDO, and CH_3OH infrared spectra and correlation with solvent basicity and hydrogen bonding, Can. J. Chem. 49 (1971) 837−856.
[49] H. Kandori, Polarized FTIR spectroscopy distinguishes peptide backbone changes in the M and N photointermediates of bacteriorhodopsin, J. Am. Chem. Soc. 120 (1998) 4546−4547.
[50] H. Kandori, N. Kinoshita, Y. Shichida, A. Maeda, R. Needleman, J.K. Lanyi, Cysteine S-H as a hydrogen bonding probe in proteins, J. Am. Chem. Soc. 120 (1998) 5828−5829.
[51] H. Kandori, N. Kinoshita, Y. Yamazaki, A. Maeda, Y. Shichida, R. Needleman, et al., Structural change of threonine 89 upon photoisomerization in bacteriorhodopsin as revealed by polarized FTIR spectroscopy, Biochemistry 38 (1999) 9676−9683.
[52] H. Kandori, N. Kinoshita, Y. Yamazaki, A. Maeda, Y. Shichida, R. Needleman, et al., Local and distant protein structural changes upon photoisomerization of the retinal in bacteriorhodopsin, Proc. Natl Acad. Sci. USA 97 (2000) 4643−4648.
[53] H. Kandori, Y. Yamazaki, Y. Shichida, M. Belenky, J. Raap, J. Lugtenburg, et al., Tight Asp85-Thr89 association during the pump switch of bacteriorhodopsin, Proc. Natl Acad. Sci. USA 98 (2001) 1571−1576.
[54] H. Kandori, M. Belenky, J. Herzfeld, Vibrational frequency and dipolar orientation of the protonated Schiff base in bacteriorhodopsin before and after photoisomerization, Biochemistry 41 (2002) 6026−6031.
[55] T. Tanimoto, M. Shibata, M. Belenky, J. Herzfeld, H. Kandori, Altered hydrogen bonding of Arg82 during the proton pump cycle of bacteriorhodopsin: A low-temperature polarized FTIR spectroscopic study, Biochemistry 43 (2004) 9439−9447.
[56] N. Mizuide, M. Shibata, N. Friedman, M. Sheves, M. Belenky, J. Herzfeld, et al., Structural changes in bacteriorhodopsin following retinal photoisomerization from the 13-cis form, Biochemistry 45 (2006) 10674−10681.
[57] H. Kandori, Hydration switch model for the proton transfer in the Schiff base region of bacteriorhodopsin, Biochim. Biophys. Acta 1658 (2004) 72−79.
[58] M. Shibata, H. Kandori, FTIR studies of internal water molecules in the Schiff base region of bacteriorhodopsin, Biochemistry 44 (2005) 7406−7413.
[59] H. Kandori, Y. Shichida, Direct observation of the bridged water stretching vibrations inside a protein, J. Am. Chem. Soc. 122 (2000) 11745−11746.
[60] T. Tanimoto, Y. Furutani, H. Kandori, Structural changes of water in the Schiff base region of bacteriorhodopsin: proposal of a hydration switch model, Biochemistry 42 (2003) 2300−2306.
[61] M. Shibata, T. Tanimoto, H. Kandori, Water molecules in the Schiff base region of bacteriorhodopsin, J. Am. Chem. Soc. 125 (2003) 13312−13313.
[62] S. Hayashi, E. Tajkhorshid, H. Kandori, K. Schulten, Role of hydrogen-bond network in energy strage of bacteriorhodopsin's light-driven proton pump revealed by ab initio normal-mode analysis, J. Am. Chem. Soc. 126 (2004) 10516−10517.
[63] M. Shibata, K. Ihara, H. Kandori, Hydrogen-bonding interaction of the protonated Schiff base with halides in a chloride-pumping bacteriorhodopsin mutant, Biochemistry 45 (2006) 10633−10640.
[64] M. Shibata, M. Yoshitsug, N. Mizuide, K. Ihara, H. Kandori, Halide binding by the D212N mutant of bacteriorhodopsin affects hydrogen bonding of water in the active site, Biochemistry 46 (2007) 7525−7535.

[65] V.A. Lorenz-Fonfria, H. Kandori, Spectroscopic and kinetic evidence on how bacteriorhodopsin accomplishes vectorial proton transport under functional conditions, J. Am. Chem. Soc. 131 (2009) 5891–5901.
[66] L.S. Brown, J. Sasaki, H. Kandori, A. Maeda, R. Needleman, J.K. Lanyi, Glutamic acid 204 is the terminal proton release group at the extracellular surface of bacteriorhodopsin, J. Biol. Chem. 270 (1995) 27122–27126.
[67] R. Rammelsberg, G. Huhn, M. Lübben, K. Gerwert, Bacteriorhodopsin's intramolecular proton-release pathway consists of a hydrogen-bonded network, Biochemistry 37 (1998) 5001–5009.
[68] O. Béjà, L. Aravind, E.V. Koonin, M.T. Suzuki, A. Hadd, L.P. Nguyen, et al., Bacterial rhodopsin: evidence for a new type of phototrophy in the sea, Science 289 (2000) 1902–1906.
[69] G. Nagel, D. Ollig, M. Fuhrmann, S. Kateriya, A.M. Musti, E. Bamberg, et al., Channelrhodopsin-1: a light-gated proton channel in green algae, Science 296 (2002) 2395–2398.
[70] G. Nagel, T. Szellas, W. Huhn, S. Kateriya, N. Adeishvili, P. Berthold, et al., Channelrhodopsin-2, a directly light-gated cation-selective membrane channel, Proc. Natl Acad. Sci. USA 100 (2003) 13940–13945.
[71] E.S. Boyden, F. Zhang, E. Bamberg, G. Nagel, K. Deisseroth, Millisecond-timescale, genetically targeted optical control of neural activity, Nat. Neurosci. 8 (2005) 1263–1268.
[72] F. Zhang, L.P. Wang, M. Brauner, J.F. Liewald, K. Kay, N. Watzke, et al., Multimodal fast optical interrogation of neural circuitry, Nature 446 (2007) 633–639.
[73] K. Deisseroth, Optogenetics, Nat. Methods 8 (2011) 26–29.
[74] P. Hegemann, A. Möglich, Channelrhodopsin engineering and exploration of new optogenetic tools, Nat. Methods 8 (2011) 39–42.
[75] D. Ikeda, Y. Furutani, H. Kandori, FTIR study of the retinal Schiff base and internal water molecules of proteorhodopsin, Biochemistry 46 (2007) 5365–5373.
[76] K. Hashimoto, A.R. Choi, Y. Furutani, K.-H. Jung, H. Kandori, Low-temperature FTIR study of *Gloeobacter* rhodopsin: presence of strongly hydrogen-bonded water and long-range structural protein perturbation upon retinal photoisomerization, Biochemistry 49 (2010) 3343–3350.
[77] M. Sumii, Y. Furutani, S.A. Waschuk, L.S. Brown, H. Kandori, Strongly hydrogen-bonded water molecule present near the retinal chromophore of *Leptosphaeria* rhodopsin, the bacteriorhodopsin-like proton pump from a Eukaryote, Biochemistry 44 (2005) 15159–15166.
[78] K. Muroda, K. Nakashima, M. Shibata, M. Demura, H. Kandori, Protein-bound water as the determinant of asymmetric functional conversion between light-driven proton and chloride pumps, Biochemistry 51 (2012) 4677–4684.
[79] M. Shibata, N. Muneda, K. Ihara, T. Sasaki, M. Demura, H. Kandori, Internal water molecules of light-driven chloride pump proteins, Chem. Phys. Lett. 392 (2004) 330–333.
[80] M. Shibata, N. Muneda, T. Sasaki, K. Shimono, N. Kamo, M. Demura, et al., Hydrogen-bonding alterations of the protonated Schiff base and water molecule in the chloride pump of Natronobacterium pharaonis, Biochemistry 44 (2005) 12279–12286.
[81] N. Muneda, M. Shibata, M. Demura, H. Kandori, Internal water molecules of the proton-pumping halorhodopsin in the presence of azide, J. Am. Chem. Soc. 128 (2006) 6294–6295.
[82] Y. Furutani, A. Kawanabe, K.-H. Jung, H. Kandori, FTIR spectroscopy of the all-trans form of *Anabaena* sensory rhodopsin at 77 K: Hydrogen bond of a water between the Schiff base and Asp75, Biochemistry 44 (2005) 12287–12296.
[83] A. Kawanabe, Y. Furutani, K.-H. Jung, H. Kandori, FTIR study of the photoisomerization processes in the 13-cis and all-trans forms of *Anabaena* sensory rhodopsin at 77 K, Biochemistry 45 (2006) 4362–4370.
[84] Y. Furutani, A.G. Bezarra Jr, S. Waschuk, M. Sumii, L.S. Brown, H. Kandori, FTIR spectroscopy of the K photointermediate of *Neurospora* rhodopsin; structural changes of the retinal, protein, and water molecules after photoisomerization, Biochemistry 43 (2004) 9636–9646.

[85] Y. Furutani, Y. Shichida, H. Kandori, Structural changes of water molecules during the photoactivation process in bovine rhodopsin, Biochemistry 42 (2003) 9619–9625.
[86] T. Ota, Y. Furutani, A. Terakita, Y. Shichida, H. Kandori, Structural changes in the Schiff base region of squid rhodopsin upon photoisomerization studied by low-temperature FTIR spectroscopy, Biochemistry 45 (2006) 2845–2851.
[87] K. Katayama, Y. Furutani, H. Imai, H. Kandori, Protein-bound water molecules in primate red- and green-sensitive visual pigments, Biochemistry 51 (2012) 1126–1133.
[88] D. Suzuki, Y. Sudo, Y. Furutani, T. Takahashi, M. Homma, H. Kandori, Structural changes of *Salinibacter* sensory rhodopsin I upon formation of the K and M photointermediates, Biochemistry 47 (2008) 12750–12759.
[89] H. Kandori, Y. Furutani, K. Shimono, Y. Shichida, N. Kamo, Internal water molecules of pharaonis phoborhodopsin studied by low-temperature infrared spectroscopy, Biochemistry 40 (2001) 15693–15698.
[90] H. Ono, K. Inoue, R. Abe-Yoshizumi, H. Kandori, FTIR spectroscopy of a light-driven compatible sodium ion-proton pumping rhodopsin at 77 K, J. Phys. Chem. B 118 (2014) 4784–4792.
[91] Y. Nomura, S. Ito, M. Teranishi, H. Ono, K. Inoue, H. Kandori, Low-temperature FTIR spectroscopy provides evidence for protein-bound water molecules in eubacterial light-driven ion pumps, Phys. Chem. Chem. Phys. 20 (2018) 3165–3171.
[92] S. Ito, S. Sugita, K. Inoue, H. Kandori, FTIR analysis of a light-driven inward proton-pumping rhodopsin at 77 K, Photochem. Photobiol. 93 (2017) 1381–1387.
[93] S. Ito, H.E. Kato, R. Taniguchi, T. Iwata, O. Nureki, H. Kandori, Water-containing hydrogen-bonding network in the active center of channelrhodopsin, J. Am. Chem. Soc. 136 (2014) 3475–3482.
[94] J. Sasaki, L.S. Brown, Y.-S. Chon, H. Kandori, A. Maeda, R. Needleman, et al., Conversion of bacteriorhodopsin into a chloride ion pump, Science 269 (1995) 73–75.
[95] K. Inoue, Y. Nomura, H. Kandori, Asymmetric functional conversion of eubacterial light-driven ion pumps, J. Biol. Chem. 291 (2016) 9883–9893.
[96] T. Okada, M. Sugihara, A.N. Bonder, M. Elstner, P. Entel, V. Buss, The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure, J. Mol. Biol. 342 (2004) 571–583.
[97] M. Murakami, T. Kouyama, Crystal structure of squid rhodopsin, Nature 453 (2008) 363–367.
[98] H.W. Choe, Y.J. Kim, J.H. Park, T. Morizumi, E.F. Pai, Krauß, et al., Crystal structure of metarhodopsin II, Nature 471 (2011) 651–655.
[99] J. Standfuss, P.C. Edwards, A. D'Antona, M. Fransen, G. Xie, D.D. Oprian, et al., Crystal structure of constitutively active rhodopsin: how and agonist can activate its GPCR, Nature 471 (2011) 656–660.
[100] R. Vogel, T.P. Sakmar, M. Sheves, F. Siebert, Coupling of protonation switches during rhodopsin activation, Photochem. Photobiol. 83 (2007) 286–292.
[101] E. Ritter, M. Elgeti, F.J. Bartl, Activity switches of rhodopsin, Photochem. Photobiol. 84 (2008) 911–920.
[102] S. Nishimura, J. Sasaki, H. Kandori, T. Matsuda, Y. Fukada, A. Maeda, Structural changes in the peptide backbone in complex formation between activated rhodopsin and transducin studied by FTIR spectroscopy, Biochemistry 35 (1996) 13267–13271.
[103] S. Nishimura, H. Kandori, A. Maeda, Interaction between photoactivated rhodopsin and the C-terminal peptide of transducin a subunit studied by FTIR spectroscopy, Biochemistry 37 (1998) 15816–15824.
[104] T. Baasov, N. Friedman, M. Sheves, Factors affecting the C:N stretching in protonated retinal Schiff base: a model study for bacteriorhodopsin and visual pigments, Biochemistry 26 (1987) 3210–3217.
[105] I. Palings, E.M.M. van den Berg, J. Lugtenburg, R.A. Mathies, Complete assignment of the hydrogen out-of-plane wagging vibrations of bathorhodposin: chromophore structure and energy storage in the primary photoproduct of vision, Biochemistry 28 (1989) 1498–1507.
[106] U.M. Ganter, E.D. Schmid, D. Perez-Sala, R.R. Rando, F. Siebert, Removal of the 9-methyl group of retinal inhibits signal transduction in the visual process. A fourier transform infrared and biochemical investigation, Biochemistry 28 (1989) 5954–5962.

[107] K. Fahmy, F. Jäger, M. Beck, T.A. Zvyaga, T.P. Sakmar, F. Siebert, Protonation states of membrane-embedded carboxylic acid groups in rhodopsin and metarhodopsin II: a Fourier-transform infrared spectroscopy study of site directed mutants, Proc. Natl Acad. Sci. USA 90 (1993) 10206–10210.
[108] P. Rath, L.L. DeCaluwe, P.H. Bovee-Geurts, W.J. Degrip, K.J. Rothschild, Fourier transform infrared difference spectroscopy of rhodopsin mutants: light activation of rhodopsin causes hydrogen-bonding change in residue aspartic acid-83 during meta II formation, Biochemistry 32 (1993) 10277–10282.
[109] F. Jäger, K. Fahmy, T.P. Sakmar, F. Siebert, Identification of glutamic acid 113 as the Schiff base proton acceptor in the metarhodopsin II photointermediate of rhodopsin, Biochemistry 33 (1994) 10878–10882.
[110] T. Nagata, A. Terakita, H. Kandori, D. Kojima, Y. Shichida, A. Maeda, Water and peptide backbone structure in the active center of bovine rhodopsin, Biochemistry 36 (1997) 6164–6170.
[111] T. Nagata, A. Terakita, H. Kandori, Y. Shichida, A. Maeda, The hydrogen bonding network of water mlecules and the peptide backbone in the region connecting Asp83, Gly120 and Glu113 in bovine rhodopsin, Biochemistry 37 (1998) 17216–17222.
[112] Y. Furutani, Y. Shichida, H. Kandori, Structural changes of water molecules during the photoactivation process in bovine rhodopsin, Biochemistry 42 (2003) 9619–9625.
[113] D. Baylor, How photons start vision, Proc. Natl Acad. Sci. USA 93 (1996) 560–565.
[114] D.A. Baylor, G. Matthews, K.W. Yau, Two components of electrical dark noise in toad retinal rod outer segments, J. Physiol. 309 (1980) 591–621.
[115] M. Yanagawa, K. Kojima, T. Yamashita, Y. Imamoto, T. Matsuyama, K. Nakanishi, et al., Origin of the low thermal isomerization rate of rhodopsin chromophore, Sci. Rep. 5 (2015) 11081.
[116] S. Gozem, I. Schapiro, N. Ferré, M. Olivucci, The molecular mechanism of thermal noise in rod photoreceptors, Science 337 (2012) 1225–1228.
[117] H. Kandori, A. Maeda, FTIR spectroscopy reveals microscopic structural changes of the protein around the rhodopsin chromophore upon photoisomerization, Biochemistry 34 (1995) 14220–14229.
[118] T. Nagata, T. Oura, A. Terakita, H. Kandori, Y. Shichida, Isomer-specific interaction of the retinal chromophore with threonine-118 in rhodopsin, J. Phys. Chem. A 106 (2002) 1969–1975.
[119] V.A. Lorenz-Fonfria, Y. Furutani, T. Ota, K. Ido, H. Kandori, Protein fluctuations as the possible origin of the thermal activation of rod photoreceptors in the dark, J. Am. Chem. Soc. 132 (2010) 5693–5703.
[120] S. Ye, E. Zaitseva, G. Caltabiano, G.F.X. Schertler, T. Sakmar, X. Deupi, et al., Tracking G-protein-coupled receptor activation using genetically encoded infrared probes, Nature 464 (2010) 1386–1390.
[121] Y. Yamazaki, T. Nagata, A. Terakita, H. Kandori, Y. Shichida, Y. Imamoto, Intramolecular interactions that induce helical rearrangement upon rhodopsin activation: light-induced structural changes in metarhodopsin IIa probed by cysteine S-H stretching vibrations, J. Biol. Chem. 289 (2014) 13792–13800.
[122] G. Wald, Molecular basis of visual excitation, Science 162 (1968) 230–239.
[123] J. Nathans, D. Thomas, D.S. Hogness, Molecular genetics of human color vision: the genes encoding blue, green, and red pigments, Science 232 (1986) 193–201.
[124] T.P. Sakmar, R.R. Franke, H.G. Khorana, The role of the retinylidene Schiff base counterion in rhodopsin in determining wavelength absorbance and Schiff base pKa, Proc. Natl Acad. Sci. USA 88 (1991) 3079–3083.
[125] S. Ahuja, M. Eilers, A. Hirshfeld, E.C. Yan, M. Ziliox, T.P. Sakmar, et al., 6-s-cis conformation and polar binding pocket of the retinal chromophore in the photoactivated state of rhodopsin, J. Am. Chem. Soc. 131 (2009) 15160–15169.
[126] S. Sekharan, K. Katayama, H. Kandori, K. Morokuma, Color vision: "OH-site" rule for seeing red and green, J. Am. Chem. Soc. 134 (2012) 10706–10712.
[127] G.G. Kochendoerfer, Z. Wang, D.D. Oprian, R.A. Mathies, Resonance Raman examination of the wavelength regulation mechanism in human visual pigments, Biochemisry 36 (1997) 6577–6587.
[128] T. Hirano, H. Imai, H. Kandori, Y. Shichida, Chloride effect on iodopsin studied by low-temperature visible and infrared spectroscopies, Biochemistry 40 (2001) 1385–1392.

[129] K. Katayama, Y. Furutani, H. Imai, H. Kandori, An FTIR study of monkey green- and red-sensitive visual pigments, Angew. Chem. Int. Ed. 49 (2010) 891–894.
[130] K. Katayama, Y. Nonaka, K. Tsutsui, H. Imai, H. Kandori, Spectral tuning mechanism of primate blue-sensitive visual pigment elucidated by FTIR spectroscopy, Sci. Rep. 7 (2017) 4904–4914.
[131] K. Katayama, T. Okitsu, H. Imai, A. Wada, H. Kandori, Identical hydrogen-bonding strength of the retinal Schiff base between primate green- and red-sensitive pigments: new insight into color tuning mechanism, J. Phys. Chem. Lett. 6 (2015) 1130–1133.
[132] D. Tischer, O.D. Weiner, Illuminating cell signalling with optogenetic tools, Nat. Rev. Mol. Cell. Biol. 15 (2014) 551–558.
[133] C.L. Tucker, J.D. Vrana, M.J. Kennedy, Tools for controlling protein interactions using light, Curr. Protoc. Cell. Biol. 64 (2014) 1–20.
[134] K. Zhang, B. Cui, Optogenetic control of intracellular signaling pathways, Trends Biotechnol. 33 (2015) 92–100.
[135] D. Edmondson, S. Ghisla, Flavoenzyme structure and function, in: S.K. Chapman, G.A. Reid (Eds.), Methods in Molecular Biology, vol. 131; Flavoprotein Protocols, Humana Press, Inc, 1999, pp. 157–179.
[136] J.M. Christie, Phototropin blue-light receptors, Annu. Rev. Plant Biol. 58 (2007) 21–45.
[137] T.E. Swartz, P.J. Wenzel, S.B. Corchnoy, W.R. Briggs, R.A. Bogomolni, Vibration spectroscopy reveals light-induced chromophore and protein structural changes in the LOV2 domain of the plant blue-light receptor phototropin 1, Biochemistry 41 (2002) 7183–7189.
[138] T. Iwata, S. Tokutomi, H. Kandori, Photoreaction of the cysteine S-H group in the LOV2 domain of Adiantum phytochrome3, J. Am. Chem. Soc. 124 (2002) 11840–11841.
[139] Y. Sato, T. Iwata, S. Tokutomi, H. Kandori, Reactive cysteine is protonated in the triplet excited state of the LOV2 domain in *Adiantum* phytochrome3, J. Am. Chem. Soc. 127 (2005) 1088–1089.
[140] T. Iwata, D. Nozaki, S. Tokutomi, T. Kagawa, M. Wada, H. Kandori, Light-induced structural changes in the LOV2 domain of Adiantum phytochrome3 studied by low-temperature FTIR and UV-visible spectroscopy, Biochemistry 42 (2003) 8183–8191.
[141] D. Nozaki, T. Iwata, T. Ishikawa, T. Todo, S. Tokutomi, H. Kandori, Role of Gln1029 in the photoactivation processes of the LOV2 domain in *Adiantum* phytochrome3, Biochemistry 43 (2004) 8373–8379.
[142] T. Iwata, D. Nozaki, S. Tokutomi, H. Kandori, Comparative investigation of the LOV1 and LOV2 domains in *Adiantum* phytochrome3, Biochemistry 44 (2005) 7427–7434.
[143] T. Iwata, D. Nozaki, Y. Sato, K. Sato, Y. Nishina, K. Shiga, et al., Identification of the C = O stretching vibrations of FMN and peptide backbone by 13C-labeling of the LOV2 domain of *Adiantum* phytochrome3, Biochemistry 45 (2006) 15384–15391.
[144] A. Yamamoto, T. Iwata, S. Tokutomi, H. Kandori, Role of Phe1010 in light-induced structural changes of the neo1-LOV2 domain of *Adiantum*, Biochemistry 47 (2008) 922–928.
[145] A. Yamamoto, T. Iwata, Y. Sato, D. Matsuoka, S. Tokutomi, H. Kandori, Light signal transduction pathway from flavin chromophore to the Jα helix of *Arabidopsis* phototropin 1, Biophys. J. 96 (2009) 2771–2778.
[146] T. Koyama, T. Iwata, A. Yamamoto, Y. Sato, D. Matsuoka, S. Tokutomi, et al., Different role of the Jα helix in the light-induced activation of the LOV2 domains in various phototropins, Biochemistry 48 (2009) 7621–7628.
[147] T. Iwata, D. Nozaki, A. Yamamoto, T. Koyama, Y. Nishina, K. Shiga, et al., Hydrogen bonding environment of the N3-H group of flavin mononucleotide in the light oxygen voltage domains of phototropins, Biochemistry 56 (2017) 3099–3108.
[148] T. Bednarz, A. Losi, W. Gartner, P. Hegemann, J. Heberle, Functional variations among LOV domains as revealed by FT-IR difference spectroscopy, Photochem. Photobiol. Sci. 3 (2004) 575–579.
[149] A. Pfeifer, T. Majerus, K. Zikihara, D. Matsuoka, S. Tokutomi, J. Heberle, et al., Time-resolved Fourier transform infrared study on photoadduct formation and secondary structural changes within the phototropin LOV domain, Biophys. J. 96 (2009) 1462–1470.

[150] A. Pfeifer, T. Mathes, Y. Lu, P. Hegemann, T. Kottke, Blue light induces global and localized conformational changes in the kinase domain of full-length phototropin, Biochemistry 49 (2010) 1024–1032.
[151] E. Herman, T. Kottke, Allosterically regulated unfolding of the A'α helix exposes the dimerization site of the blue-light-sensing aureochrome-LOV domain, Biochemistry 54 (2015) 1484–1492.
[152] J.N. Iuliano, A.A. Gil, S.P. Laptenok, C.R. Hall, J. Tolentino Collado, A. Lukacs, et al., Variation in LOV photoreceptor activation dynamics probed by time-resolved infrared spectroscopy, Biochemistry 57 (2018) 620–630.
[153] M. Gauden, I.H.M. van Stokkum, J.M. Key, D.C. Lührs, R. van Grondelle, P. Hegemann, et al., Hydrogen-bond switching through a radical pair mechanism in a flavin-binding photoreceptor, Proc. Natl Acad. Sci. USA 103 (2006) 10895–10900.
[154] W. Laan, M.A. van der Horst, I.H. van Stokkum, K.J. Hellingwerf, Initial characterization of the primary photochemistry of AppA, a blue-light-using flavin adenine dinucleotide-domain containing transcriptional antirepressor protein from *Rhodobacter sphaeroides*: a key role for reversible intramolecular proton transfer from the flavin adenine dinucleotide chromophore to a conserved tyrosine? Photochem. Photobiol. 78 (2003) 290–297.
[155] S. Masuda, K. Hasegawa, T.A. Ono, Light-induced structural changes of apoprotein and chromophore in the sensor of blue light using FAD (BLUF) domain of AppA for a signaling state, Biochemistry 44 (2005) 1215–1224.
[156] J.S. Grinstead, M. Avila-Perez, K.J. Hellingwerf, R. Boelens, R. Kaptein, Light-induced flipping of a conserved glutamine sidechain and its orientation in the AppA BLUF domain, J. Am. Chem. Soc. 128 (2006) 15066–15067.
[157] R. Takahashi, K. Okajima, H. Suzuki, H. Nakamura, M. Ikeuchi, T. Noguchi, FTIR study on the hydrogen bond structure of a key tyrosine residue in the flavin-binding blue light sensor TePixD from *Thermosynechococcus elongatus*, Biochemistry 46 (2007) 6459–6467.
[158] A.L. Stelling, K.L. Ronayne, J. Nappa, P.J. Tonge, S.R. Meech, Ultrafast structural dynamics in BLUF domains: transient infrared spectroscopy of AppA and its mutants, J. Am. Chem. Soc. 129 (2007) 15556–15564.
[159] T. Iwata, A. Watanabe, M. Iseki, M. Watanabe, H. Kandori, Strong donation of the hydrogen bond of tyrosine during photoactivation of the BLUF domain, J. Phys. Chem. Lett. 2 (2011) 1015–1019.
[160] T. Domratcheva, E. Hartmann, I. Schlichting, T. Kottke, Evidence for tautomerisation of glutamine in BLUF blue light receptors by vibrational spectroscopy and computational chemistry, Sci. Rep. 6 (2016) 22669.
[161] T. Iwata, T. Nagai, S. Ito, S. Osoegawa, M. Iseki, M. Watanabe, et al., Hydrogen bonding environments in the photocycle process around the flavin chromophore of the AppA-BLUF domain, J. Am. Chem. Soc. 140 (2018) 11982–11991.
[162] C. Lin, T. Todo, The cryptochromes, Genome Biol. 6 (2005) 220.
[163] T. Kottke, A. Batschauer, M. Ahmad, J. Heberle, Blue-light-induced changes in *Arabidopsis* cryptochrome 1 probed by FTIR difference spectroscopy, Biochemistry 45 (2006) 2472–2479.
[164] A. Hense, E. Herman, S. Oldemeyer, T. Kottke, Proton transfer to flavin stabilizes the signaling state of the blue light receptor plant cryptochrome, J. Biol. Chem. 290 (3) (2015) 1743–1751.
[165] C. Thöing, S. Oldemeyer, T. Kottke, Microsecond deprotonation of aspartic acid and response of the α/β subdomain precede C-terminal signaling in the blue light sensor plant cryptochrome, J. Am. Chem. Soc. 137 (2015) 5990–5999.
[166] D. Yamada, Y. Zhang, T. Iwata, K. Hitomi, E.D. Getzoff, H. Kandori, FTIR study of the photoactivation process of Xenopus (6-4) photolyase, Biochemistry 51 (2012) 5774–5783.
[167] I.M.M. Wijaya, Y. Zhang, T. Iwata, J. Yamamoto, K. Hitomi, S. Iwai, et al., Detection of distinct α-helical rearrangements of CPD photolyase upon substrate binding by FTIR spectroscopy, Biochemistry 52 (2013) 1019–1027.

[168] I.M.M. Wijaya, T. Iwata, J. Yamamoto, K. Hitomi, S. Iwai, E.D. Getzoff, et al., Flavin adenine dinucleotide chromophore charge controls the conformation of cyclobutane pyrimidine dimer photolyase α-helices, Biochemistry 53 (2014) 5864−5875.
[169] D. Yamada, J. Yamamoto, Y. Zhang, T. Iwata, K. Hitomi, E.D. Getzoff, et al., Structural changes of the active center during the photoactivation of Xenopus (6-4) photolyase, Biochemistry 55 (2016) 715−723.
[170] I.M. Mahaputra Wijaya, T. Domratcheva, T. Iwata, E.D. Getzoff, H. Kandori, Single hydrogen bond donation from flavin N5 to proximal asparagine ensures FAD reduction in DNA photolyase, J. Am. Chem. Soc. 138 (2016) 4368−4376.
[171] E. Schleicher, B. Hessling, V. Illarionova, A. Bacher, S. Weber, G. Richter, et al., Light-induced reactions of *Escherichia coli* DNA photolyase monitored by Fourier transform infrared spectroscopy, FEBS J. 272 (2005) 1855−1866.
[172] Y. Zhang, J. Yamamoto, D. Yamada, T. Iwata, K. Hitomi, T. Todo, et al., Substrate assignment of the (6-4) photolyase reaction by FTIR spectroscopy, J. Phys. Chem. Lett. 183 (2011) 2774−2777.
[173] I.M.M. Wijaya, T. Iwata, J. Yamamoto, K. Hitomi, S. Iwai, E.D. Getzoff, et al., FTIR study of CPD photolyase with substrate in single strand DNA, Biophysics 11 (2015) 39−45.
[174] D. Yamada, T. Iwata, J. Yamamoto, K. Hitomi, T. Todo, S. Iwai, et al., Structural role of two histidines in the (6-4) photolyase reaction, Biophys. Physicobiol. 12 (2015) 139−144.
[175] D. Yamada, H.M. Dokainish, T. Iwata, J. Yamamoto, T. Ishikawa, T. Todo, et al., Functional conversion of CPD and (6-4) photolyases by mutation, Biochemistry 55 (2016) 4173−4183.
[176] H.M. Dokainish, D. Yamada, T. Iwata, H. Kandori, A. Kitao, Electron fate and mutational robustness in the mechanism of (6-4) photolyase-mediated DNA repair, ACS Catal. 7 (2017) 4835−4845.

This page intentionally left blank

CHAPTER 3

Quantum cascade laser-based infrared transmission spectroscopy of proteins in solution

Andreas Schwaighofer and Bernhard Lendl

Institute of Chemical Technologies and Analytics, Vienna University of Technology, Vienna, Austria

3.1 Quantum cascade lasers and their advantages for mid-infrared transmission measurements

The recently emerging availability of quantum cascade lasers (QCLs) has opened up new and exciting possibilities for mid-infrared (mid-IR) spectroscopy in protein research. Experimental realization of the first QCL in 1994 was the starting point for the development of compact, room temperature-operated semiconductor lasers in the mid-IR region [1]. The working principle of QCLs is fundamentally different from conventional semiconductor lasers, where light is generated from an electron interband transition. Here, the high-energy electron in the conduction band recombines with a hole in the valence band. Consequently, for conventional lasers the emission wavelength is determined by the band gap energy of the utilized materials, thus restricting their application mainly to the UV−vis and near-infrared range. In contrast, QCLs are not composed of bulk semiconductor material in their optically active region, but made of a series of thin semiconductor layers of varying material composition. Upon application of an electrical field, electrons are injected into the upper state of the quantum well system. The ensuing relaxation of the radiative electron into a lower intersubband state generates mid-IR photons with an energy corresponding to the energy difference between initial and final state. While these optical transitions take place in the active regions, the injector regions maintain population inversion, which is achieved through extraction of electrons from the active region by resonant tunneling subsequent to the optical transition. The extraction process occurs approximately 10 times faster than the radiative transition, hence electrons accumulate in the successive injector region, thereby preserving population inversion. In QCLs a cascade of typically up to 40 repetitions of this characteristic active region-injection sequence is realized in the laser gain medium. Thus a single electron can undergo

Vibrational Spectroscopy in Protein Research.
DOI: https://doi.org/10.1016/B978-0-12-818610-7.00003-7

multiple radiative transitions generating mid-IR photons, in contrast to diode lasers, where each electron results in single-photon emission only [2]. Following this principle, the emission wavelength of QCLs is mainly determined by the thickness of the layers, rather than the material, and consequently allows tailoring of the emission wavelength of QCLs over a wide range in the same material system [3–5]. The quantum well structure of the gain medium can be designed so that amplification of the emitted radiation over a broad range of several hundreds of wave numbers is possible. Whereas the material system based on InGaAs/AlInAs heterostructure grown on InP substrate is the most mature technology and compatible with mass production requirements [6], alternative active region designs based on InAs/AlSb have been presented [7]. Most recently the first InAs/AlSb QCLs directly grown on a Si substrate were introduced [8].

There are available several types of QCLs for different spectroscopic purposes. On the one hand, there are distributed feedback (DFB)-QCLs, where the wave number is selected by an internal Bragg grating applied to the gain medium and which provide a wave number tunability of approximately 5 cm^{-1} by changing the injection current and/or the operation temperature [9]. Due to this narrow wavelength tunability, this type of laser is predominantly used for gas-phase analysis [10,11], however, multiple DFB-QCL could be combined in an array to achieve a wider spectral coverage, which would be beneficial for the analysis of liquid samples that show broader spectral features than gas-phase analytes [12]. For liquid-phase analysis, commonly external cavity (EC)-QCLs are employed that provide broadband spectral tuning by an external diffraction grating. Consequently, with this type of QCL complete broadband spectra can be acquired when operated in sweep or "step-and-measure" mode. Alternatively, measurement points for individual absorbance values or time traces at multiple wave numbers can be flexibly chosen in single wavelength mode. In sweep mode the laser emission wavelength is tuned across a set region. Tuning is achieved by a continuous slew motion of the diffraction grating relative to the QCL chip, which allows tuning ranges of up to several hundred wave numbers. The maximum spectral coverage of a single chip in current state-of-the-art EC-QCLs could be extended to a maximum of 556 cm^{-1} in the 3–4 nm region [13] and 432 cm^{-1} in the 7.6–11.4 nm region [14]. However, even larger tuning ranges of more than 1000 cm^{-1} can be provided by a single device by beam combination of up to four individual EC-QCL modules. Regarding the speeds for wavelength tuning, currently rates of up to 25,000 cm^{-1}/s (LaserTune, Block Engineering) and 5000 cm^{-1}/s (MIRcat-QT, Daylight Solutions) are commercially available. A drawback of the emission characteristics of EC-QCLs operated in sweep mode are mode-hops that can occur during tuning. Mode-hops originate from the competition of different optical modes for the available net gain in the laser medium. Consequently, the laser emission "hops" from one mode to another, although the gain spectrum itself is homogeneous, which leads to distinctive steps in the emission curve that appear in the measurement in the form of laser intensity variations [15]. In single

wavelength mode the grating is adjusted to a specific position so that the laser emits at the desired wavelength. When performing repeated scans with this "step-and-measure" emission mode, it may occur that the laser does not emit at the same light mode (i.e., exact wavelength) when aiming for the set nominal wavelength at every cycle, which adds to the measurable laser noise. As a further mid-IR light source based on this technology, most recently QCL frequency combs were introduced [16], which provide simultaneous emission of equally spaced lines (e.g., $\Delta cm^{-1} = 0.25$) with a well-defined phase relationship among themselves across a spectral region of up to 230 cm^{-1} [17].

These fundamentally different operation principles of QCLs provide unique properties in comparison to other mid-IR light sources. Unlike previously available mid-IR lasers, such as CO_2 and lead salt lasers, as well as light sources based on optical parametric generation (difference frequency generation; and optical parametric oscillation), QCLs combine wide tunability, stable operation at room temperature, and versatility in pulsed and continuous wave (CW) mode with modulation capability up to the MHz-regime. Compared to thermal light sources employed in conventional Fourier-transform infrared (FTIR) spectrometers, QCLs offer coherent and polarized light with an increased spectral power density by a factor of 10^4 and more.

The high available spectral power is beyond doubt the most exploited unique characteristic of QCLs for transmission measurements in the liquid phase. The spectral power density provided by QCLs (mW-W/cm^{-1} range) is several orders of magnitude higher than the one of globars employed in FTIR spectrometers (μW/cm^{-1} range) [18]. In direct absorption spectroscopy the increased laser power translates to a logarithmic increase of the absorbance by increasing the available path length according to Beer–Lambert's law. While mid-IR analysis in aqueous solutions is for most applications limited to path lengths in the range of tens of micrometers, this possible increase of the optical path is particularly beneficial for the analysis of proteins. Commonly employed path lengths for FTIR transmission measurements of the protein amide I band in aqueous solutions are below 10 μm to avoid total IR absorption in the region of the HOH-bending vibration [19]. This experimental constraint arises from the combination of two aspects: first, the HOH-bending band of water near 1643 cm^{-1} with a high molar absorption coefficient, which overlaps with the protein amide I band; and second, the low emission power provided by conventional thermal light sources (globars) that are used in FTIR spectrometers. Cell and sample handling under these conditions is rather laborious because thin spacers are extremely susceptible to electrostatic charging making the assembly of a tight flow cell difficult. Further, highly concentrated protein solutions are viscous and lead to the formation of air bubbles when filling the cell. As a consequence the low feasible path lengths for FTIR transmission measurements of proteins in aqueous solution are a considerable impairment for the robustness of analysis and impede flow-through measurements and high-throughput applications. In order to formulate a theoretical

estimation of the achievable improvement of the path length by employing a high intensity light source [20], the Beer–Lambert's law is followed, which states that the absorption A is proportional to the decadic molar absorption coefficient ε, the sample concentration c, and the interaction length d:

$$A = -\log\left(\frac{I}{I_0}\right) = \varepsilon \cdot c \cdot d$$

I_0 and I denote the recorded intensity of the background and sample measurement, respectively. Further considerations are based on the assumption that a certain minimum incident intensity I_D is required to reach the detector, thus focusing on the influence of the light source. Upon switching from a FTIR spectrometer with the source intensity ($I_{\mathrm{FTIR}} = I_0$) to a QCL-based setup, the intensity of the light source (I_{QCL}) is higher by the factor (f_I). Consequently, the maximum achievable optical path length d_{max} for a certain wavelength λ can be expressed as

$$d_{max,\mathrm{QCL}} = \frac{1}{\varepsilon c}\left(\log\frac{I_{\mathrm{FTIR}}}{I_D} + \log f_I\right) \text{with } I_{\mathrm{QCL}} = f_I \cdot I_{\mathrm{FTIR}}$$

Following from that, the maximum path length consists of two terms: the path length achievable with the FTIR spectrometer and an additional term considering the higher available power of the QCL. Assuming typical peak power levels of QCLs operated in pulsed mode (100 mW/cm^{-1}) and of a thermal emitter in a FTIR spectrometer (10 μW/cm^{-1}—designation per cm^{-1} to be comparable with QCLs), the power enhancement factor amounts to $f_{\mathrm{I}} = 10^4$. Consequently, the maximum achievable path length is calculated as

$$d_{max,\mathrm{QCL}} = \frac{1}{\varepsilon c}\left(\log\frac{I_{\mathrm{FTIR}}}{I_D} + \log f_I\right) = d_{max,\mathrm{FTIR}} + \frac{1}{\varepsilon c} \cdot 4 \text{ with } d_{max,\mathrm{FTIR}} + \frac{1}{\varepsilon c} \cdot \log\frac{I_{\mathrm{FTIR}}}{I_D}$$

When inserting the values for pure water as a sample substance with a molarity of 55.5 mol/L and a molar coefficient of 21.8 L/mol/cm^{-1} at 1643 cm^{-1} (maximum of the HOH-bending band in the protein amide I region) [21] and furthermore, a minimum required intensity meeting the detector of $I_D = 850$ nW (corresponding to an absorbance of 1.1), the maximum achievable optical path for FTIR spectrometers $d_{max,\mathrm{FTIR}} = 8.8$ μm and QCL-based setups $d_{max,\mathrm{QCL}} = 41.9$ μm is received. This derivation describes the higher signals in terms of absorbance obtainable with QCL-based IR spectroscopy due to the larger path length. However, for a full evaluation of the performance (i.e., SNR) in comparison to conventional FTIR spectroscopy also the noise level needs to be considered. In the latter case the noise level introduced by low intensity thermal light sources can usually be neglected compared to the detector noise. Furthermore, FTIR spectroscopy benefits from the Fellgett's advantage (all wavelengths passing through the sample simultaneously) that partially compensates for the low spectral power densities provided by thermal emitters.

High intensity laser sources, on the other hand, contribute to the overall measurement noise, particularly when operated in pulsed mode due to pulse-to-pulse intensity fluctuations. This partially cancels out the signal-enhancement due to the increase in optical path length enabled by QCLs in liquid-phase transmission measurements [22].

Most state-of-the-art QCLs can be operated in CW and pulsed mode at room temperature without cryogenic cooling but employing thermoelectric temperature control and additional water cooling when operated in CW mode. Both operation modes have their advantages and limitations. Generally, the spectral coverage of QCLs is larger when operated in pulsed operation compared to CW mode. Pulsed operation suffers from pulse-to-pulse intensity fluctuations [23], but requires lower energy and no water cooling, so this mode offers advantages for portable devices. For measurements close to the strongly absorbing HOH-bending band of water (centering at 1643 cm^{-1}), large absorption values can lead to excess heat transfer to the sample, which is undesired due to possible alterations of the sample. Consequently, pulsed mode operation is employed for recording broadband IR spectra across the amide I and II regions, particularly in sweep mode, as low average power avoids unwanted heat transfer to the sample which might lead to sample alteration while maintaining high peak power levels that allow for large optical path lengths in transmission measurements. For measurements of highly time-resolved (tr) kinetic traces at individual wavelength, usually the CW mode is employed, because it provides an emission at a particularly constant emission power and high wavelength stability.

A further inherent quality of QCLs is the coherent nature of the emitted light. This characteristic enables advanced measurement schemes beyond conventional direct absorption measurements such as Mach–Zehnder interferometry [24] or optical heterodyne detection [25,26]. The latter technique was made possible by the development of frequency combs that were employed for dual-comb spectroscopy for single-shot broadband IR spectroscopy on protein reactions with sub μs time resolution [27]. Mach–Zehnder interferometry was recently introduced for liquid-phase samples [28], but this approach has not yet been translated for the analysis of proteins.

Finally, the radiation emitted by QCLs is linearly polarized, which principally enables the development of QCL-based setups for vibrational circular dichroism measurements. This method was employed for monitoring the pH-induced change in peptides [29], but has not yet been adopted for the investigation of proteins.

The following sections provide an overview about the different applications of QCL for mid-IR spectroscopic analysis of proteins. Even though evanescent wave–based measurements were attempted [30,31], the focus is on transmission measurements, because the advantages of QCLs, most notably the high emission power, can be efficiently harnessed in this measurement mode and consequently the majority of applications were performed this way.

3.2 Steady-state broadband infrared transmission spectroscopy of the protein amide bands

Broadband IR spectroscopy of the protein amide region is employed for qualitative and quantitative analysis of proteins. Regarding the analysis of the protein structure, the amide I band (1700−1600 cm^{-1}—originating from the C = O stretching and N−H in-phase bending vibration of the amide group) has been recognized to be most sensitive to the secondary structure. Furthermore, also the amide II band (1600−1500 cm^{-1}—arising from N−H bending and C−N stretching vibrations) was shown to be sensitive to protein structure. The sensitivity to individual secondary structure elements originates in differing patterns of hydrogen bonding, dipole−dipole interactions, and geometric orientations in the α-helices, β-sheets, turns, and random coil structures that induce different frequencies of the C = O vibrations [32]. The quantitative estimation of the individual secondary structure elements can be either performed by curve fitting [33], or by using more sophisticated chemometric approaches such as partial least squares regression [34] or multivariate curve resolution-alternating least squares (MCR-ALS) [35] of the amide I region. However, it has been shown that additional and more in-depth information about protein secondary structure can be gained by combined analysis with the amide II band, particularly when performed with chemometric analysis [36−38].

Acquisition of steady-state laser-based broadband IR transmission spectra of proteins was achieved with EC-QCLs operated in sweep mode as well as with the step-and-measure method, in combination with a microfluidic modulation approach.

3.2.1 External cavity-quantum cascade laser-based infrared transmission spectroscopy of proteins recorded in sweep mode

The first setup for recording EC-QCL-based broadband IR transmission spectra of proteins was introduced by the Lendl group in 2015 [39]. The system consisted of an EC-QCL covering the amide I region, which was operated in pulsed mode at a duty cycle of 5%. The IR light was recorded by a Peltier-cooled mercury cadmium telluride (MCT) detector and processed by a two-channel boxcar detector (see Fig. 3.1A). Due to the high optical power provided by the QCL light source, the optical path of the transmission cell employed in this setup could be extended to 38 μm, which is approximately four times larger than applicable with conventional FTIR spectrometers.

This experimentally obtained increase of the transmission path is well in accordance with the theoretical considerations outlined in the previous section. An important part of the setup was a custom-made temperature-controlled flow cell with relative temperature stability of 0.005°C to ensure minimal temperature difference between the reference and sample measurement. For the employed first-generation EC-QCL the fastest tuning time

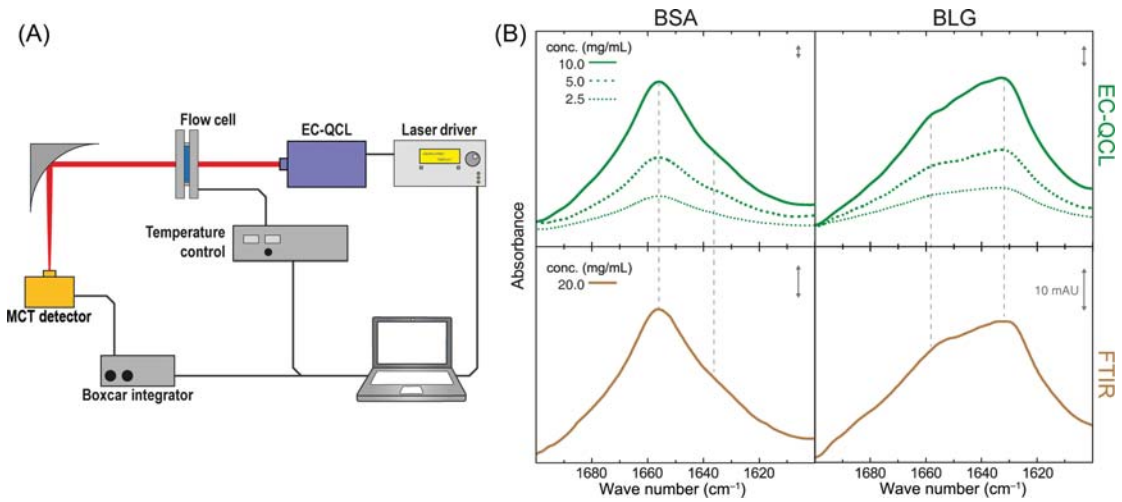


Figure 3.1

(A) Schematic of the experimental first-generation external cavity-quantum cascade laser (EC-QCL)-based setup for broadband mid-infrared (mid-IR) transmission measurements. (B) IR absorbance spectra of 10 mg/mL (*green solid line*), 5 mg/mL (*green dashed line*), and 2.5 mg/mL (*green dotted line*) protein solutions acquired by the EC-QCL setup (*top*), and 20 mg/mL (*brown solid line*) protein solutions acquired by Fourier-transform infrared (FTIR) spectroscopy (*bottom*). Gray double-headed arrows indicate an absorbance of 10 mAU. *Source: Figure A adapted from M.R. Alcaráz, A. Schwaighofer, H. Goicoechea, B. Lendl, EC-QCL mid-IR transmission spectroscopy for monitoring dynamic changes of protein secondary structure in aqueous solution on the example of beta-aggregation in alcohol-denaturated alpha-chymotrypsin, Anal. Bioanal. Chem. 408 (15) (2016) 3933–3941 [40]. Figure B adapted with permission from M.R. Alcaráz, A. Schwaighofer, C. Kristament, G. Ramer, M. Brandstetter, H. Goicoechea, et al., External cavity-quantum cascade laser spectroscopy for mid-IR transmission measurements of proteins in aqueous solution, Anal. Chem. 87 (13) (2015) 6980–6987 [39]. Copyright (2015) American Chemical Society.*

across the accessible spectral range of $\sim 165\ \mathrm{cm}^{-1}$ was 1.5 seconds. The tuning characteristic of the laser when operated in the sweep mode was determined by step-scan FTIR spectroscopy. In doing so, the time axis of a scan could be converted to a wave number axis as required for the protein spectra. To decrease the noise level, 100 scans were recorded amounting to a total measurement time of approximately 8 minutes (after every recorded forward scan, the grating needs to go back to its initial position). However, it turned out that here the recorded scans cannot directly be averaged as in FTIR spectroscopy, because the individual scans recorded from this early EC-QCL light source were slightly shifted in time, that is, wave number axis. This offset originates in the manner in which the spectral tuning across the wave number range is performed, which is by a quasicontinuous motion of the EC grating. Apparently mechanical imperfections as well as triggering issues lead to a spectral mismatch in the fine mode-hop structure of consecutive

scans. Since these fluctuations are not of a stochastic nature, averaging of a high number of scans is not a suitable approach to decrease its disturbance. Furthermore, it was found that the shift is not constant throughout the entire scan, that is, a uniform offset, but the extent changes along the wave number axis, which indicated irregularities in the motion of the EC grating. In order to rectify this experimental drawback, a spectral alignment routine was implemented based on the rubberband-type alignment algorithm "correlation optimized warping." This approach utilizes the inherent mode-hop structure in the EC-QCL emission spectrum to align a given scan with a selected reference scan. This alignment procedure was employed for scan to scan alignment before averaging as well as for background to sample alignment before calculating the absorbance. After this alignment procedure, only minor Fourier filtering was required to obtain high-quality spectra. In a final step of data processing, deviations in the wave number scale were calibrated to the (considered as absolute) scale of FTIR spectroscopy. For the EC-QCL-based setup, deviations in the wave number scale may occur from inaccuracies in the light source and data acquisition as well as data processing. In data acquisition, triggering inaccuracies and internal data processing routines may introduce time lags. As mentioned above, the temporal axis directly corresponds to the wave number axis and consequently any time lags directly translate into wave number deviations. For correction of these deviations in the wave number scale, the recorded EC-QCL spectra were calibrated to the absorption lines of water vapor, which proved to be suitable in the investigated spectral region. To this end, reference spectra can be obtained by an FTIR spectrometer, but in principle reference data for various gases available from the HITRAN database could also be employed for this calibration step.

The developed EC-QCL-based IR transmission setup was employed to record IR spectra of several proteins with different secondary structure (see Fig. 3.1B depicting bovine serum albumin (BSA) for α-helical structure and β-lactoglobulin (BLG) for β-sheet structure), and the obtained results were compared with FTIR spectra. The comparability of the IR spectra acquired by EC-QCL and FTIR spectroscopy was quantitatively evaluated by calculating the degree of spectral overlap, which obtained values >0.99 for all investigated proteins, indicating the high consistency of the recorded spectra by the two different techniques. Fig. 3.1B shows the recorded IR spectra of BSA and BLG at different concentrations and a comparison of the band shape with FTIR spectra. With this first-generation EC-QCL-based setup, IR spectra with well-defined spectral features for qualitative analysis of the secondary structure could be obtained for protein concentrations down to 2.5 mg/mL. Regarding quantitative analysis, the limit of detection (LOD) was determined to be 0.117 mg/mL.

The first-generation EC-QCL IR transmission setup was further employed to monitor slow dynamic changes of protein secondary structure. To this end, protein denaturation of α-chymotrypsin (α-CT) in the presence of the cosolvent 2,2,2-trifluoroethanol (TFE) was monitored [40]. After exposure of β-sheet rich proteins such as α-CT to TFE, there occurs

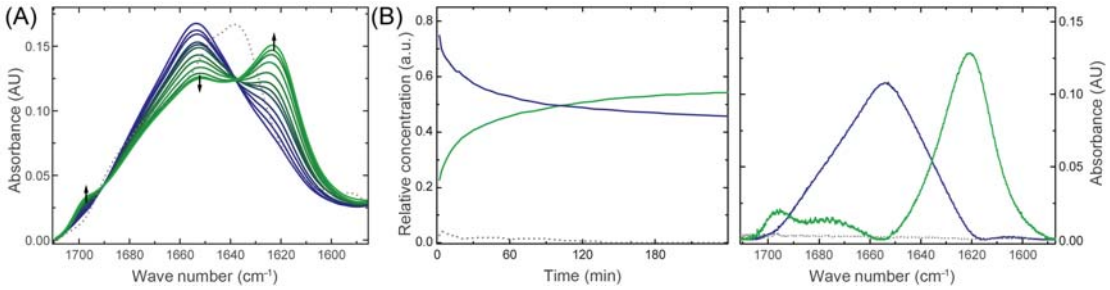


Figure 3.2

(A) Time-dependent infrared (IR) absorbance spectra of 20 mg/mL α-chymotrypsin (α-CT) in 50% 2,2,2-trifluoroethanol (TFE)/buffer solution, pH 7.8 at 25°C (*solid lines*). The spectra were recorded at time periods between 2 and 240 min after the protein was dissolved in TFE/buffer. Blue solid lines show the spectrum of α-CT with TFE-induced α-helical structure. Green solid lines indicate the spectrum of the protein after gradual formation of intermolecular β-sheets. Gray dashed lines represent spectra of the native protein in aqueous buffer. Black arrows illustrate directions of absorbance changes as a function of time. (B) Temporal (*left*) and spectral (*right*) profiles retrieved by multivariate curve resolution-alternating least square (MCR-ALS) for 20 mg/mL α-CT in 50% TFE/buffer solution at pH 8.2. Solid blue and solid green lines show individual contributions of α-helical and intermolecular β-sheet conformation, respectively. Dashed gray lines indicate the instrumental noise of the system. *Source: Adapted from M.R. Alcaráz, A. Schwaighofer, H. Goicoechea, B. Lendl, EC-QCL mid-IR transmission spectroscopy for monitoring dynamic changes of protein secondary structure in aqueous solution on the example of beta-aggregation in alcohol-denaturated alpha-chymotrypsin, Anal. Bioanal. Chem. 408 (15) (2016) 3933–3941 [40].*

rapid formation of nonnative α-helical structures in the time range of milliseconds, which is not directly observable with the employed setup. This fast conformational change is followed by slower β-sheet aggregation that was monitored by the EC-QCL-based setup. Fig. 3.2A shows the gradual conformational change from α-helical secondary structure (band maximum at 1654 cm^{-1}) to intermolecular antiparallel β-sheet secondary structure (band maximum at 1623 cm^{-1} and shoulder at 1697 cm^{-1}) of the protein throughout a time period of 240 minutes. The effect of different pH values and protein concentration on the kinetics of β-sheet aggregation has been investigated employing EC-QCL IR spectroscopy and the results were evaluated by extended MCR-ALS. In the extended variant of MCR-ALS, seven individual tr-IR measurements at different pH values were combined to an augmented data matrix to obtain the temporal evolution of β-aggregation as well as spectral profiles (Fig. 3.2B). The obtained spectral profiles of pure spectra are consistent with band maxima and shapes of the associated protein secondary structure components identified in the EC-QCL IR spectra. The temporal evolution shows a decline of the α-helical secondary structure, that goes along with the emergence of the β-sheet secondary structure component.

For analysis of protein structure and dynamics, as an alternative to measurements in aqueous solution, IR measurements of proteins are commonly performed in deuterated

solution. D_2O is chosen as a solvent as it has mostly similar physical and chemical properties when compared to regular water. Regarding IR spectroscopy, this solvent provides the advantage that the DOD-bending band is located at $\sim 1200\ cm^{-1}$, thus creating a region of relatively low solvent absorbance between 1500 and $1800\ cm^{-1}$. Upon solvent exchange, the amide II band (predominantly originating from N−H vibrations) is shifted from ~ 1550 to $\sim 1450\ cm^{-1}$, then referred to as the amide II′ band. Since the amide I band is mainly composed of C = O vibrations, its shift is relatively small ($5-10\ cm^{-1}$) compared to the amide II band. The shift to lower wave numbers of the intense absorption bands of the solvent inherently make it possible to employ larger path length for IR measurements in the amide region. Consequently, for recording IR spectra in a deuterated solution the optical path could be extended by the same factor compared to FTIR spectra acquisition (commonly used path length: $\sim 100\ \mu m$) as for measurements in aqueous solution. Thus the previously introduced first-generation EC-QCL-based setup [39] was equipped with a transmission cell with the path length of 478 μm [41]. The setup was employed for recording IR spectra of proteins featuring different secondary structures. Fig. 3.3A shows QCL-IR spectra of BSA and concanavalin A (ConA) in a deuterated medium at protein concentrations ranging from 0.25 to 2 mg/mL. The distinctive band shapes of the α-helical (BSA) and β-sheet (ConA) secondary structures compare well with the respective FTIR spectra. Beside measuring these model proteins, the setup was further employed for monitoring dynamic changes in the secondary structure of poly-L-lysine (PLL) induced by thermal denaturation. This polypeptide is a prominent model system, because its secondary structure can be readily altered by variation of the external conditions such as pH/pD value and temperature, among others. At pD 12.4, PLL adopts α-helical structure at room temperature, which is transformed to β-sheet secondary structure upon heating to 50°C. Fig. 3.3B shows dynamic QCL-IR spectra of the temperature-induced α−β transition of PLL acquired with the EC-QCL-based setup. The spectrum recorded at 20°C shows a maximum of the amide I′ band at $1635\ cm^{-1}$, characteristic for the α-helical structure adopted by PLL at these conditions. Further, upon increasing the temperature, the band shape attributed to the α-helix configuration decreases as two bands at 1611 and $1681\ cm^{-1}$ emerge, attributed to the high- and low-frequency bands of intermolecular, antiparallel β-sheets in PLL. The inset depicts the progress of the conformational transition with increasing temperature as evaluated by the absorbance of the low-frequency β-sheet band at $1611\ cm^{-1}$. In this study the influence of the polypeptide concentration on the point of inflection of the temperature-induced α−β transition was investigated at equal heating rates. By employing QCL-IR, FTIR, and electronic circular dichroism spectroscopy, it could be shown that the transition point decreases with increasing PLL concentrations [41].

Apart from these investigations of conformational changes of protein secondary structure after different types of external perturbation, the developed EC-QCL-based IR transmission setup was also employed for the quantification of individual proteins in commercial bovine

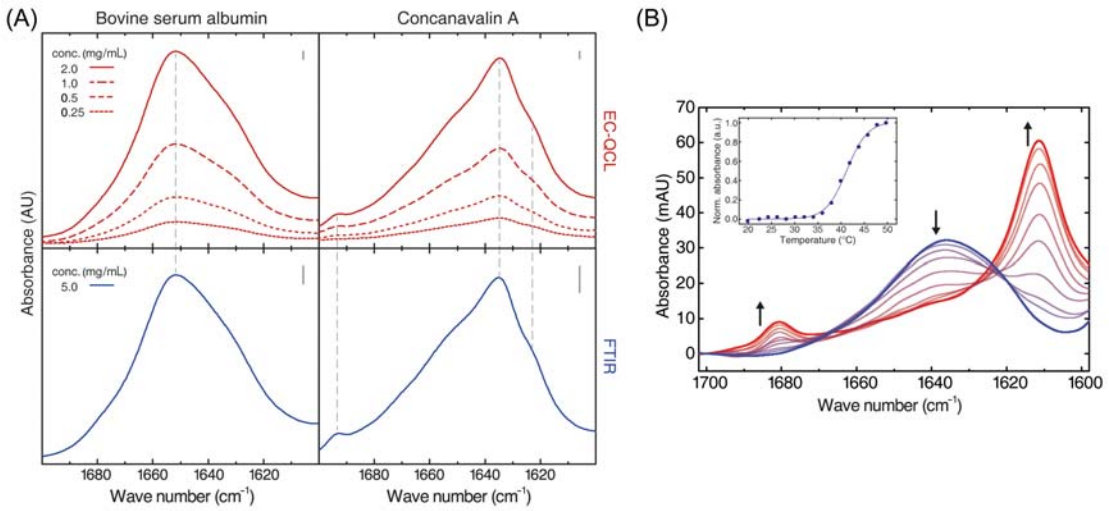


Figure 3.3

(A) Infrared (IR) absorbance spectra recorded with the external cavity-quantum cascade laser (EC-QCL)-based setup of 2 mg/mL (*red solid line*), 1 mg/mL (*red dashed line*), 0.50 mg/mL (*red short dashed line*), and 0.25 mg/mL (*red dotted line*) protein solutions (*top*), and 5 mg/mL (*blue solid line*) protein solution obtained by Fourier-transform infrared (FTIR) spectroscopy (*bottom*). Gray bars indicate the absorbance of 10 mAU. (B) Infrared (IR) spectra of 0.50 mg/mL poly-L-lysine (PLL) in deuterated phosphate buffer pD 12.4 recorded as function of the temperature by using the EC-QCL-based IR transmission setup. The arrows indicate directions of spectral changes with temperature. The inset shows the progression of the relative change in IR absorbance obtained by evaluation of the band height at 1611 cm^{-1}, at PLL concentration of 0.50 mg/mL. Dashed lines represent a fitted sigmoidal curve. *Source: Adapted from A. Schwaighofer, M.R. Alcaráz, C. Araman, H. Goicoechea, B. Lendl, External cavity-quantum cascade laser infrared spectroscopy for secondary structure analysis of proteins at low concentrations, Sci. Rep. 6 (2016) 33556 [41].*

milk [42]. Here the advantage of the increased applicable optical path could be directly harnessed for a real-life application and it enabled the analysis of the amide I region of milk samples in transmission mode, which is not possible with FTIR spectroscopy. Milk is a liquid with high viscosity and a complex matrix composition that features a diverse chemical makeup involving fat, protein, and lactose. When it is inserted into a transmission cell with low optical paths, it tends to clog or form air bubbles, thus preventing robust sample handling. For this reason, FTIR spectroscopy could be employed merely for investigation of the mid-IR fingerprint region at wave numbers $<1600\ \text{cm}^{-1}$ using a transmission path length of $\sim 50\ \text{cm}^{-1}$. Consequently, with FTIR spectroscopy, it was not possible to access the secondary structure sensitive amide I region. By employing the first-generation laser-based IR transmission setup, the amide I band of milk samples can be measured in a robust way, and with partial least squares (PLS) modeling the individual concentrations of casein, BLG, as well as the total protein content could be determined in

commercial bovine milk samples. Casein is an economically important component of milk and BLG is an important indicator for heat treatment. Prior to quantification, the subtraction of the spectral contribution of the milk matrix from the overall IR signal proved to be beneficial. Matrix subtraction was necessary as constituents of the matrix, even when not absorbing themselves in the investigated spectral region, led to variations in the absorbance intensity of the H_2O-solvent band at 1643 cm^{-1} due to the displacement of water caused by milk matrix compounds relative to the H_2O background spectrum [43]. This step was performed with the multivariate science-based calibration method. This strategy of matrix subtraction followed by performing separate PLS modeling for casein, BLG, and total protein content enabled quantification of major milk proteins within 10 minutes and prediction accuracies of approx. 1 mg/mL (root mean square error of cross validation; RMSECV). By determination of the concentration of the temperature-sensitive protein BLG, different commercial bovine milk types could be discriminated according to the experienced heat load during processing, such as ultrahigh temperature milk (high heat load) versus pasteurized milk (low heat load). Protein concentrations obtained by EC-QCL-based IR transmission spectroscopy for analysis of different commercial bovine milk types compared well with results received from the standard reference method for protein quantification in milk, high-performance liquid chromatography (HPLC) [44]. With this application example, laser-based IR spectroscopy combined with chemometric analysis demonstrates the possibilities and potential beyond the scope of currently employed standardized methods in terms of rapid analyte quantification in complex liquids, as for example, HPLC requires several time-consuming manual processing steps prior to chromatographic analysis.

In recent years there has been the commercial realization and availability of EC-QCLs with improved emission characteristics, such as wider spectral emission ranges, faster tuning rates, as well as enhanced repeatability of the steering and control of the EC. Regarding the qualitative and quantitative analysis of proteins, particularly in combination with chemometric methods, additional and more in-depth information can be gained by analysis of a broader spectral region covering the amide I and amide II region. To this end, the Lendl group introduced a novel EC-QCL-based IR transmission setup employing a latest-generation commercial EC-QCL with a spectral tuning range between 1730 and 1470 cm^{-1}. Here, the laser was operated at a repetition rate of 100 kHz with a duty cycle of 50% and the signal was processed by a lock-in amplifier. An experimental challenge with the expanded spectral region was to reconcile the high attenuation of the laser emission in the region of the HOH-bending band in the amide I region and the high (unattenuated) emission power of the laser across the amide II region with the dynamic range of the MCT detector. A solution for this issue was to implement an optical element that selectively attenuates the laser power in the amide II region in order to match the strong solvent absorption in the amide I region. For this purpose, a wedged sapphire window was employed that acts as a

shortwave pass filter with a cutoff at approximately 1650 cm^{-1}. In this context of accessible spectral range, potential detector saturation, and limited dynamic range of the detector, the optimum transmission path was found to be 31 μm. Compared to the first-generation laser, the employed latest-generation laser featured improved emission characteristics. On the one hand, the laser allowed an increased scanning speed of 1200 cm^{-1}/s which made it possible to record 100 scans for averaging within 53 seconds. Furthermore, the control over the motion of the EC was evidently improved. Consequently, a rubberband-type alignment routine for postprocessing of the recorded scans was not required. However, evaluation of repeated background and sample scans, respectively, according to their similarity allowed deviating scans to be revealed, and their remove prior to further data processing to lower the noise level. By implementation of these experimental and processing measures, IR spectra of proteins could be successfully recorded across the amide I and amide II regions. IR spectra were obtained of several proteins with different secondary structures in the concentration range between 1 and 5 mg/mL and the spectral features were well comparable with FTIR spectra [45]. With the short measurement time and the low noise levels, the performance of this laser-based IR transmission setup employing a latest-generation EC-QCL could be directly benchmarked against FTIR spectroscopy in terms of LOD acquirable at a defined time period (see Table 3.1). For evaluation of the root mean square (RMS) noise of the different setups, 100% lines are employed which are obtained by calculating the absorbance spectrum of two subsequent single beam spectra of the same sample (in this case the solvent water) under identical conditions. Under ideal conditions, the result would be a flat line at 100% transmittance, corresponding to zero absorbance. As summarized in Table 3.1, when comparing the two described EC-QCL-based setups, with the latest-generation laser, lower noise levels by a factor of $\sim$6 could be achieved employing approximately one-tenth of the measurement time. This EC-QCL setup was also compared to a high-end FTIR spectrometer

Table 3.1: Characteristic parameters and relevant results for comparing the performance of external cavity-quantum cascade laser (EC-QCL)-based infrared (IR) transmission setups and conventional FTIR spectroscopy.

	Meas. time (s)/scans	RMS-noise 10^{-5} (AU)	Path length (μm)	LOD (mg/mL)	Spectral range (cm^{-1})	Detector type/temp. (°C)/cooling
QCL setup first gen.	500/100	37	38	0.117	1700–1600	MCT/–60/TE
QCL setup latest gen.	53/100	6.2	31	0.025	1700–1500	MCT/–78/TE
High-end FTIR	53/81	2.7	8	0.041	4000–600	MCT/–196/LN_2
Routine FTIR	53/21	12.1	8	0.182	4000–600	DLaTGS/25/–

LOD, Limit of detection; *MCT*, mercury cadmium telluride; *RMS*, root mean square; *FTIR*, Fourier-transform infrared.

(Bruker Vertex 80v). For FTIR measurements, a path length of 8 μm was employed. Even though the noise level obtained with the high-end FTIR spectrometer was approximately 2.3 times lower than with the laser-based setup, due to the higher applicable path length, with the laser-based setup a LOD better by a factor of approximately two could be reached compared to the high-end FTIR spectrometer.

This latest-generation EC-QCL-based setup for measurement of the amide I and amide II region was employed for analysis of commercial bovine milk samples. PLS analysis of the expanded spectral region also allowed the quantification of α-lactalbumin (ALA) in addition to the previously accessible proteins casein, BLG, as well as the total protein content. In this approach, the complex matrix of milk and its effect on the IR spectrum were directly considered in the PLS model by including milk samples at different dilutions. This approach is beneficial, in cases where matrix compositions are not exactly known or not reproducible for the PLS calibration set. With this method, quantification could be performed within 1 minute data acquisition time without any sample processing with prediction accuracies of approximately 0.25 mg/mL (RMSECV) [46]. By simultaneous determination of the concentration of the two temperature-sensitive proteins BLG and ALA, it was possible to perform a fine discrimination between different commercial bovine milk types as shown in Fig. 3.4. First, it is possible to distinguish milk experiencing high heat

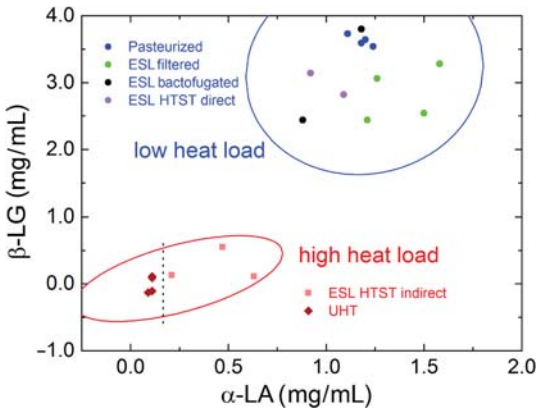


Figure 3.4

β-Lactoglobulin (β-LG) concentration plotted versus α-lactalbumin (α-LA) concentration. Milk types experiencing low heat load appear at high β-LG and α-LA concentrations, indicating soft heat processing. Further subdivision of the milk samples analyzed within this study into high temperature–short time (HTST) indirect and ultrahigh temperature (UHT) samples is indicated by the black dashed line. Source*: M. Montemurro, A. Schwaighofer, A. Schmidt, M.J. Culzoni, H.K. Mayer, B. Lendl, High-throughput quantitation of bovine milk proteins and discrimination of commercial milk types by external cavity-quantum cascade laser spectroscopy and chemometrics, Analyst 144 (2019) 5571-5579.*

load from low heat load. Furthermore, within the high temperature–short time (HTST) milk type, it is possible to distinguish different heating technologies (i.e., direct and indirect heating). And finally, within the milk processed with high heat load, further discrimination can be obtained to identify ultrahigh temperature (UHT) milk.

3.2.2 QCL-based infrared transmission spectroscopy of proteins recorded in step-and-measure mode with microfluidic modulation

An alternative approach for recording broadband IR transmission spectra of proteins in the amide I region was recently presented in the form of microfluidic modulation spectroscopy (MMS) [47]. Here, the experimental setup consists of a tunable QCL light source and a thermoelectrically cooled MCT detector. As depicted in Fig. 3.5A, the key component of the setup is a microfluidic controller that is capable of rapidly exchanging the reference liquid with the sample solution. Throughout the measurement, the buffer and reference

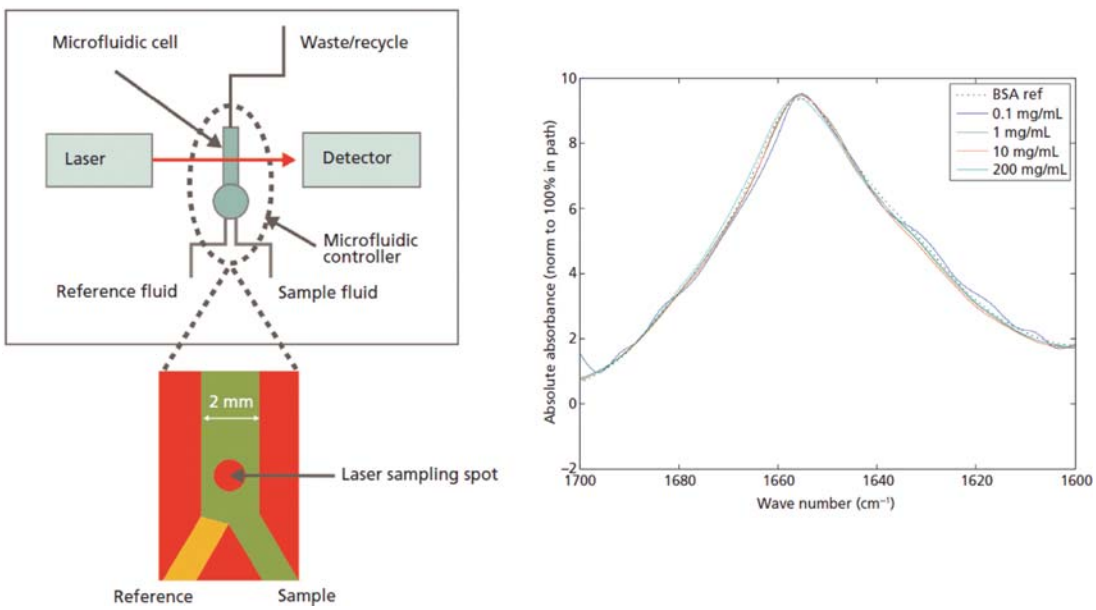


Figure 3.5

(A) Schematic of the experimental setup of an quantum cascade laser (QCL)-based infrared (IR) transmission system involving microfluidic modulation spectroscopy. (B) Microfluidic modulation spectroscopy (MMS)-based quantum cascade laser-infrared (QCL-IR) spectra of bovine serum albumin (BSA) recorded at a concentration range between 0.1 and 200 mg/mL and a Fourier-transform infrared (FTIR) spectrum for comparison of the band shape (BSA ref). *Source: Adapted from E. Ma, L.B. Wang, B. Kendrick, Enhanced protein structural characterization using microfluidic modulation, Spectroscopy 33 (7) (2018) 46–52 [47]. Reprinted with permission of the authors.*

solution are introduced into the transmission cell (path length: 25 μm) under continuous flow with a modulation frequency of 1−10 Hz. For acquisition of one IR spectrum the QCL, which is operated in CW mode, is stepped across the amide I region (1590−1714) to record 31 discrete wave numbers with a data point spacing of 4 cm^{-1}. This way of operation, that is, the immediate-successive and alternate measurement of the background and sample solution, provides several advantages. First, for acquisition of the background and sample intensity, the laser merely needs to be tuned once to the respective wave number. Traditionally, for acquisition of an absorbance spectrum, in the first step the intensity of the background solution is measured across the entire wave number region and subsequently the same is performed for the intensity of the sample solution. However, in this step-and-measure emission mode, it may occur that different modes are emitting upon repeated tuning to the same wave number, which adds to the noise of the final absorbance spectrum. Consequently, this source of noise can be circumvented by exchanging the liquids for background and sample measurement while the laser is tuned once to a respective wave number. Secondly, the applied modulation further reduces the sensitivity to ambient drifts, particularly when no temperature control of the fluid or measurement cell is applied [47]. Most recently, this measurement platform equipped with an auto sampler became commercially available (AQS^3pro, RedShiftBio).

It was shown that this approach provides good repeatability, low LODs, and a wide dynamic range. Fig. 3.5B shows overlays of IR spectra of BSA covering the concentration range from 0.1 to 200 mg/mL and comparison of the band shape with FTIR spectra. Furthermore, the linearity of recorded data in this wide concentration range was demonstrated [47]. In a first reported application, MMS-based QCL-IR spectroscopy was employed to monitor the change of secondary structure of an antibody−drug conjugate (ADC) after thermal and pH-induced stress [48]. After storing the ADC for 20 minutes at 70°C, the predominant parallel β-sheet secondary structure changed to antiparallel β-sheets, as determined by Gaussian-based curve-fitting. This finding was supported by size exclusion chromatography, which found that a large part of the sample was aggregated. A similar secondary structure change was not found when storing the ADC for 10 days at 50°C. In the pH stress experiment, a conformational change of parallel to antiparallel β-sheet secondary structure was found after exposing the ADC to pH 3 for 3 days.

3.3 Time-resolved laser-based infrared spectroscopy to monitor protein dynamics

tr-IR spectroscopy is a powerful tool to monitor fast, dynamic changes of protein structure in real-time. The capability of IR spectroscopy to investigate biological processes across the entire time range between femtoseconds and hours is a pronounced advantage compared to many other techniques. Ultrafast IR spectroscopy at time resolutions in the femto- and

picosecond timescale is performed by dedicated pump-probe setups, for example, to study energy and electron transfer in photosynthetic systems [49]. Employing conventional FTIR instruments, the most commonly applied methods for tr-IR spectroscopy are the rapid-scan and step-scan techniques.

When recording FTIR spectra in the rapid-scan mode, the process of data acquisition is comparable to conventional FTIR spectroscopy, however the modulation speed of the mirror is increased to achieve time resolutions of 10 ms at a spectral resolution of 4 cm^{-1}, provided the detector has a fast response and the analog-to-digital converter is fast enough to correctly digitize the interferogram [50,51]. Here, the limitations for time resolution are posed by the optical components in the FTIR spectrometer such as the achievable velocity and distance covered by the mirror. Consequently, the time resolution can be slightly improved by lowering the spectral resolution.

For tr-FTIR spectroscopy at a time resolution from tens of nanoseconds to hundreds of milliseconds, the step-scan technique is performed. In this approach, time traces at individual interferogram data points (corresponding to specific mirror positions) are recorded separately. To achieve this, the mirror is moved to a particular position prior to starting the investigated reaction by an external trigger (light flash, potential step, etc.) and the time trace is recorded while the mirror is at rest. Data points from multiple repetitions can be coadded at a constant retardation to increase the SNR. Subsequently, the mirror is moved to the next position, and the reaction is initiated once again. This process is repeated until time traces are recorded for each sampling point of the interferogram. Finally, the obtained data matrix containing the temporal changes for all mirror retardations is rearranged to yield interferograms at specific times after the reaction start. A clear requirement for this acquisition mode is the repeatability and reproducibility of the observed sample reaction, as it must be repeated a large number of times, that is, mirror stop positions multiplied by the number of repetitions, thus 50,000−100,000 repetitions are the norm for most applications [52,53]. A further prerequisite for successful data acquisition is the long-term stability of the (biological) system. Particularly if longer recovery kinetics are required for the protein system under investigation, the data acquisition time can add up to hours and even days [54]. Given these conditions, at each interferogram point spectral data acquisition is decoupled from the time domain, so theoretically any degree of spectral resolution as well as time resolution with any SNR is possible. However, practical limitations in the step-scan mode arise from the electrical components of the instrument such as the speed and sensitivity of detectors and amplifiers.

The thereby obtained tr-IR spectra contain a multitude of absorption bands from the protein under investigation as well as other biomolecules present in the sample. Hence, a filter is needed to focus the attention on the molecular groups involved in the induced reaction. This kind of selectivity is commonly obtained by difference spectroscopy. To this

end, the IR spectrum of the sample in the equilibrium state is subtracted from the sample in the intermediate state. Consequently, the constant absorbance background of the sample is abolished and the dynamic spectral changes are highlighted. To be able to perform this kind of evaluation, the external conditions such as sample volume, concentration temperature, pH, etc. need to be strictly maintained throughout the reaction. This requirement is more easily fulfilled for reactions induced by light or potential steps than for reactions that involve mixing of two reaction substrates or a change in the buffer conditions [50,55].

Generally, tr-FTIR spectroscopy compared to other IR techniques benefits from the multiplex advantage. However, it is limited in the SNR by the low emission power of the broadband thermal light source [56]. For this reason, in recent years, QCLs were increasingly used also for investigation of fast protein dynamics. In the corresponding measurements, protein kinetics are monitored at single frequencies with superior SNR. For acquisition of multiple wavelengths, mostly broadly tunable EC-QCLs are employed and the reaction is repeated for acquisition of single wave numbers of interest. Consequently, the number of repetitions and thus data acquisition time is greatly reduced. Firstly, due to the higher emission power, the reaction needs to be repeated less often to yield a satisfying SNR. Secondly, kinetic traces at individual wave numbers need to be recorded separately. This seems to be a disadvantage compared to multiplexing in FTIR spectroscopy. However, with this approach, the recorded wave numbers or wave number range can be individually selected. Consequently, time traces can be only recorded in spectral regions that are of interest for the specific problem. Thus to monitor a desired spectral region of 100 cm^{-1} with a wave number spacing of 0.5 cm^{-1}, 201 time traces need to be recorded. In contrast, for acquisition of a step-scan FTIR spectrum covering 2000 cm^{-1} (even though possibly only 100 cm^{-1} are of interest) with a spectral resolution of 4 cm^{-1}, time traces at approximately 1000 mirror positions need to be recorded. As in step-scan FTIR spectroscopy, for QCL-based measurements the time resolution is only limited by detector electronics and can reach the range from 10 ns to 1 seconds, thus covering the most relevant timescales relevant in protein function.

The Heberle group developed an QCL-based flash-photolysis setup for studying light-induced reactions in proteins such as channelrhodopsin (ChR), bacteriorhodopsin (bR), and myoglobin. A schematic of the setup for investigation of the latter two model proteins is shown in Fig. 3.6A [57]. It consists of two EC-QCLs with spectral tuning ranges of 1660−1840 and 1840−1980 cm^{-1}, that are operated in CW mode. Before focusing the emitted laser beam on the sample, the laser light is directed through a servocontrolled adjustable aperture to keep the power hitting the detector constant and to prevent saturation. For initiation of the photoreaction, the second harmonic of a quality switched neodymium-doped yttrium aluminum garnet laser was used which emits $\sim$6 ns at 532 nm. Time-dependent signal changes were recorded by an MCT detector. In order to enable the

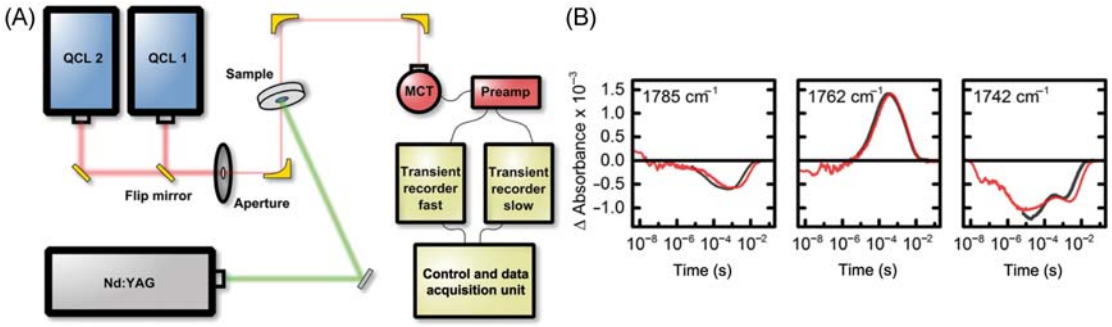


Figure 3.6

(A) Schematic of the developed external cavity-quantum cascade laser (EC-QCL)-based setup for studying protein dynamics. (B) Time traces for the photocycle of a bR film sample probed by the quantum cascade laser-infrared (QCL-IR) setup (*red*) at 1785 cm^{-1} (release of an excess proton from the internal hydrogen-bonded network of bR), 1762 cm^{-1} (protonation of primary acceptor D85 of the retinal Schiff base in the M intermediate of the photocycle), and 1742 cm^{-1} (C = O vibration of carboxylic acid side chain). Black: Time traces extracted from step-scan Fourier-transform infrared (FTIR) spectroscopy and multiplied by 3.5. *Source: Adapted from B.J. Schultz, H. Mohrmann, V.A. Lorenz-Fonfria, J. Heberle, Protein dynamics observed by tunable mid-IR quantum cascade lasers across the time range from 10 ns to 1s, Spectrochim. Acta Part A 188 (2018) 666–674 [57].*

calculation of difference absorbance values for the time traces of each individual wave number, pretrigger points were collected prior to initiating the reaction. Data sampling was performed by oscilloscopes employing a fast (250 MS/s) and a slow (1 MS/s) channel. Prior to further data processing, the dataset from the different channels were merged to obtain time traces covering the range between tens of nanoseconds and hundreds of milliseconds. In order to obtain spectra from the individual time traces, the data was convoluted in Fourier space by a Gaussian function. Further data postprocessing steps included averaging, smoothing, and singular value decomposition.

This setup was tested to study bR [57], a light-driven proton pump which was well-researched in recent decades in order to resolve its functional mechanism. To this end, a rehydrated film of bacteriorhodopsin in the native purple membrane was prepared on a CaF_2 window. Here, kinetics were recorded between 1700 and 1800 cm^{-1} in steps of 0.5 cm^{-1}. With 100 repetitions performed for every individual wavelength (amounting to 20,100 repetitions in total), and a repetition frequency of 4 Hz, the total data acquisition time was $\sim$4 hours. The obtained time traces were compared with step-scan FTIR spectroscopy (see Fig. 3.6B), which provided time-resolved data from 6.25 μs to 50 ms (here, 43,000 repetitions were performed to obtain spectra covering the spectral range of 2250–850 cm^{-1} with sufficient SNR). In contrast, the EC-QCL setup covered a time span from 4 ns to 250 ms. Three relevant time traces recorded at selected frequencies show

excellent agreement between the two data sets acquired with the different setups. From this it can be concluded that the absorption of the intense IR emission from the QCL does not significantly increase the temperature of the sample, which would influence the observed kinetics. In a further experiment, the developed setup was employed to study the photodissociation of bound carbon monoxide [57]. Photodissociation of carbonmonoxymyoglobin by a short laser pulse yields a short-lived photoproduct, which relaxes to deoxymyoglobin. This experiment was performed in a microfluidic cell with an optical path of 50 μm, and the sample was exchanged after every experiment. Individual time traces were recorded every 2 cm^{-1} in the spectral region from 1880 to 1960 cm^{-1}. Due to the high emission power of the QCL, only 10 averages were sufficient to obtain kinetic traces at high SNR, resulting in a total acquisition time of the dataset of only 30 minutes. Kinetic analysis of the C = O stretching vibration band of photodissociated carbon monoxide at 1943 cm^{-1} revealed a (pseudo) first order kinetic reaction with two time constants that agree with step-scan FTIR experiments [58].

The developed flash-photolysis setup was further employed to study kinetic and vibrational isotope effects in the proton transfer reactions of channelrhodopsin-2 (ChR2) upon exchanging H_2O with D_2O [59]. For these experiments, the setup was equipped with two EC-QCLs covering the spectral regions of 1600−1700 and 1690−1790 cm^{-1}. In the latter frequency range, transient protonation and hydrogen bonding changes of amino acids were observed for this protein. The results of the QCL-based setup were benchmarked against step-scan FTIR spectroscopy, highlighting the larger observable time range, 20 ns−300 ms with the QCL-based setup and 6 μs−300 ms with the step-scan FTIR technique. Subsequently, deuterium effects in carboxylic region were investigated. Fig. 3.7A shows the tr-IR differences of ChR2 in H_2O and D_2O, respectively. As expected, the vibrational frequencies are (predominantly) downshifted after the H/D exchange due to increase of the reduced mass induced by the heavier isotope. The higher mass further leads to a slowdown of the kinetics of the vibrational bands associated with the proton transfer. This delay is illustrated in Fig. 3.7B, which shows the delayed kinetics of proton transfer of the aspartic acid D156. The decay and subsequent recovery of the bands at 1736/1738 cm^{-1} in the millisecond time range were attributed to deprotonation and reprotonation of the amino acid. These differences in kinetics further indicate that D156 is accessible to H/D exchange.

The superior time resolution of this QCL-based setup was also employed to analyze a ChR2 variant with a E123T mutation. The resulting ChR2 mutant exhibits a five-time faster but still native-like photocycle [60]. For monitoring of fast changes with a time resolution of 15 ns in the amide I and carboxylic region, the setup was equipped with EC-QCLs emitting from 1610 to 1680 and from 1700 to 1780 cm^{-1}. In this study, it could be confirmed that conformational changes in a ChR take place in the nanosecond time range.

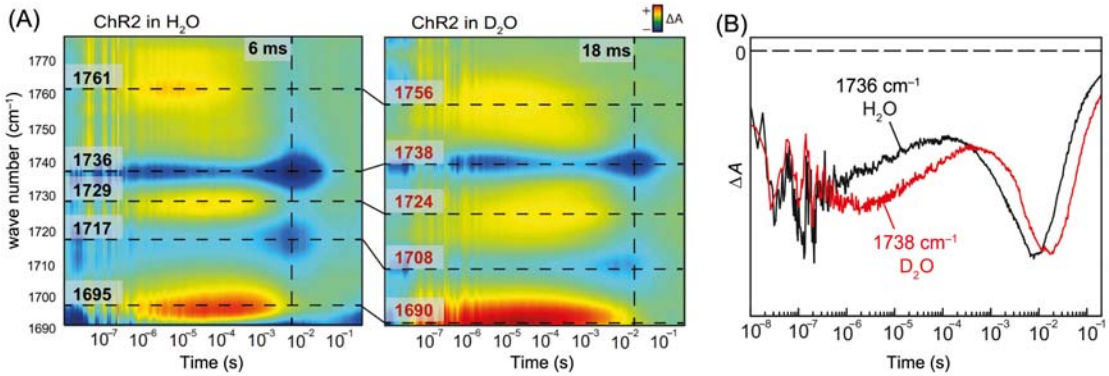


Figure 3.7

(A) Contour plot of the infrared (IR) absorption differences of ChR2 in H_2O (*left*) and D_2O (*right*) in the carboxylic region from 1690 to 1775 cm^{-1}. Positive and negative absorbance changes are indicated in red and blue, respectively. (*Dashed horizontal lines*) Maxima and minima of difference bands at particular wave numbers. (B) Transient absorption changes of the C = O stretching vibrational band of D156 (proton donor to the retinal Schiff base) probed at 1736 cm^{-1} in H_2O (*black*) and 1738 cm^{-1} in D_2O (*red*). *Source: Adapted from T. Resler, B.-J. Schultz, V.A. Lórenz-Fonfría, R. Schlesinger, J. Heberle, Kinetic and vibrational isotope effects of proton transfer reactions in channelrhodopsin-2, Biophys. J. 109 (2) (2015) 287–297 [59].*

The Hauser group developed a QCL-based spectrometer for analysis of folding dynamics of polypeptides and structural motifs after external perturbation, predominantly by pH- or T-jumps (Fig. 3.8A) [62,63]. To probe the relaxation kinetics in the amide I region, an EC-QCL with a spectral tuning range between 1580 and 1715 cm^{-1} was operated in CW mode. To generate the laser-excited T-jump, a Q-switched Ho:YAG laser with an emission wavelength of 2090 nm was used to excite an overtone vibration of D_2O yielding the rapid temperature jump in the sample volume. Different neutral density attenuators were employed to adjust the magnitude of the T-jump. For achieving homogenous heating of the sample volume, the pump pulse was split by a 50:50 beam splitter into two counterpropagating beams. The extent of the temperature jump was controlled by adjusting the intensity and the illuminated area of the laser beam. The final temperature was determined by referencing the temperature change to temperature-dependent FTIR spectra of the solvent D_2O. Photochemically induced pH jumps were achieved with a Q-switched Nd:YAG laser equipped with a fourth harmonic generator to obtain an emission wavelength of 266 nm. Photolysis of *o*-nitrobenzaldehyde (*o*-NBA) results in the release of caged protons and thus reduction of the pH value. The magnitude of the pH jump was estimated indirectly by calculating the number of protons released into the solution. These measurements were performed in D_2O, employing a flow-through cell with a path length of 100 μm. To separate peptide from solvent dynamics, the solvent solution was measured separately and then subtracted. For each type of relaxation experiment, as well as for the

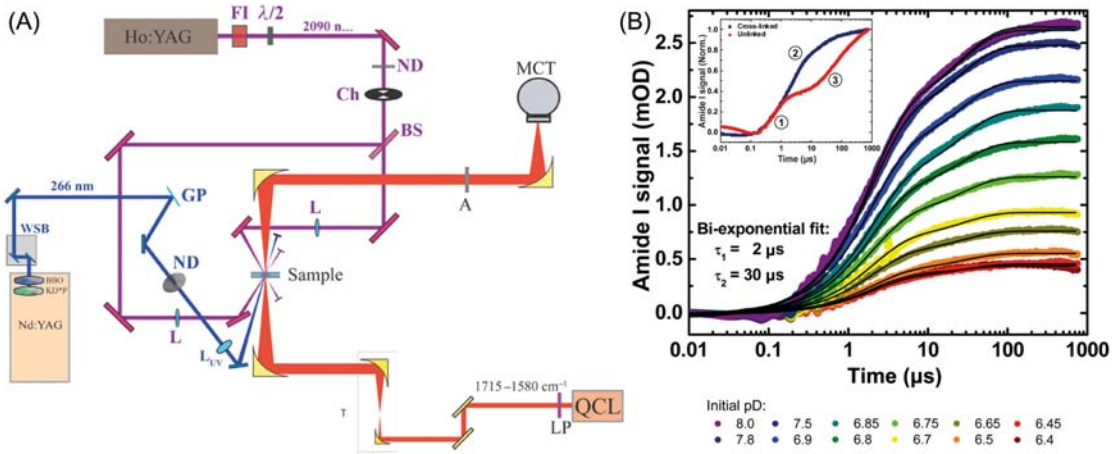


Figure 3.8

(A) Schematic of the quantum cascade laser (QCL)-based spectrometer combined with laser-excited perturbation techniques employed for pH- and T-jump studies. A QCL is used to monitor the relaxation kinetics (*red beamline*) initiated by a T-jump (excitation at 2090 nm, *purple beamline*) or by a photochemically induced pH jump (excitation at 266 nm, *blue beamline*). (B) Amide I′ transients recorded at 1629 cm^{-1} for the cross-linked peptide at various starting pDs. A global biexponential fit (*black lines*) revealed time constants $\tau_1 = 2$ μs and $\tau_1 = 30$ μs. The inset shows a comparison of the cross-linked and unlinked peptides. *Source: Figure A adapted from T. Resler, B.-J. Schultz, V.A. Lórenz-Fonfría, R. Schlesinger, J. Heberle, Kinetic and vibrational isotope effects of proton transfer reactions in channelrhodopsin-2, Biophys. J. 109 (2) (2015) 287–297. Figure B adapted with permission from M.L. Donten, S. Hassan, A. Popp, J. Halter, K. Hauser, P. Hamm, pH-Jump induced leucine Zipper folding beyond the diffusion limit, J. Phys. Chem. B 119 (4) (2015) 1425–1432 [61]. Copyright (2014) American Chemical Society.*

solvent background signal, several thousand transients were recorded at a repetition rate of 10 Hz and averaged in a quasilogarithmic manner.

The described setup was employed for investigating folding mechanisms and kinetics of polypeptides and structural motifs after pH- [61,62] and temperature- [61,63–66] induced perturbation. In particular, QCL-IR spectroscopy was employed for studying helix folding and unfolding of poly-L-glutamic acid (PGA) induced by a pH- and T-jump, respectively [62]. PGA is a model peptide, that is frequently employed to study helix formation. The helix content of the polypeptide can be modulated by variation of the pH value and temperature. To monitor the change of α-helical content after perturbation, transients of the relaxation dynamics were recorded at the single wave number of 1630 cm^{-1}. By inducing a T-jump of 6.6°C from 19 to 25.6°C at pD 5.8, the helical content was decreased by 4% (unfolding). The pH jump from pD 6.0 to pD 5.8 at a final temperature of 24.9°C increased the helicity by 20% (folding). Evaluation of the recorded relaxation transients by a biexponential fit revealed two components with time constants of a few hundred

nanoseconds and a few microseconds, respectively. This additional, fast kinetic step could first be resolved due to the good SNR obtained with the QCL light source. The observed time constants for the two experiments are similar and indicate that helix folding and unfolding occur with the same molecular mechanism.

The experimental setup was further employed to study the folding mechanism of a leucine zipper structural motif after a pH jump [61]. Here, the goal was to assess the stability of the tertiary structural motif after pH change and compare the kinetics of the cross-linked and unlinked peptide. As described above, the pH jump was induced by photolysis of *o*-NBA. For these measurements, peptide concentrations between 0.5 and 2 mg/mL were employed and solutions were exchanged after each measurement. Fig. 3.8B shows kinetic traces of the cross-linked peptide after perturbation with different initial pD values. Here, the band at 1629 cm^{-1} is related to the neutralization of the polyglutamatic acid side chain. For all measurements, the folding kinetics follow a biexponential function, indicating that the initial pD does not have an impact on the temporal evolution of the signal, except for the amplitude. The measurements were also performed with the unlinked peptide (see inset of Fig. 3.8B). In this case, a rate-limiting, diffusion-controlled step of the association of the two peptides is recognizable after approx. 5 μs.

3.4 Time-resolved infrared spectroscopy of protein dynamics by dual-comb spectroscopy

Dual-comb spectroscopy is an emerging spectroscopy method for broadband spectroscopy, which is based on frequency combs. This approach purely depends on time domain scanning and does not require any moving parts, in contrast to FTIR spectroscopy and also EC-QCL-based broadband spectroscopy. Based on this characteristic, particularly high time resolutions can be achieved at fast acquisition times, even with single-shot experiments without the need for experiment repetitions. This technique exploits the broad bandwidths as well as the high resolution and frequency accuracy of frequency comb light sources for recording sensitive broadband spectra [67]. Consequently, its application strongly depends on the availability of suitable light sources, that is, frequency combs and was first implemented at longer wavelengths. Initially employed comb sources in the mid-IR region were based on downconversion of near-infrared lasers using nonlinear crystals, which were limited in versatility due to the dependence on a multitude of optical components. The emergence of QCL-based frequency combs directly emitting mid-IR light in the recent years has enabled the development of compact dual-comb spectrometers in this spectral region [26,68]. Frequency combs provide simultaneous emission of equally spaced emission lines with a well-defined phase relationship. The line spacing can be tuned by the length of the QCL and by its group refractive index. Commonly employed line spacings for liquid-phase spectroscopy are in the range of a few GHz ($\sim$0.1 cm^{-1}) and several tens of GHz

($\sim 1\ \mathrm{cm}^{-1}$). QCL frequency combs are usually operated in CW mode and provide emission powers of several milliwatts per emission line [26]. For dual-comb spectroscopy (principle shown in Fig. 3.9A), two QCL frequency combs are employed, which provide an overlapping spectral emission range and slightly different line spacings. The line spacing of the first frequency comb (f_r) differs from that of the second comb ($f_r + \Delta f_r$) by Δf_r, which is in the region of several MHz ($\sim 0.0001\ \mathrm{cm}^{-1}$). By interference of the two frequency combs on the detection element, a heterodyne beat spectrum is generated in the radio frequency (rf) region. This rf spectrum is easily accessible with rf electronics and contains the relevant absorption information of the optical comb spectra. The time resolution of this measurement principle is limited by the requirement to resolve neighboring comb lines in the beating spectrum. Following from that, the shortest measurement time for full spectral evaluation of the rf beating signal is $t_{\min} = 2/\Delta f_r$. Furthermore, the comb line spacing f_r is related to the difference in line spacing Δf_r by $N \times \Delta f_r < f_r/2$. Consequently, for f_r in the GHz range and Δf_r in the MHz range, a time

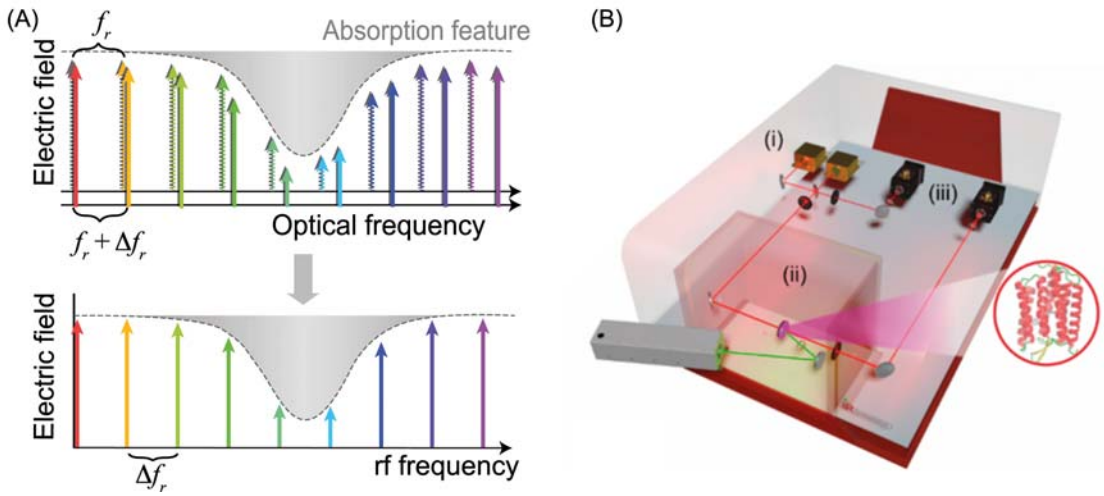


Figure 3.9

(A) Multiheterodyne beatnotes form a radio frequency (rf) comb with a line spacing of Δf_r, created by two frequency combs in the optical frequency region with slightly different comb spacings (f_r,$f_r + \Delta f_r$). Resulting from the comb structure, each pair of comb lines yields a rf heterodyne beatnote at a unique rf frequency. (B) Spectrometer with two quantum cascade laser frequency combs, (i) sample chamber with (ii) optically excited bacteriorhodopsin sample and (iii) two fast mercury cadmium telluride (MCT) photodetectors. *Source: Figure A was adapted with permission from I. Coddington, N. Newbury, W. Swann, Dual-comb spectroscopy, Optica 3 (4) (2016) 414–426, The Optical Society. Figure B was adapted from J.L. Klocke, M. Mangold, P. Allmendinger, A. Hugi, M. Geiser, P. Jouy, et al., Single-shot sub-microsecond mid-infrared spectroscopy on protein reactions with quantum cascade laser frequency combs, Anal. Chem. 90 (17) (2018) 10494–10500 [27]. Copyright (2018) American Chemical Society.*

resolution of ~ 1 μs can be achieved by detecting several hundred multiheterodyne beatnotes [27,67]. Based on this concept, a dual-comb spectrometer was commercialized (IRis-F1, IRsweep, Zurich, Switzerland, see Fig. 3.9B), and was employed to study photoinduced kinetics of bacteriorhodopsin with μs time resolution with single-shot spectroscopy [27]. For this application, two QCL combs with center wavelengths at approximately 1220 cm^{-1} with an emission range of 70 cm^{-1} are employed. The two combs overlap in the spectral region from 1185 to 1240 cm^{-1}. For spectra acquisition, the sample is introduced into the optical beam path of one of the frequency combs. The beams from the two frequency combs are combined by a 50:50 beam splitter, and are detected by two fast, high-bandwidth MCT detectors for noise reduction. Information about the sample absorption in the IR spectral region of the comb light is encoded to the rf region and recovered by heterodyne detection and data analysis. This setup delivers full spectra every 320 ns for a data acquisition time of 20 ms. The inset in Fig. 3.10C shows Allan plots demonstrating the achievable noise levels of single comb lines. At native spectral resolution, a RMS-noise of 1 mAU at 100 μs time resolution in a single shot and 50 μAU in 500 shots is observed. The noise can be further improved by averaging neighboring multiheterodyne lines and sacrificing the high spectral resolution, obtaining 400 μAU in a single shot and 15 μAU in 500 shots at 4.5 cm^{-1} resolution. These noise levels are in the range of routinely obtained values by FTIR spectroscopy, however, at much longer acquisition times.

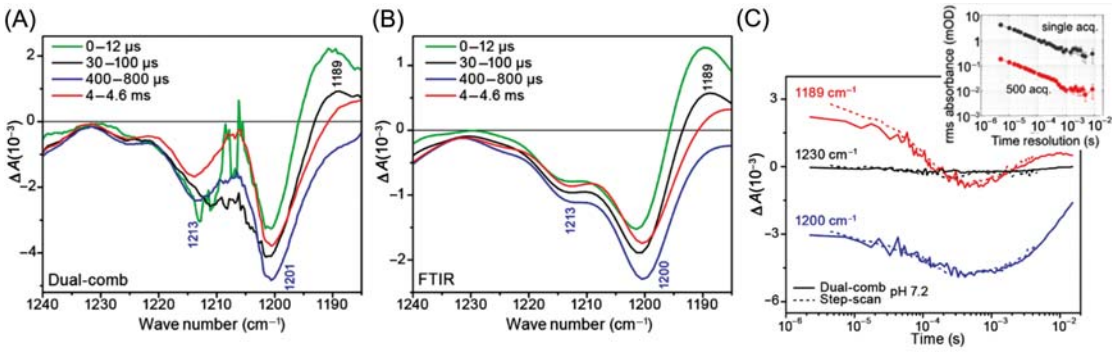


Figure 3.10

Infrared (IR) difference spectra representative for different intermediates (time constants) in the photocycle of bacteriorhodopsin obtained by dual-comb and step-scan Fourier-transform infrared (FTIR) spectroscopy. (C) Comparison of time traces obtained by dual-comb spectroscopy (500 averages) at three selected wave numbers with those from step-scan FTIR experiments. Inset: Allan deviation of a single comb line during a single acquisition and 500 acquisitions of 20 ms duration average. *Source: Figure was adapted from J.L. Klocke, M. Mangold, P. Allmendinger, A. Hugi, M. Geiser, P. Jouy, et al., Single-shot sub-microsecond mid-infrared spectroscopy on protein reactions with quantum cascade laser frequency combs, Anal. Chem. 90 (17) (2018) 10494−10500 [27]. Copyright (2018) American Chemical Society.*

For protein measurements, fragments of purple membrane containing bacteriorhodopsin were dried and rehydrated in a BaF_2 cell with a path length of 16 μm. The sample was excited at 532 nm with the second harmonic of a pulsed Nd:YAG laser. In the experiment, 5 ms of reference transmission were recorded prior to the 10-ns laser pulse and then the spectral evolution was monitored for 15 ms. The measurements were repeated 500 times at 2 Hz. IR spectra evaluated for different time constants associated with different intermediates in the photocycle of bacteriorhodopsin are depicted in Fig. 3.10A. The band shapes and positions show good accordance with step-scan FTIR spectra. Whereas the dual-comb spectra were recorded with 500 sample excitations, acquisition of the step-scan FTIR spectra required ~20,000 excitations of the protein. General discrimination of the characteristic profiles of the individual intermediates could be also accomplished by a single-shot experiment. Notably, dual-comb spectroscopy is the first experimental approach to offer single-shot analysis of protein kinetics in the mid-IR range with microsecond time resolution. Furthermore, three individual time traces were extracted from the dataset and compared with results obtained by step-scan FTIR experiments, showing good agreement of the two techniques for monitoring reaction kinetics.

3.5 Conclusions and future developments

Technological advancements and increasing commercial availability in recent years have led to the development of many setups and instruments for application in protein research. Maturation of EC-QCLs enabled sensitive analysis of the protein amide I and amide II bands. Combination with chemometric analysis allowed the elucidation of slow thermally and chemically induced secondary structure changes and quantification of multiple individual proteins in a complex matrix. Sample modulation techniques were developed to improve the SNR levels of the measurement. Further, EC-QCLs were employed for time-resolved monitoring of fast protein and polypeptide kinetics with time resolutions in the nanosecond range. It was demonstrated that laser-based setups could be used to record individual time traces of marker bands that were identified previously by steady-state or step-scan FTIR spectroscopy. Finally, the emergence of QCL frequency combs for use in dual-comb spectroscopy allows the recording of broadband spectra with high time resolutions providing the potential for single-shot investigations of irreversible protein systems.

After successfully establishing laser-based IR spectroscopy techniques in the form of research-based laboratory setups, the first compact table-top instruments already became commercially available. Future developments may be directed toward further miniaturization of the setups to monolithically integrated devices that comprise a QCL light source, a waveguide for sample interaction, and a suitable detector, possibly a quantum cascade detector. The resulting on-chip sensors could provide fast available readouts that

could be beneficial, for example, in industrial (protein analysis in cow milk) or medical applications (protein analysis in blood or human breast milk) and could even supersede currently used standard methods that involve several manual processing steps.

References

[1] J. Faist, F. Capasso, D.L. Sivco, C. Sirtori, A.L. Hutchinson, A.Y. Cho, Quantum cascade laser, Science 264 (5158) (1994) 553–556.

[2] J. Faist, Quantum Cascade Lasers, Oxford University Press, Oxford, UK, 2013.

[3] A. Schwaighofer, M. Brandstetter, B. Lendl, Quantum cascade lasers (QCLs) in biomedical spectroscopy, Chem. Soc. Rev. 46 (2017) 5903–5924.

[4] R. Pecharroman-Gallego, An overview on quantum cascade lasers: origins and development, in: V.N. Stavrou (Ed.), Quantum Cascade Lasers, InTechOpen, Rijeka, 2017, pp. 3–24.

[5] K. Isensee, N. Kröger-Lui, W. Petrich, Biomedical applications of mid-infrared quantum cascade lasers—a review, Analyst 143 (24) (2018) 5888–5911.

[6] B. Meng, Q.J. Wang, Broadly tunable single-mode mid-infrared quantum cascade lasers, J. Opt. 17 (2) (2015) 023001.

[7] T. Kruczek, K.A. Fedorova, G.S. Sokolovskii, R. Teissier, A.N. Baranov, E.U. Rafailov, InAs/AlSb widely tunable external cavity quantum cascade laser around 3.2 mu m, Appl. Phys. Lett. 102 (1) (2013) 011124.

[8] H. Nguyen-Van, A.N. Baranov, Z. Loghmari, L. Cerutti, J.-B. Rodriguez, J. Tournet, et al., Quantum cascade lasers grown on silicon, Sci. Rep. 8 (1) (2018) 7206.

[9] P. Fuchs, J. Friedl, S. Hofling, J. Koeth, A. Forchel, L. Worschech, et al., Single mode quantum cascade lasers with shallow-etched distributed Bragg reflector, Opt. Express 20 (4) (2012) 3890–3897.

[10] G. Duxbury, N. Langford, M.T. McCulloch, S. Wright, Quantum cascade semiconductor infrared and far-infrared lasers: from trace gas sensing to non-linear optics, Chem. Soc. Rev. 34 (11) (2005) 921–934.

[11] J.P. Waclawek, R. Lewicki, H. Moser, M. Brandstetter, F.K. Tittel, B. Lendl, Quartz-enhanced photoacoustic spectroscopy-based sensor system for sulfur dioxide detection using a CW DFB-QCL, Appl. Phys. B Lasers Opt. 117 (1) (2014) 113–120.

[12] R. Lewicki, M. Witinski, B. Li, G. Wysocki, Spectroscopic benzene detection using a broadband monolithic DFB-QCL array, in: A. Belyanin, P.M. Smowton (Eds.), Spectroscopic Benzene Detection Using a Broadband Monolithic DFB-QCL Array, SPIE, San Francisco, USA, 2016.

[13] S. Riedi, A. Hugi, A. Bismuto, M. Beck, J. Faist, Broadband external cavity tuning in the 3-4 mu m window, Appl. Phys. Lett. 103 (3) (2013) 031108.

[14] A. Hugi, R. Maulini, J. Faist, External cavity quantum cascade laser, Semicond. Sci. Technol. 25 (8) (2010) 083001.

[15] F. Fuchs, B. Kirn, C. Mann, Q. Yang, W. Bronner, B. Raynor, et al., Spectral tuning and mode competition of quantum cascade lasers studied by time-resolved Fourier transform spectroscopy, in: M. Analoui, A. Belyanin, R.A. Drezek, C. Gmachl, J.P. Robinson (Eds.), Spectral Tuning and Mode Competition of Quantum Cascade Lasers Studied by Time-Resolved Fourier Transform Spectroscopy, SPIE, Boston, USA, 2006.

[16] J. Faist, G. Villares, G. Scalari, M. Rosch, C. Bonzon, A. Hugi, et al., Quantum Cascade Laser Frequency Combs, Nanophotonics 5 (2) (2016) 272–291.

[17] E.T.H. Zürich, Quantum Optoelectronics Group, Frequency Combs. <https://www.ethz.ch/content/specialinterest/phys/quantum-electronics/quantum-optoelectronics/en/research/frequency_combs.html>, 2019 (accessed 02.05.19).

[18] M. Brandstetter, A. Genner, K. Anic, B. Lendl, Tunable external cavity quantum cascade laser for the simultaneous determination of glucose and lactate in aqueous phase, Analyst 135 (12) (2010) 3260–3265.

[19] H. Fabian, W. Mäntele, Infrared spectroscopy of proteins, Handbook of Vibrational Spectroscopy, John Wiley & Sons, Ltd, Hoboken, NJ, USA, 2006.
[20] M. Brandstetter, C. Koch, A. Genner, B. Lendl, Measures for optimizing pulsed EC-QC laser spectroscopy of liquids and application to multi-analyte blood analysis, Proc. SPIE 8993 (2014) 89931u.
[21] S.Y. Venyaminov, F.G. Prendergast, Water (H2O and D2O) molar absorptivity in the 1000-4000 cm(-1) range and quantitative infrared spectroscopy of aqueous solutions, Anal. Biochem. 248 (2) (1997) 234–245.
[22] M. Brandstetter, L. Volgger, A. Genner, C. Jungbauer, B. Lendl, Direct determination of glucose, lactate and triglycerides in blood serum by a tunable quantum cascade laser-based mid-IR sensor, Appl. Phys. B Lasers Opt. 110 (2) (2013) 233–239.
[23] J. Ropcke, P.B. Davies, N. Lang, A. Rousseau, S. Welzel, Applications of quantum cascade lasers in plasma diagnostics: a review, J. Phys. D Appl. Phys 45 (42) (2012) 423001.
[24] J. Hayden, S. Hugger, F. Fuchs, B. Lendl, A quantum cascade laser-based Mach–Zehnder interferometer for chemical sensing employing molecular absorption and dispersion, Appl. Phys. B Lasers Opt. 124 (2) (2018) 29.
[25] G. Villares, J. Wolf, D. Kazakov, M.J. Suess, A. Hugi, M. Beck, et al., On-chip dual-comb based on quantum cascade laser frequency combs, Appl. Phys. Lett. 107 (25) (2015) 251104.
[26] G. Villares, A. Hugi, S. Blaser, J. Faist, Dual-comb spectroscopy based on quantum-cascade-laser frequency combs, Nat. Commun. 5 (2014) 5192.
[27] J.L. Klocke, M. Mangold, P. Allmendinger, A. Hugi, M. Geiser, P. Jouy, et al., Single-shot sub-microsecond mid-infrared spectroscopy on protein reactions with quantum cascade laser frequency combs, Anal. Chem. 90 (17) (2018) 10494–10500.
[28] S. Lindner, J. Hayden, A. Schwaighofer, T. Wolflehner, C. Kristament, M.G. Gonzalez-Cabrera, et al., External cavity quantum cascade laser based mid-infrared dispersion spectroscopy for qualitative and quantitative analysis of liquid-phase samples, Anal. Spectrosc. (2019).
[29] A. Rüther, M. Pfeifer, V.A. Lórenz-Fonfría, S. Lüdeke, pH titration monitored by quantum cascade laser-based vibrational circular dichroism, J. Phys. Chem. B 118 (14) (2014) 3941–3949.
[30] J. Haas, A. Müller, L. Sykora, B. Mizaikoff, Analytical performance of μ-groove silicon attenuated total reflection waveguides, Analyst 144 (2019) 3398–3404.
[31] Á.I. López-Lorente, P. Wang, M. Sieger, E. Vargas Catalan, M. Karlsson, F. Nikolajeff, et al., Mid-infrared thin-film diamond waveguides combined with tunable quantum cascade lasers for analyzing the secondary structure of proteins, Phys. Status Solidi A 213 (2016) 2117–2123.
[32] S. Bal Ram, Basic aspects of the technique and applications of infrared spectroscopy of peptides and proteins, Infrared Analysis of Peptides and Proteins, American Chemical Society, 1999, pp. 2–37.
[33] B. Srour, S. Bruechert, S.L.A. Andrade, P. Hellwig, Secondary structure determination by means of ATR-FTIR spectroscopy, in: J.-J. Lacapere (Ed.), Membrane Protein Structure and Function Characterization: Methods and Protocols, Springer New York, New York, NY, 2017, pp. 195–203.
[34] K.E. Wilcox, E.W. Blanch, A.J. Doig, Determination of protein secondary structure from infra-red spectra using partial least squares regression, Biochemistry 55 (27) (2016) 3794–3802.
[35] M. Shariati-Rad, M. Hasani, Application of multivariate curve resolution-alternating least squares (MCR-ALS) for secondary structure resolving of proteins, Biochimie 91 (7) (2009) 850–856.
[36] M.R. Alcaráz, A. Schwaighofer, H. Goicoechea, B. Lendl, Application of MCR-ALS to reveal intermediate conformations in the thermally induced α-β transition of poly-l-lysine monitored by FT-IR spectroscopy, Spectrochim. Acta Part A 185 (2017) 304–309.
[37] F. Dousseau, M. Pezolet, Determination of the secondary structure content of proteins in aqueous solutions from their amide I and amide II infrared bands. Comparison between classical and partial least-squares methods, Biochemistry 29 (37) (1990) 8771–8779.
[38] S. Navea, R. Tauler, E. Goormaghtigh, A. de Juan, Chemometric tools for classification and elucidation of protein secondary structure from infrared and circular dichroism spectroscopic measurements, Proteins 63 (3) (2006) 527–541.

[39] M.R. Alcaráz, A. Schwaighofer, C. Kristament, G. Ramer, M. Brandstetter, H. Goicoechea, et al., External cavity-quantum cascade laser spectroscopy for mid-IR transmission measurements of proteins in aqueous solution, Anal. Chem. 87 (13) (2015) 6980−6987.
[40] M.R. Alcaráz, A. Schwaighofer, H. Goicoechea, B. Lendl, EC-QCL mid-IR transmission spectroscopy for monitoring dynamic changes of protein secondary structure in aqueous solution on the example of beta-aggregation in alcohol-denaturated alpha-chymotrypsin, Anal. Bioanal. Chem. 408 (15) (2016) 3933−3941.
[41] A. Schwaighofer, M.R. Alcaráz, C. Araman, H. Goicoechea, B. Lendl, External cavity-quantum cascade laser infrared spectroscopy for secondary structure analysis of proteins at low concentrations, Sci. Rep. 6 (2016) 33556.
[42] A. Schwaighofer, M.R. Alcaráz, J. Kuligowski, B. Lendl, Recent advancements of EC-QCL based mid-IR transmission spectroscopy of proteins and application to analysis of bovine milk, Biomed. Spectrosc. Imaging 7 (1) (2018) 35−45.
[43] J. Kuligowski, A. Schwaighofer, M.R. Alcaráz, G. Quintás, H. Mayer, M. Vento, et al., External cavity-quantum cascade laser (EC-QCL) spectroscopy for protein analysis in bovine milk, Anal. Chim. Acta 963 (2017) 99−105.
[44] A. Schwaighofer, J. Kuligowski, G. Quintás, H.K. Mayer, B. Lendl, Fast quantification of bovine milk proteins employing external cavity-quantum cascade laser spectroscopy, Food Chem. 252 (2018) 22−27.
[45] A. Schwaighofer, M. Montemurro, S. Freitag, C. Kristament, M.J. Culzoni, B. Lendl, Beyond FT-IR spectroscopy: EC-QCL based mid-IR transmission spectroscopy of proteins in the amide I and amide II region, Anal. Chem. 90 (11) (2018) 7072−7079.
[46] M. Montemurro, A. Schwaighofer, A. Schmidt, J. Culzoni, H.K. Mayer, B. Lendl, High-throughput quantitation of bovine milk proteins and discrimination of commercial milk types by external cavity-quantum cascade laser spectroscopy and chemometrics, Analyst 144 (2019) 5571−5579.
[47] E. Ma, L.B. Wang, B. Kendrick, Enhanced Protein structural characterization using microfluidic modulation, Spectroscopy 33 (7) (2018) 46−52.
[48] K.S. Shah, Applications of microfluidic modulation spectroscopy for antibody-drug conjugate structural characterization, Am. Pharm. Rev. 21 (6) (2018) 120−125.
[49] M. Di Donato, M.L. Groot, Ultrafast infrared spectroscopy in photosynthesis, Biochim. Biophys. Acta Bioenerg. 1847 (1) (2015) 2−11.
[50] T. Kottke, V.A. Lórenz-Fonfría, J. Heberle, The grateful infrared—sequential protein structural changes resolved by ir difference spectroscopy, J. Phys. Chem. B 121 (2) (2016) 335−350.
[51] A. Mezzetti, W. Leibl, Time-resolved infrared spectroscopy in the study of photosynthetic systems, Photosynth. Res. 131 (2) (2017) 121−144.
[52] E. Ritter, L. Puskar, F.J. Bartl, E.F. Aziz, P. Hegemann, U. Schade, Time-resolved infrared spectroscopic techniques as applied to channelrhodopsin, Front. Mol. Biosci. 2 (2015) 38.
[53] G.D. Smith, R.A. Palmer, Fast time-resolved mid-infrared spectroscopy using an interferometer, Handbook of Vibrational Spectroscopy, John Wiley & Sons, Ltd, 2006.
[54] V.A. Lorenz-Fonfria, T. Resler, N. Krause, M. Nack, M. Gossing, G.F. von Mollard, et al., Transient protonation changes in channelrhodopsin-2 and their relevance to channel gating, Proc. Natl Acad. Sci. U.S.A 110 (14) (2013) E1273−E1281.
[55] C. Kotting, K. Gerwert, Proteins in action monitored by time-resolved FTIR spectroscopy, ChemPhysChem 6 (5) (2005) 881−888.
[56] J. Heberle, Time-resolved and surface-enhanced infrared spectroscopy, Pumps Channels Transp. (2015).
[57] B.J. Schultz, H. Mohrmann, V.A. Lorenz-Fonfria, J. Heberle, Protein dynamics observed by tunable mid-IR quantum cascade lasers across the time range from 10ns to 1s, Spectrochim. Acta Part A 188 (2018) 666−674.
[58] M. Schleeger, C. Wagner, M.J. Vellekoop, B. Lendl, J. Heberle, Time-resolved flow-flash FT-IR difference spectroscopy: the kinetics of CO photodissociation from myoglobin revisited, Anal. Bioanal. Chem. 394 (7) (2009) 1869−1877.

[59] T. Resler, B.-J. Schultz, V.íctor A. Lórenz-Fonfría, R. Schlesinger, J. Heberle, Kinetic and vibrational isotope effects of proton transfer reactions in channelrhodopsin-2, Biophys. J. 109 (2) (2015) 287−297.
[60] V.A. Lórenz-Fonfría, B.-J. Schultz, T. Resler, R. Schlesinger, C. Bamann, E. Bamberg, et al., Pre-gating conformational changes in the ChETA variant of channelrhodopsin-2 monitored by nanosecond IR spectroscopy, J. Am. Chem. Soc. 137 (5) (2015) 1850−1861.
[61] M.L. Donten, S. Hassan, A. Popp, J. Halter, K. Hauser, P. Hamm, pH-Jump induced leucine zipper folding beyond the diffusion limit, J. Phys. Chem. B 119 (4) (2015) 1425−1432.
[62] A. Popp, D. Scheerer, B. Heck, K. Hauser, Biomolecular dynamics studied with IR-spectroscopy using quantum cascade lasers combined with nanosecond perturbation techniques, Spectrochim. Acta Part A 181 (2017) 192−199.
[63] A. Popp, D. Scheerer, H. Chi, T.A. Keiderling, K. Hauser, Site-specific dynamics of -sheet peptides with (D)Pro-Gly turns probed by laser-excited temperature-jump infrared spectroscopy, ChemPhysChem 17 (9) (2016) 1273−1280.
[64] A. Popp, L. Wu, T.A. Keiderling, K. Hauser, Effect of hydrophobic interactions on the folding mechanism of β-hairpins, J. Phys. Chem. A 118 (49) (2014) 14234−14242.
[65] D. Scheerer, H. Chi, D. McElheny, T.A. Keiderling, K. Hauser, Isotopically site-selected dynamics of a three-stranded β-sheet peptide detected with temperature-jump infrared-spectroscopy, J. Phys. Chem. B 122 (46) (2018) 10445−10454.
[66] D. Scheerer, H. Chi, D. McElheny, A. Samer, T.A. Keiderling, K. Hauser, Role of aromatic cross-links in structure and dynamics of model three-stranded β-sheet peptides, J. Phys. Chem. A 122 (2) (2018) 543−553.
[67] I. Coddington, N. Newbury, W. Swann, Dual-comb spectroscopy, Optica 3 (4) (2016) 414−426.
[68] A. Hugi, G. Villares, S. Blaser, H.C. Liu, J. Faist, Mid-infrared frequency comb based on a quantum cascade laser, Nature 492 (7428) (2012) 229−233.

CHAPTER 4

Theoretical simulation of protein two-dimensional infrared spectroscopy

Zhijun Pan and Wei Zhuang

State Key Laboratory of Structural Chemistry, Fujian Institute of Research on the Structure of Matter, Chinese Academy of Sciences, Fuzhou, China

4.1 Introduction

Understanding protein conformational dynamics is one fundamental issue in the biophysical sciences. Extensive research efforts have been devoted in this field [1–14]. A number of structural characterization techniques, including X-ray crystallography [15–18], nuclear magnetic resonance (NMR) [8,19–21], fluorescence [22–24], infrared (IR) [25–29], Raman [30–34], circular dichroism [35–39], single molecule spectroscopy [40–42], and cryogenic electron microscopy [43–45], have been developed to study protein conformational dynamics.

Infrared vibrational spectroscopy, which probes vibrational transitions and their responses to the surrounding environment, is one of the most commonly used spectroscopic tools for the characterization of protein conformations and their variations [46–50]. Vibrational transitions are sensitive to the local molecular structure as well as the hydrogen bonding and electrostatic environments [51–56]. They are therefore ideal for distinguishing between various protein secondary structural motifs and monitoring the effects of changing environments.

Four vibrational bands related to the localized amide vibrations on each peptide unit, namely the amide I, II, III, and A bands, are commonly used as the probes in protein infrared spectroscopy [48,57–61]. The amide I band between 1600 and 1700 cm^{-1} mostly originates from the stretching motion of the C = O peptide bond (coupled to in-phase N–H bending and C–H stretching, see Fig. 4.2) and is particularly useful for the investigation of peptide structure due to its strong transition dipole moment, high sensitivity to the secondary structure, as well as the large separation from other vibrational bands. The amide I band of the α-helix motif absorbs near 1650 cm^{-1}, while that of the β-sheets absorbs at 1620 and 1675 cm^{-1} [52,62–65]. The up to 20 cm^{-1} variation of the amide I frequency for

Vibrational Spectroscopy in Protein Research.
DOI: https://doi.org/10.1016/B978-0-12-818610-7.00004-9

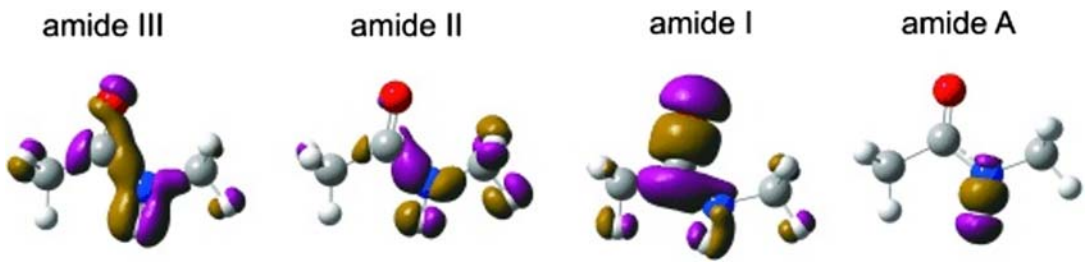


Figure 4.1
Transition charge densities (TCD) for the four amide modes of *N*-methyl acetamide (NMA). The 0.01 esu/Bohr contour is shown. Violet and brown contours represent positive and negative values, respectively. Source: *Reprinted from W. Zhuang, T. Hayashi, S. Mukamel, Coherent multidimensional vibrational spectroscopy of biomolecules: concepts, simulations, and challenges, Angew. Chem. Int. Ed. 48 (21) (2009), 3750–3781 [66].*

different secondary structures is widely used as a marker in polypeptide and protein structure determination. Other useful amide vibrational bands (Fig. 4.1) include the amide II ($\sim 1500\ \mathrm{cm}^{-1}$) band from the bending motion of the C–N–H coupled to the mainly C–N stretching, the amide III ($\sim 1200\ \mathrm{cm}^{-1}$) band related to the C–N stretching coupled to the mainly the C–N–H bending motion, as well as the amide A ($\sim 3500\ \mathrm{cm}^{-1}$) band from the N–H bond stretching. Compared with the amide I band, the amide II and III bands are more delocalized (with the amide III band typically weaker than the amide I and II bands), and the amide A has a much smaller transition dipole.

In 1950 Ambrose adopted the IR technique to detect the α–β transformation procedure of insulin [25]. Krimm and coworkers later investigated theoretically the peptide vibrational bands [67–69]. Following the invention of the world's first commercial Fourier-transform infrared (FTIR) spectrometer, made by Digilab, higher signal-to-noise ratio amide bands in the protein were measured [70,71]. Femtosecond laser pulses generated from mode-locked oscillators then became the ideal tools for IR detection [72,73]. Hamm and Hochstrasser [74] in 1998 carried out the first frequency–frequency two-dimensional infrared (2DIR) measurement of apamin, scyllatoxin, and bovine pancreatic trypsin inhibitor. 2DIR spectroscopy, as a cutting-age dynamic infrared spectroscopic technique, enhances the resolution of the peptide conformation compared with the conventional FTIR technique [57,59,66,75–84]. A heterodyne-detected 2DIR experiment has three laser pulses (with wave vectors k_1, k_2, and k_3) that interact with the peptide [57,59,75–78]. The signal $S(t_3; t_2; t_1)$ is then generated coherently in the directions $k_4 = \pm k_1 \pm k_2 \pm k_3$. The signal depends parametrically on the time intervals between pulses t_1, t_2, and t_3, and is highly oscillatory. For a clearer picture, the signal is usually double Fourier-transformed with respect to two time variables to generate a frequency/frequency correlation plot such as $S(\omega_1;t_2;\omega_3)$, where ω_1 and ω_3 are the frequency conjugates to t_1 and t_3 (holding t_2 fixed).

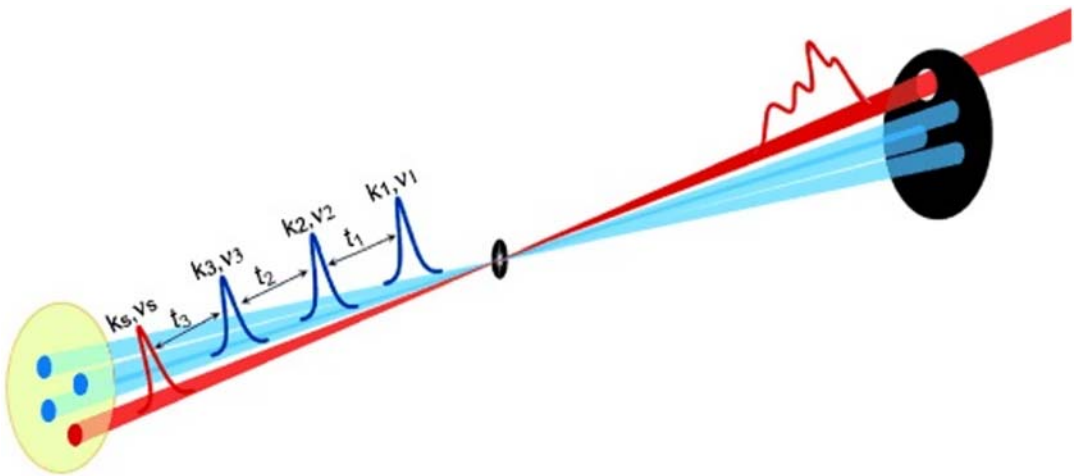


Figure 4.2
Schematic experimental setup for a heterodyne-detected four-wave mixing experiment. Signals are recorded as the function of three time delays and displayed as two-dimensional (2D) correlation plots involving the double Fourier transform of two time delays, holding the third fixed. Source: *Reprinted from W. Zhuang, T. Hayashi, S. Mukamel, Coherent multidimensional vibrational spectroscopy of biomolecules: concepts, simulations, and challenges, Angew. Chem. Int. Ed. 48 (21) (2009), 3750–3781 [66].*

In the 2DIR signals, new resonances are generated by the coupled vibrational modes at the combination of the single mode frequencies, which lead to the characteristic features in the off-diagonal regions (so-called "cross peaks"). The intensities and profiles of these cross peaks give well resolved zero-background information about the correlations between transitions. Correlation plots of dynamical events occurring within the controlled evolution time periods can be described via terms of multipoint correlation functions. These carry considerably more information than the two-point correlation functions of linear spectroscopy. Therefore the 2DIR signals have the potential to distinguish between many different protein configurations whose 1D responses are virtually identical. 2DIR is a rapidly expanding class of new ultrafast coherent vibrational spectroscopies with broad applications in studies of peptides [57,75,76], proteins [58, 59,77,78], ion channels [81–84], DNA [85,86] chemical exchange kinetics [87–89], hydrogen bonding [90–93], vibrational energy transport dynamics [94,95], and chemical reaction mechanisms [96,97] (Fig. 4.2).

Due to the congested information of optical transitions, the interpretation of peptide vibrational spectra can benefit greatly from the theoretical modeling. Since 2DIR and atomistic level MD simulations can be carried out on the same timescale, it becomes highly desirable to develop MD-based theoretical tools for simulating and analyzing the 2DIR

signals, which help assign the 2DIR features and reveal the underlying fast dynamics. In the following section, we introduce the computational protocols of simulating peptide 2DIR based on MD simulation.

4.2 Theoretical simulation

In the commonly applied theoretical tools for protein 2DIR simulations, a trajectory of the system under study (protein and solvent) is first created using the molecular dynamics (MD) simulation. The effective fluctuating vibrational-exciton Hamiltonian, $\hat{H}_s(t)$ and the transition dipole matrix $\mu(t)$ for the relevant states are then constructed for each snapshot in the trajectory. For the peptide amide vibrations this is usually carried out based on the FE model [98], which was originally developed to describe the coupled localized transitions on the oligomers or polymers with similar repeat units. The parameters required are generated from the electronic structure calculations of individual chromophores, instead of the whole system. This treatment greatly reduces the computational demands.

Using the calculated fluctuating Hamiltonian, the four-point transition dipole correlation functions, representing the quantum Liouville-space pathways [99] relevant for the chosen technique, are constructed. The response function as well as the signals are then calculated by taking the proper combinations of the four-point correlation functions. The spectra calculated consist of well-separated bands of energy levels representing single excitations, double excitations, and higher excitations. The molecular Hamiltonian conserves the number of excitations, which can only be altered by the optical fields. The lowest (single-exciton) manifold is accessible by linear optical techniques such as absorption spectroscopy, whereas the doubly excited (two-exciton) and higher manifolds only show up in the nonlinear spectra. In the following we use the peptide amide I band as the example to discuss the details of this procedure.

4.2.1 Hamiltonian construction

Calculation of the fluctuating frequency begins by partitioning the total Hamiltonian into the system Hamiltonian H_S and the system−field interaction H_F:

$$\hat{H} = \hat{H}_s + \hat{H}_F \tag{4.1}$$

the system Hamiltonian is:

$$\hat{H}_S = \sum_m \varepsilon_m \hat{B}_m^\dagger \hat{B}_m + \sum_{mn}^{m\neq n} J_{mn} \hat{B}_m^\dagger \hat{B}_m - \frac{1}{2}\sum_m \Delta_m \hat{B}_m^\dagger \hat{B}_m^\dagger \hat{B}_n \hat{B}_n \tag{4.2}$$

and the interaction Hamiltonian is:

$$\hat{H}_F = -E(t)\sum_m \mu_m(\hat{B}_m^\dagger + \hat{B}_n) \tag{4.3}$$

The first two terms in $\hat{H}_s$ describe the free excitons where ε_m is the local vibrational frequency and J_{mn} is the harmonic intersite couplings which induces exciton hopping. Δ_m and μ_m represent the anharmonicity and transition dipole moment, respectively. $\hat{B}_m^\dagger$ and $\hat{B}_m$ are creation and annihilation Boson operators for the mth mode. In the Cartesian coordinates $\hat{B}_m^\dagger$ and $\hat{B}_m$ are expressed as $\hat{B}_m^\dagger = \frac{1}{\sqrt{2}}(q_m - ip_m)$ and $\hat{B}_m = \frac{1}{\sqrt{2}}(q_m + ip_m)$, which satisfy the Bose commutation relations $[\hat{B}_m, \hat{B}_n^\dagger] = \delta_{mn}$.

4.2.1.1 Vibrational frequency ε_m

Frequencies of the local amide I modes, ε_m, are related to the surrounding (especially electrostatic) environment. The Stark model describes the local electrostatic environment by an electric field/potential at a reference point [100], then a time-dependence of $\Delta\omega(t) = kE(t)$ is used to derive the instant frequency. Such a simple model works fairly well for the small molecules [101–103].

For large systems such as the proteins, on the other hand, the nonuniform electric field across the peptide bond should be taken into account. An analytical frequency-field/potential relation can be established based on a large number of ab initio solute/solvent cluster calculations [104–106]. The frequencies for the target vibrational mode of the solute molecule are generated for all cluster configurations, and then fitted to the electric fields imposed on the solute from the solvent molecules to determine the frequency-field relation (the "ab initio map"). Using this map, a fluctuating frequency trajectory is conveniently generated, based on the MD simulation result.

Minhaeng Cho, for instance, calculated ε_m as a linear function of the electrostatic potentials at the C, O, N, and H atoms (CHO_4) [62], or adding two more methyl sites (CHO_6) [56]. The least squares fit is used for the normal mode frequencies of *N*-methyl acetamide (NMA)-water clusters calculated at the restricted Hartree–Fock (RHF) level. For the (CHO_4) map [62], as an example, the frequency of the mth local amide I mode is:

$$\omega_m = \omega_0 + \sum_{j=1}^{4} l_{j(m)}\phi_{j(m)} \tag{4.4}$$

where the coefficient is

$$l_{j(m)} = \frac{g_I}{4\pi c M_I^2(\omega_I^0)^3}\left(\frac{\partial c_{j(m)}}{\partial Q_m}\right)_0^{\text{eff}} \tag{4.5}$$

and the electric field is

$$\phi_{j(m)} = \frac{1}{4\pi\epsilon_0} \sum_{n \neq m}^{N} \sum_{k(n)} \frac{c_{k(n)}}{r_{j(m),k(n)}} \tag{4.6}$$

Here, $C_{k(n)}$ is the partial charge of the kth site of the nth peptide, $r_{j(m),k(n)}$ is the distance between the jth site of the mth peptide and the kth site of the nth peptide.

Similarly, Skinner [51] constructed a map of the *N*-methyl acetamide-D; $CH_3(CO)ND(CH_3)$ amide I frequency, where electric fields instead of potential are fitted at the C, O, N, and H atoms. Watson and Hirst proposed that the accuracy of NMA amide I frequencies in water can be improved by adding sampling points at the middle points of CO, CN, and NH [52]. Hayashi et al. [53] have constructed a density functional theory (DFT) electrostatic map, where the electric potential on the NMA unit is expanded to sixth order in the five normal coordinates, and requires 19 independent parameters to fit. Jansen and Knoester [54] also constructed the amide frequency map, in which the frequency is parameterized by the electric field and gradients at the C, O, N, and H atoms.

As the generalization and rationalization of this treatment, solvent-induced perturbations of molecular spectroscopic properties can be closely linked to the potential of mean force imposed on a solute by the solvent molecules. This intermolecular interaction-induced spectroscopic phenomenon is called the solvatochromism [107]. Various theories [108–112] have been proposed to consider the attractive and repulsive electrostatic interactions underlying the vibrational frequency shift. Ben-Amotz [108] summarized that the attractive solvation forces contribute a mean-field perturbation with an almost linear density dependence. The repulsive force, on the other hand, is induced by a change in pressure. As an observation of vibrational solvatochromism in proteins, Eaton et al. [113] found that the amide I mode frequency of the NMA in liquid water exhibits a strong redshift by $\sim 80\ \mathrm{cm}^{-1}$, in comparison to that of the gas-phase NMA molecule.

To rationalize the origin of the frequency shift observed in the peptide amide I vibrational band in water, Minhaeng Cho developed a systematic theory based on the coarse-grained model of vibrational solvatochromism combined with the effective fragment potential (EFP) method [114]. In this theory, the overall frequency shift is considered to be contributed by five components: $\Delta\omega_j^{\mathrm{EFP}} = \Delta\omega_j^{\mathrm{Coul}} + \Delta\omega_j^{\mathrm{Rep}} + \Delta\omega_j^{\mathrm{Ind}} + \Delta\omega_j^{\mathrm{Disp}} + \Delta\omega_j^{\mathrm{CT}}$. In which $\Delta\omega_j^{\mathrm{Coul}}$ is the Coulomb interaction between unperturbed charge densities of solute and solvent molecules; $\Delta\omega_j^{\mathrm{Rep}}$ is the exchange–repulsion interaction; $\Delta\omega_j^{\mathrm{Ind}}$ is the second-order induction; $\Delta\omega_j^{\mathrm{Disp}}$ is the dispersion interactions; and $\Delta\omega_j^{\mathrm{CT}}$ is the charge transfer effect.

The charge transfer contribution, $\Delta\omega_j^{\mathrm{CT}}$, is negligible for the NMA amide I vibrational frequency shift by water. The electrostatic interaction energy, $\Delta\omega_j^{\mathrm{Coul}}$, is calculated by considering the distributed multipole expansion of electron charge density distributions: [104]

$$\Delta\omega_j^{\mathrm{Coul}} = \Delta\omega_j^{\mathrm{SolCAMM}} - \frac{1}{2M_j\omega_j}\sum_{x\in\text{solute}}\sum_{y\in\text{solvent}}\left[\sum_i \frac{g_{ijj}}{M_i\omega_i^2}\sum_{r^{-n}\text{terms}} f_i^{(x,y)} - \sum_{r^{-n}\text{terms}} K_{jj}^{(x,y)}\right] \tag{4.7}$$

where $\Delta\omega_j^{\mathrm{SolCAMM}}$ is the Coulomb interaction-induced frequency shift calculated from cumulative atomic multipole moments (CAMM) [115]. The potential derivative coefficients are defined as

$$f_i = \left(\frac{\partial U}{\partial Q_i}\right)_{\mathbf{Q_0}}, K_{ij} = \left(\frac{\partial^2 U}{\partial Q_i \partial Q_j}\right)_{\mathbf{Q_0}}, g_{ijk} = \left(\frac{\partial^3 U}{\partial Q_i \partial Q_j \partial Q_k}\right)_{\mathbf{Q_0}} \tag{4.8}$$

where U is the solute−solvent intermolecular interaction energy, V is the total energy of the solute, and M_j and ω_j are the reduced mass and harmonic frequency of the jth normal mode with coordinate Q_j.

The exchange−repulsion interaction $\Delta\omega_j^{\mathrm{Rep}}$ between molecules X (solute) and Y (solvent) can be given approximately as [105,106]:

$$\begin{aligned} U_{\mathrm{Ex-rep}} = \sum_{a\in X}\sum_{b\in Y}\Bigg[& -4\sqrt{\frac{-2\ln|S_{ab}|}{\pi}}\frac{S_{ab}^2}{r_{ab}} - 2S_{ab}\left(\sum_{c\in X} F_{ac}^X S_{cb} + \sum_{d\in Y} F_{bd}^Y S_{da} - 2T_{ab}\right) \\ & + 2S_{ab}^2\left(-\sum_{y\in Y}\frac{Z_y}{r_{ay}} - \sum_{x\in X}\frac{Z_x}{r_{xb}} + 2\sum_{d\in Y}\frac{1}{r_{ad}} + 2\sum_{c\in X}\frac{1}{r_{cb}} - \frac{1}{r_{ab}}\right)\Bigg] \end{aligned} \tag{4.9}$$

where S_{ab} and T_{ab} are the one-electron overlap and the kinetic integrals between the ath and bth localized molecular orbital (LMO), $F_{ab}^{X(Y)}$ is the Fock matrix element of molecule $X(Y)$, and $Z_{x(y)}$ is the atomic number of the $x(y)$th atom. The relative distance r is defined as $r_{xy} = |r_x - r_y|$. Here, the indices a, b, c, and d refer to the LMOs, whereas indices x and y refer to the atoms in molecules X and Y, respectively. Note also that the ath and cth LMOs belong to the solute and bth and dth LMOs to the solvent molecule.

The frequency shift originating from solute−solvent repulsive interactions can be cast as:

$$\Delta\omega_j^{\mathrm{Ex-rep}} = -\frac{1}{2M_j\omega_j}\sum_i \frac{f_i^{\mathrm{Ex-rep}} g_{ijj}}{M_i\omega_i^2} - K_{jj}^{\mathrm{Ex-rep}} \approx -\frac{1}{2M_j\omega_j}\sum_i \frac{f_i^{\mathrm{Ex-rep}} g_{ijj}}{M_i\omega_i^2} \tag{4.10}$$

Similarly, the definitions of $f_i^{\mathrm{Ex-rep}}$ and $K_{jj}^{\mathrm{Ex-rep}}$ are given by the derivative from the relative potential term.

The repulsive component of the solvatochromic frequency shift is

$$\Delta\omega_j^{\mathrm{Ex-rep}} \approx -\frac{1}{2M_j\omega_j}\sum_i \frac{g_{ijj}}{M_i\omega_i^2}\frac{\partial}{\partial Q_i}\left\{\sum_{a\in X}\sum_{b\in Y}\left[-4\sqrt{\frac{-2\ln|S_{ab}|}{\pi}}\frac{S_{ab}^2}{r_{ab}} - 2S_{ab}\left(\sum_{c\in X}F_{ac}^X S_{cb} + \sum_{d\in Y}F_{bd}^Y S_{da} - 2T_{ab}\right) + 2S_{ab}^2\left(-\sum_{y\in Y}\frac{Z_y}{r_{ay}} - \sum_{x\in X}\frac{Z_x}{r_{xb}} + 2\sum_{d\in Y}\frac{1}{r_{ad}} + 2\sum_{c\in X}\frac{1}{r_{cb}} - \frac{1}{r_{ab}}\right)\right]\right\}_{\mathbf{Q_0}} \tag{4.11}$$

The polarization interaction energy in the distributed dipole−dipole static polarizability approximation can be simply recast in the following form:

$$U_{\mathrm{Pol}} = -\frac{1}{2}\mathbf{p}\cdot\mathbf{F} \tag{4.12}$$

where **p** is the induced dipole vectors at polarizability centers, while **F** is the electric field vector, created by the distributed static multipole moments of interacting molecules at each polarizable center. The induction component of the frequency shift is therefore:

$$\Delta\omega_j^{\mathrm{Ind}} = -\frac{1}{4M_j\omega_j}\sum_i \frac{g_{ijj}}{M_i\omega_i^2}\left\{\left(\frac{\partial\mathbf{F}^T}{\partial Q_i}\right)_{\mathbf{Q_0}}\cdot\left[\mathbf{D}^{-1} + \left(\mathbf{D}^{-1}\right)^T\right] - \mathbf{F}^T\cdot\mathbf{D}^{-1}\left(\frac{\partial\mathbf{D}}{\partial Q_i}\right)_{\mathbf{Q_0}}\mathbf{D}^{-1}\right\}\cdot\mathbf{F} \tag{4.13}$$

Here, matrix **D** is defined as the second rank dipole−dipole interaction tensor.

The dispersion interaction-induced frequency shift is proportional to the isotropic polarizability α_S of a solvent molecule, $\Delta\omega_j^{\mathrm{Disp}} = A_\alpha\rho_S\alpha_S$, where ρ_S is the solvent density and A_α represents the proportionality constant from experiment data. Applying multipole expansion on this relation, and considering only the isotropic part of the dipole−dipole polarizabilities leads to:

$$\Delta\omega_{j,6}^{\mathrm{Disp}} \approx \frac{\hbar}{4M_{j,0}\omega_{j,0}}\sum_i \frac{g_{ijj}}{M_{i,0}\omega_{i,0}^2}\sum_{a\in A}\sum_{b\in B}\sum_{\alpha\beta\gamma\delta}^{x,y,z}\left[T_{\alpha\beta}^{ab}T_{\gamma\delta}^{ab}\int_0^\infty\left(\frac{\partial\alpha_{\alpha\gamma}^b(i\Omega)}{\partial Q_i}\right)_{\mathbf{Q_{0A}}}\alpha_{\beta\delta}^b(i\Omega)d\Omega + \left(\frac{\partial\left(T_{\alpha\beta}^{ab}T_{\gamma\delta}^{ab}\right)}{\partial Q_i}\right)_{\mathbf{Q_{0A}}}\int_0^\infty\alpha_{\alpha\gamma}^b(i\Omega)\alpha_{\beta\delta}^b(i\Omega)d\Omega\right] \tag{4.14}$$

4.2.1.2 Couplings between the local vibrational transitions J_{mn}

The nearest neighbor coupling constant can be calculated using the DFT technique as the finite difference in energies of the peptide configurations displaced along the coordinate of the amide I normal modes. For example, by performing ab initio molecular orbital

calculations for model glycine di- and tripeptides with various ϕ and ψ dihedral Ramachandran angles, Torii and Tasumi [116] constructed the nearest peptide groups' coupling constants for all configurations on a grid of the Ramachandran angles in the ranges 0 degree $\leq \varphi \leq$ 180 degrees and -180 degrees $\leq \psi \leq$ 180 degrees. Partial geometry optimization was performed for GLDP with each specified pair of ϕ and ψ values. The ϕ and ψ angles were taken at intervals of 30 degrees. In the geometry optimization, the structural parameters of the peptide and the methyl groups were fixed to those of the fully optimized structure of NMA. Only the structural parameters of the methylene group (CH_2), including the N−C and C−C bond lengths but excluding the ϕ and ψ angles, were optimized. Both peptide groups of GLDP (which are denoted peptide groups 1 and 2) were displaced along the coordinate of the amide I mode. The vibrational pattern was taken from the amide I normal mode of NMA. Only the positions of the C, O, N, and H atoms were displaced, since the motions of the α-carbons (the methyl groups in NMA) are negligible in the amide I mode. Four structures were generated by this procedure. The energies of these four structures were calculated at the HF/6-31(+)G** level.

In 2002 Minhaeng Cho constructed a similar map, with a smaller angular interval, at the RHF/6-311++G** level [63,64]. Performing comprehensive quantum chemical calculations (MP2 and B3LYP) for glycine dipeptide, Gerhard Stock obtained a similar potential map [13]. Thomas Jansen proposed another map of nearest-neighbor couplings at the DFT BPW91/6-31G(d,p) level [117,118]. Hayashi and Mukamel identified the intermode coupling constants, estimated the coupling constants between the adjacent polypeptide units for amide I, II, III, and A vibrational modes by performing the DFT calculations for the anharmonic vibrational Hamiltonian of glycine dipeptide [119].

The electrostatic model is commonly used for the calculation of nonneighbor couplings. Krimm first proposed the transition dipole coupling (TDC) model [28]:

$$J_{mn} = \frac{1}{4\pi\epsilon_0} \frac{(\boldsymbol{\mu}_m \cdot \boldsymbol{\mu}_n) - 3[\boldsymbol{\mu}_m \cdot \mathbf{e}_{mn}][\boldsymbol{\mu}_n \cdot \mathbf{e}_{mn}]}{r_{mn}^3} \tag{4.15}$$

where μ_m is the transition dipole in (D/Å/u$^{1/2}$) units, r_{mn} is the distance between dipoles (in Å), $\mathbf{e}_{mn}$ is the unit vector connecting m and n, and $= 1$ is the dielectric constant. The angle between the transition dipole and C = O bond usually results in a range from 10 to 20 degrees, with a most accepted angle of $\sim$15 degrees. Torii and Tasumi later on proposed a similar model [116] to calculate the amide I transition dipole μ_m and the nonnearest neighbor couplings J_{mn}.

The transition charge density couplings model was proposed by Woutersen and Hamm [120], which expands the multipoles to the fourth rank:

$$J_{mn} = \sum_{ij} \int_0^\infty \mathbf{d}r \int_0^\infty \mathbf{d}r' \frac{\partial\rho(r)}{\partial \mathbf{q}_{mi}} \frac{\partial\rho(r')}{\partial \mathbf{q}_{nj}} \mathbf{q}_{mi}\mathbf{q}_{nj} \tag{4.16}$$

This results in $\sim {r_{mn}}^{-3}$ (dipole−dipole), $\sim {r_{mn}}^{-4}$ (dipole−quadrupole), and $\sim {r_{mn}}^{-5}$ (quadrupole−quadrupole and dipole−octupole) interaction terms. An alternative treatment was proposed by Jansen et al. [121] to calculate the amide I mode coupling between two local amide I modes by considering the transition charge−charge (TCC) interactions. In 2005 Gerhard Stock [13] constructed a higher quantum chemical computational level-based multipole generalizations map with transferability. The integral equation formalism polarized continuum mode solvent model was employed for fully optimized calculations. Eight different B3LYP basis sets have been adopted from 6-31G(d) to 6-311++G(3df,2pd) level of DFT to study the base set dependency. Stock also investigated the transferability of these maps to dipeptides with hydrophilic and hydrophobic side chains as well as to tripeptides with charged end groups. Although the long-range vibrational coupling between a pair of local amide I modes is predicted successfully by the transition charge or dipole coupling models, these models cannot fully describe the through-bond coupling between neighboring polypeptides since the transition charge coupling theory is strictly based on the electrostatic interactions.

The anharmonicity Δ_m of the amide I mode is usually fixed to the measured value of $-16\ \mathrm{cm}^{-1}$ [74]. The transition dipole moments can be calculated from their definition as the dipole derivative by coordinate $\left(\partial\boldsymbol{\mu}/\partial Q\right)$. For example, in the TDC model, the optimized magnitude is 2.73 D/A and direction with 10.0 degrees as the angle to the C = O bond.

Mendelsohn and coworkers estimated the amide I local vibrational mode frequencies by subtracting an ad hoc value which was associated with the strengths of hydrogen bonding interactions (calculated by ab initio methods) [121,122]. A number of refinements have been made in the calculations of these interactions, such as a finite size for the transition dipoles, an explicit distance dependent expression for the dielectric constant due to intervening matter in the through-space electrical interaction, a geometry dependence of the through-bond interactions, a frequency shift in the acceptor hydrogen bonded groups, and a $\pi-\pi$ interaction across extended H-bonds. Six globular proteins and two peptides were calculated to benchmark this method. The excellent agreement between calculated and observed spectra suggests that the general nature of the interactions between oscillators is reasonably well determined. The mode frequencies and the coupling constants were calculated from the Hessian matrix reconstruction method, the method was based on the results of ab initio vibrational analysis, developed by Minhaeng Cho and coworkers [56,62]. The coupling constants and local mode frequencies achieved by this method have been successfully applied to simulate the IR and 2DIR spectra of polypeptides and proteins.

4.2.2 Calculation of third-order optical response functions

With the calculated fluctuating vibrational Hamiltonian and transition dipole moment trajectories, the third-order response function and the signals are then calculated. As the

most commonly used theoretical approach for simulating and analyzing the nonlinear response function, the sum-over-states (SOS) treatment describes the response functions as the combinations of correlation functions [46,57,66,99,123–126]. Based on SOS, if the frequency fluctuations of all states involved obey Gaussian statistics, the cumulant expansions of Gaussian fluctuations (CGF) technique can be used to calculate the response functions. The CGF approach describes only the diagonal fluctuations of the energy, but not the fluctuations of the off-diagonal coupling and the transition dipole moments. It also doesn't count in the non-Gaussian part of the fluctuations. Furthermore, CGF uses a fixed exciton basis and neglects the eigenstate fluctuations. Large fluctuations in protein structures might therefore not be considered. These limits can be lifted by the numerical integration of the Schrödinger equation (NISE) [54,127]. Further, for large systems, propagation in Liouville space using the stochastic Liouville equation (SLE) [128] will be more convenient when calculating the response function. In the following, we introduce these techniques in detail.

Coherent optical signals can be classified according to their power law dependence on the driving field intensities [99]. The signals are related to the polarization, $P(\boldsymbol{r},t)$, induced by the external electric fields. The induced polarization can be obtained perturbatively by expanding the density matrix $\rho(t)$ in powers of the external fields [99]:

$$P_{\nu_4}^{(3)}(\boldsymbol{r},t)=\sum_{\nu_3\nu_2\nu_1}\int_0^\infty dt_3\int_0^\infty dt_2\int_0^\infty dt_1 S_{\nu_4\nu_3\nu_2\nu_1}^{(3)}(t_3,t_2,t_1) \\ E_{\nu_3}(\boldsymbol{r},t-t_3)E_{\nu_2}(\boldsymbol{r},t-t_2-t_3)E_{\nu_1}(\boldsymbol{r},t-t_1-t_2-t_3) \tag{4.17}$$

where r and t_1, t_2, t_3 represent the coordinates and the interaction time intervals between successive interactions with the optical pulses, $E(\boldsymbol{r}, t)$. ν_j are the Cartesian components of the fields and polarizations. The response functions are system property-tensors that contain all relevant molecular information. The third-order response function $S^{(3)}$ represents the lowest order contribution to the induced polarization in isotropic systems [99]:

$$S_{\nu_4\nu_3\nu_2\nu_1}^{(3)}(\tau_4,\tau_3,\tau_2,\tau_1)=\left(\frac{i}{\hbar}\right)^3\left\langle\left[\left[\left[v_{\nu_4}(\tau_4),v_{\nu_3}(\tau_3)\right],v_{\nu_2}(\tau_2)\right],v_{\nu_1}(\tau_1)\right]\right\rangle \tag{4.18}$$

where ν_i are the Cartesian polarization indices, eight Liouville-space pathways can be derived from three nested commutators in the response functions:

$$S_{\nu_4\nu_3\nu_2\nu_1}^{(3)}(t_3,t_2,t_1)=\left(\frac{i}{\hbar}\right)^3\theta(t_1)\theta(t_2)\theta(t_3)\sum_{\alpha=1}^{8}R_\alpha(t_3,t_2,t_1) \tag{4.19}$$

Here, step function $\theta(t)$ is defined as $\theta(t>0)=1,\theta(t<0)=0$.

The dominant contributions to resonant signals only come from terms obtained when the field and molecular frequencies in Eq. (4.17) have an opposite sign. Other (same-sign) highly oscillatory terms may be safely neglected. Using this rotating wave approximation (RWA) [99], each phase-matching signal is described by a specific combination of Liouville-space pathways, depending on the pulse configuration and the detection mode. For instance, the rephasing signal S^{k_I} is created by the pulse sequence with the phase matching condition, $k_{\mathrm{I}} = -k_1 + k_2 + k_3$ in which the coherence evolves during t_1 and t_3 with the conjugate phase. This signal is contributed by three pathways [66]:

$$\begin{aligned} S^{K_I}_{\nu_4\nu_3\nu_2\nu_1}(\tau_4,\tau_3,\tau_2,\tau_1) = \left(\frac{i}{\hbar}\right)^3 \sum_{n_1 n_2 n_3 n_4} \mu_{n_1;\nu_1}(\tau_1)\mu_{n_2;\nu_2}(\tau_2)\mu_{n_3;\nu_3}(\tau_3)\mu_{n_4;\nu_4}(\tau_4) \\ \times \Big(\Big\langle g|B_{n_1}\mathcal{U}(\tau_1,\tau_2)B^{\dagger}_{n_2}\mathcal{U}(\tau_2,\tau_4)B_{n_4}\mathcal{U}(\tau_4,\tau_3)B^{\dagger}_{n_3}|g\Big\rangle \\ + \Big\langle g|B_{n_1}\mathcal{U}(\tau_1,\tau_3)B^{\dagger}_{n_3}\mathcal{U}(\tau_3,\tau_4)B_{n_4}\mathcal{U}(\tau_4,\tau_2)B^{\dagger}_{n_2}|g\Big\rangle \\ - \Big\langle g|B_{n_1}\mathcal{U}(\tau_1,\tau_4)B_{n_4}\mathcal{U}(\tau_4,\tau_3)B^{\dagger}_{n_3}\mathcal{U}(\tau_3,\tau_2)B^{\dagger}_{n_2}|g\Big\rangle \Big) \end{aligned} \tag{4.20}$$

The time evolution operator of the vibrations denoted by $\mathcal{U}(\tau_2,\tau_1)$ is:

$$\mathcal{U}(\tau_2,\tau_1) = \exp_+\left(-\frac{i}{\hbar}\int_{\tau_1}^{\tau_2} H(\tau)d\tau\right) \tag{4.21}$$

where $\exp_+$ represents the time-ordered exponential operator. The three terms in Eq. (4.20) correspond to the contributions of the ground state bleaching (GSB), excited state absorption (ESA), and excited state emission (ESE), respectively.

$$S^{K_I}_{\mathrm{GSB}}(t_3,t_2,t_1) = -\left(\frac{\mathrm{i}}{\hbar}\right)^3 \Big\langle \mu_{\mathrm{ge}}(\tau_1)\mathcal{U}_{ee}(\tau_1,\tau_2)\mu_{eg}(\tau_2)\mu_{ge}(\tau_4)\mathcal{U}_{ee}(\tau_4,\tau_3)\mu_{eg}(\tau_3)\Big\rangle \Gamma_{\mathrm{GSB}}(t_3,t_2,t_1) \tag{4.22}$$

$$S^{K_I}_{\mathrm{ESE}}(t_3,t_2,t_1) = -\left(\frac{\mathrm{i}}{\hbar}\right)^3 \Big\langle \mu_{ge}(\tau_1)\mathcal{U}_{ee}(\tau_1,\tau_3)\mu_{eg}(\tau_3)\mu_{ge}(\tau_4)\mathcal{U}_{ee}(\tau_4,\tau_2)\mu_{eg}(\tau_2)\Big\rangle \Gamma_{\mathrm{ESE}}(t_3,t_2,t_1) \tag{4.23}$$

$$\begin{aligned} S^{K_I}_{\mathrm{ESA}}(t_3,t_2,t_1) = -\left(\frac{\mathrm{i}}{\hbar}\right)^3 \\ \Big\langle \mu_{ge}(\tau_1)\mathcal{U}_{ee}(\tau_1,\tau_4)\mu_{ef}(\tau_4)\mathcal{U}_{\mathrm{ff}}(\tau_4,\tau_3)\mu_{fe}(\tau_3)\mathcal{U}_{ee}(\tau_3,\tau_2)\mu_{eg}(\tau_2)\Big\rangle \Gamma_{\mathrm{ESA}}(t_3,t_2,t_1) \end{aligned} \tag{4.24}$$

Here, τ_1,τ_2,τ_3, and τ_4 are the moments when the system interacts with the external fields. The vibrational lifetimes accounted for by the relaxation factor are given as:

$$\Gamma_{\mathrm{GSB}}(t_3,t_2,t_1)=\exp\Big(-(t_3+2t_2+t_1)/2T_1\Big) \tag{4.25}$$

$$\Gamma_{\mathrm{ESE}}(t_3,t_2,t_1)=\exp\Big(-(t_3+2t_2+t_1)/2T_1\Big) \tag{4.26}$$

$$\Gamma_{\mathrm{ESA}}(t_3,t_2,t_1)=\exp\left(-\frac{(t_3+2t_2+t_1)}{2T_1}-\frac{t_3}{2T_2}\right) \tag{4.27}$$

T_1 and T_2 are the lifetimes of the singly and doubly excited states, respectively.

Fig. 4.3 shows the three contributions from distinct Feynman diagrams for the k_I technique in the state of the density matrix during each time interval. The amide band energy level scheme consists of three well-separated bands as shown, g is the ground state, e is the first excited state manifold, and f is the second excited state manifold. The transitions that can be induced by the pulses are μ_{ge} and μ_{ef}. Only transitions between the ground state, g, and the first excited states manifold, e, and between the first and second excited state manifold, f, are allowed. The response functions may be calculated by summing over all possible transitions among vibrational eigenstates (therefore "sum-over-states"). The nonlinear response vanishes for harmonic vibrations and may thus be attributed to the anharmonicities.

4.2.3 Cumulant expansion of Gaussian fluctuation of third-order response functions

The response function for a special class of Gaussian fluctuation models may be solved exactly with a compact closed-form of expressions. Such a method has been successfully applied for small peptides with fewer than 30 residues. The systems should have only

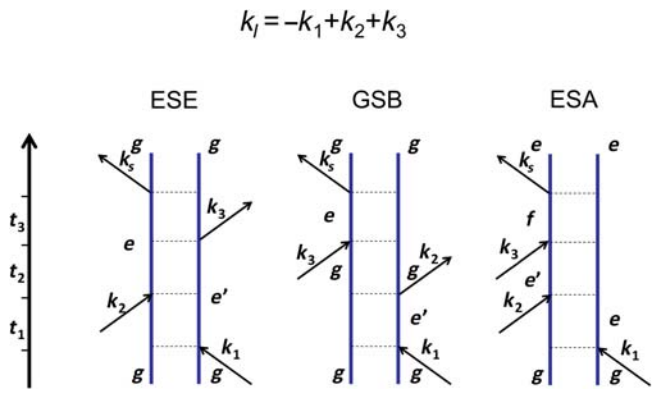


Figure 4.3
Double-sided Feynman diagrams representing the Liouville-space pathways contributing to the k_I signal in the rotating wave approximation (RWA). The three pathways are known as excited-state emission (ESE), the ground-state bleaching (GSB), and excited-state absorption (ESA).

diagonal energy fluctuations with Gaussian statistics. The fluctuations should be small with respect to energy-level spacings so that curve-crossing is avoided. In this case the energies are modulated by collective coordinates expressed as sums of harmonic coordinates.

Due to the central limit theorem, on the other hand, the CGF method should be applicable for a broader range of systems, where the collective coordinates can be expressed as the sums of many bath coordinates with small contributions. When the collective coordinate is the electric field at a certain site, for instance, given by the sum of contributions from all charges in the environment, Condon approximation suggests that we can neglect fluctuations with the magnitude as large as the transition dipole.

The response functions can be calculated exactly using the second-order cumulant expansion. We can introduce $\Delta\omega_{ma}(t) \equiv \omega ma(t) - \overline{\omega} ma$ to represent the fluctuations of the transition frequencies, where $\omega_{\overline{\alpha\beta}}$ is the average transition frequency. The two-time correlation function of $\Delta\omega$ is [99]:

$$C_{mn}(\tau_1, \tau_2) = \frac{1}{\hbar^2}\left\langle \Delta\omega_{ma}(\tau_1)\Delta\omega_{na}(\tau_2)\right\rangle \equiv C^{'}(t_1) + iC^{''}(t_1) \tag{4.28}$$

where $C'(t)$ and $C''(t)$ are the real and imaginary parts of C and $t_1 = \tau_1 - \tau_2$. We further define the line-broadening functions:

$$g_{mn}(t) = \int_0^t d\tau_2 \int_0^{\tau_1} d\tau_2 C_{mn}(\tau_2) \tag{4.29}$$

using the fluctuation dissipation relation between C' and C''. $g_{mn}(t)$ can be expressed as

$$g_{mn}(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi}\frac{1-\cos(\omega t)}{\omega^2}\coth\left(\frac{\hbar\omega}{2k_BT}\right)C_{mn}^{''}(\omega) - i\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\frac{\sin(\omega t)-\omega t}{\omega^2}C_{mn}^{''}(\omega) \tag{4.30}$$

where

$$C_{mn}^{''}(\omega) = 2\int_0^{\infty} dt\ \ \sin(\omega t)C_{mn}^{''}(t) \tag{4.31}$$

is the *spectral density*. The real and the imaginary parts of $g_{nm}(t)$ are responsible for line-broadening and shift, respectively. The third-order nonlinear response functions can be expressed in terms of $g(t)$.

The simulated k_I signals of NMA amide I band (Fig. 4.4A) in water is presented in Fig. 4.4B [129]. The finite vibrational lifetime is significant in this case. For systems with dynamics that are separable into multiple time regimes, we can adopt the inhomogeneous CGF protocol, and calculate the spectrum by summing over contributions of the slowly interconverting configurations, each represented by the CGF. We can illustrate this protocol for the 2DIR spectra of the trpzip2 peptide, which has a β-hairpin conformation. Trpzips are the smallest peptides to adopt a unique tertiary fold without requiring metal binding,

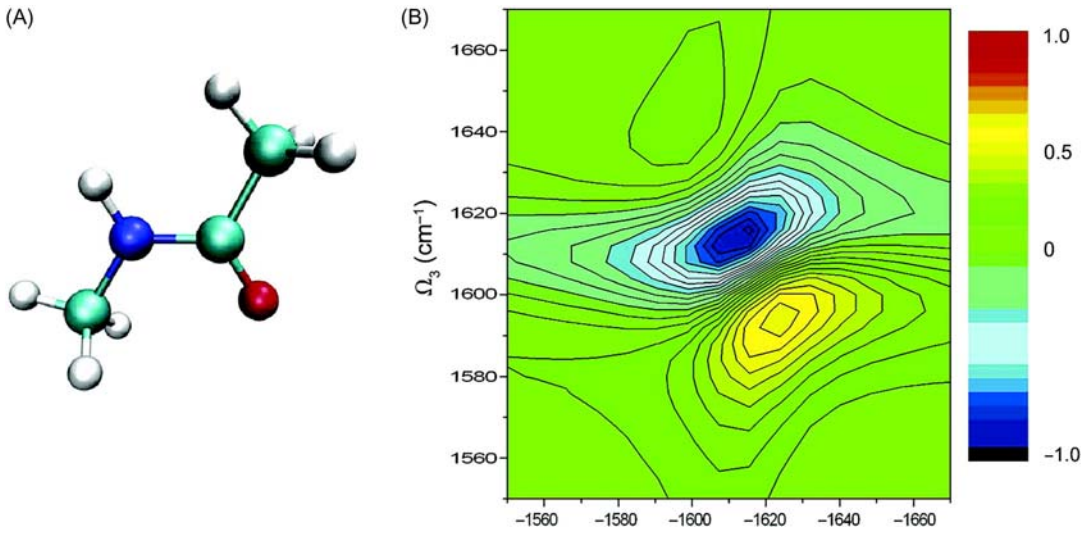


Figure 4.4

(A) *N*-methyl acetamide (NMA) molecule structure. (B) Cumulant expansion of Gaussian fluctuation (CGF) simulated k_I imaginary spectrum of NMA with vibrational relaxation. Source: *Reprinted from W. Zhuang, D. Abramavicius, T. Hayashi, S. Mukamel, Simulation protocols for coherent femtosecond vibrational spectra of peptides, J. Phys. Chem. B 110 (7) (2006) 3362−3374 [129].*

unusual amino acids, or disulfide cross-links. The unlabeled samples, second residue (Trp2), and the seventh residue (Gly7) isotopic labeled samples are denoted as UL, L2, and L7, respectively. The top row in Fig. 4.5 shows the experimental 2DIR spectra of the trpzip2 ^{13}C isotopomers. In panel (A), the diagonal signals are due to 0−1 (red) and 1−2 transitions (blue). The two fundamental 0−1 frequencies agree with the experimental absorption. The cross peaks are induced by pairwise vibrational couplings among local amide I modes. The diagonal and off-diagonal peaks change upon ^{13}C-labeling as shown in panels (B) and (C). The spectra simulated using the Ham and Cho model [56] are shown in the middle row and those made using the Hayashi and Mukamel model [53] are in the bottom row. The main 2DIR characteristics of the UL, L2, and L7 are reasonably well reproduced by both models.

4.2.4 The numerical integration of the Schrödinger equation

The response functions can also be calculated by NISE [130]. This approach explicitly takes the nonadiabatic effects into consideration. For the *n*-exciton wave function, the time-dependent Schrödinger equation is expressed as:

$$i\hbar\frac{\partial}{\partial t}|\psi_{\nu_1}^{(n)}(t;\tau_1)\rangle = H(t)|\psi_{\nu_1}^{(n)}(t;\tau_1)\rangle \tag{4.32}$$

where $|\psi_{\nu_1}^{(n)}(t;t_0)\rangle = \sum_m |\psi_{m;\nu_1}^{(n)}(t;t_0)B_m^\dagger|\mathrm{g}\rangle$.

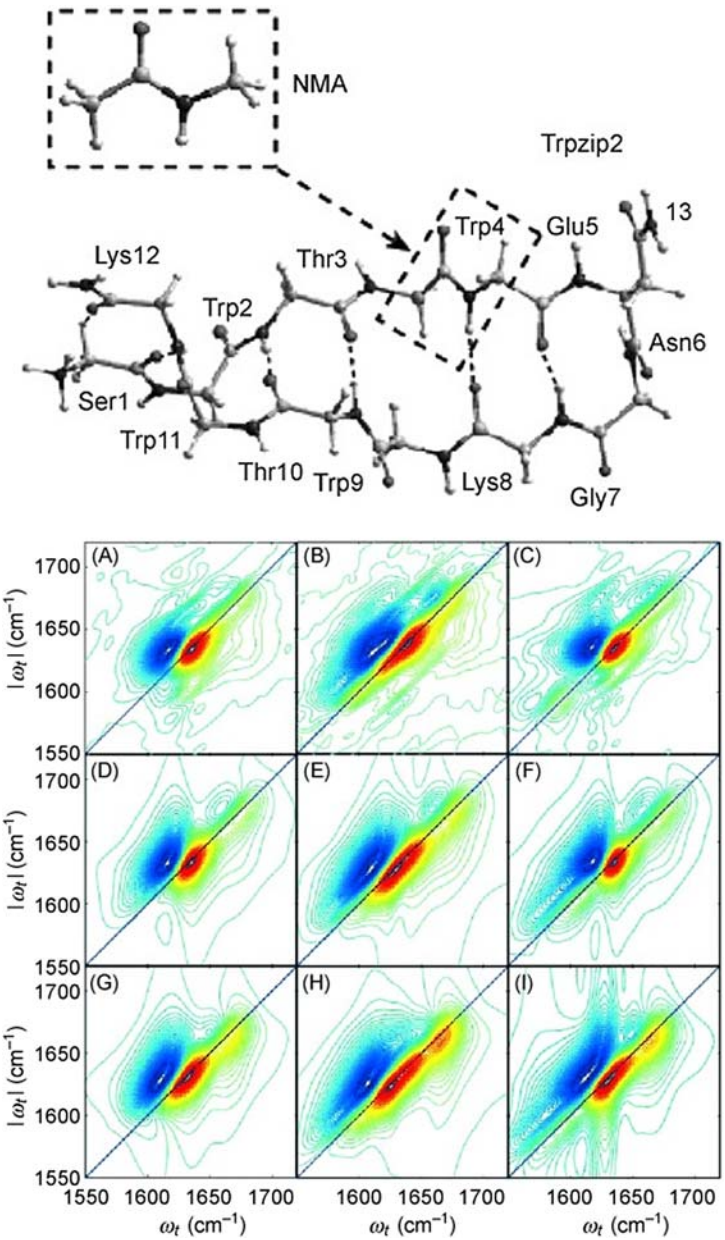


Figure 4.5

(A–C) Experimental $k_I + k_{II}$ spectra of trpzip2 ^{13}C isotopomers. (D–F) simulations using the Ham and Cho Hamiltonian [56]; (G–I) simulations results using the Hayashi and Mukamel Hamiltonian [53]. The three columns from left to right are isotopic species denoted as UL, L2, and L7. ω_τ is Ω_1 in the notation, while ω_τ is Ω_3. Source: *Reprinted from W. Zhuang, T. Hayashi, S. Mukamel, Coherent multidimensional vibrational spectroscopy of biomolecules: concepts, simulations, and challenges, Angew. Chem. Int. Ed. 48 (21) (2009), 3750–3781 [66].*

This wave function can be written in terms of the original fixed site basis functions $|\varphi>$ for the considered excitation manifold:

$$|\psi_i^{(n)}(t;\tau_2;\tau_1)\rangle = \sum_j c_{ij}(t;\tau_2;\tau_1)|\varphi_j^{(n)}\rangle \tag{4.33}$$

For $t = 0$, the expansion coefficients fulfill $c_{ji}(0) = \delta_{ji}$. Upon insertion of Eq. (4.33) in the time-dependent Schrödinger equation, an equation for the expansion coefficients is obtained:

$$i\hbar\frac{\partial}{\partial t}c_{ij}(t) = \sum_k H_{ik}(t)c_{jk}(t) \tag{4.34}$$

Changing the partial ∂t to Δt, the numerical integration over one such time step of duration Δt yields:

$$c_{ij}((n+1)\Delta t) = \exp\left(-\frac{i}{\hbar}\sum_k H_{ik}(n\Delta t)\Delta t\right)c_{jk}(n\Delta t) \equiv \sum_k \mathcal{U}_{ik}((n+1)\Delta t, n\Delta t)c_{jk}(n\Delta t) \tag{4.35}$$

Here, the time evolution operators $\mathcal{U}((n+1)\Delta t, n\Delta t)$ propagate the time-dependent coefficients from the time $n\Delta t$ to $(n+1)\ \Delta t$. For numerical integration of this equation using small time increments, the Hamiltonian can be treated as a constant during each integration step. The propagator in Eq. (4.21) can then be written as the product of ordinary exponential factors:

$$U(n\Delta t, 0) = \exp_+\left(-\frac{i}{\hbar}\int_0^{n\Delta t} \boldsymbol{H}(\tau)d\tau\right) \tag{4.36}$$

The Hamiltonian $H_{ik}(n\Delta t)$ usually changes with times, so the numerical integration to a converged response function may be very time-consuming. There is an approach called time-averaging approximation (TAA) [131,132] to overcome such problems. With a frequency averaged over some suitably chosen fixed time T, one can arrive at an approximate expression for the line shape that involves replacing the above frequency, averaged over time t. The time evolution operator of the vibrations $\mathcal{U}(\tau_2, \tau_1)$ in Eq. (4.21) becomes:

$$\mathcal{U}(t,0) = \left\{\exp_{\mathrm{T}}\left(-\frac{it}{\hbar}\frac{1}{t}\int_0^t H(\tau)d\tau\right)\right\} \approx \left\{\exp_{\mathrm{T}}\left(-\frac{it}{\hbar}\frac{1}{t}\int_0^t \omega(t')dt'\right)\right\} \tag{4.37}$$

$$\omega_T = \frac{1}{T}\int_0^T \omega(t')dt' \tag{4.38}$$

Then the propagation only involves the average frequency ω_T. This expression is suggested by Auer and Skinner [131] using the average approximating over a fixed time interval T. TAA accelerates the numerical procedure in a very efficient way accounting for dynamic effects in linear absorption spectra. Jansen and Ruszel [132] used a slightly more general expression which introduces a weight function such as the exponential function, a box shaped function or a Gaussian function.

$$\omega_{\tau A}^{\mathrm{EXP}} = \frac{1}{\tau_A}\int_{\mathrm{t}}^{\infty} \exp(-(t^{'}-t)/\tau_A)\omega(t^{'})dt^{'} \tag{4.39}$$

$$\omega_{\tau A}^{\mathrm{BOX}} = \frac{1}{\tau_A}\int_{\mathrm{t}}^{t+\tau_A} \omega(t^{'})dt^{'} \tag{4.40}$$

$$\omega_{\tau A}^{\mathrm{GAUSS}} = \frac{2}{\sqrt{2\pi}\tau_A}\int_{\mathrm{t}}^{\infty} \exp\left[-\frac{1}{2}\left(\frac{t^{'}-t}{\tau_A}\right)\right]\omega(t^{'})dt^{'} \tag{4.41}$$

Jansen and Ruszel [132] found that the Gaussian weight function gave the best results for commonly occurring overdamped Brownian model cases over a broad range of fluctuation timescales. For the Kubo model of an isolated chromophore this TAA works, for an appropriately chosen averaging time, as long as one is not in the extreme-narrowing limit. The power of this approach is especially evident for complicated problems with many interacting chromophores.

Fig. 4.6 shows the TAA spectrum in the perpendicular polarization configuration compared with the NISE spectrum. The left spectrum is calculated with the time-average procedure using a 167 fs averaging time with a Gaussian weight function. This corresponds to an

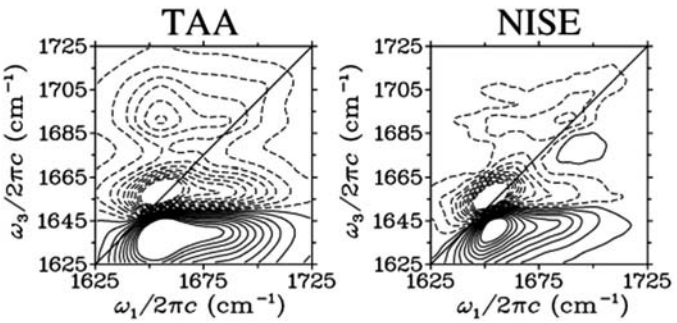


Figure 4.6

The two-dimensional infrared (2DIR) spectrum of the amide I region for β-hairpin trpzip2 calculated in the time-averaging approximation (TAA) and using the numerical integration of the Schrödinger equation (NISE) [132] by Jansen and Ruszel. Source*: Reprinted from T.L.C. Jansen, W. M. Ruszel, Motional narrowing in the time-averaging approximation for simulating two-dimensional nonlinear infrared spectra, J. Chem. Phys. 128 (21) (2008), 214501.*

inhomogeneous line width of 20 cm^{-1}. The TAA reproduces the basic structure of the spectrum. However, the peaks are generally too broad and the cross peak at (ω_1,ω_1) = (1650, 1690 cm^{-1}) is too intense in the TAA calculation. These differences can either be due to the neglect of nonadiabatic effects during the time delays t_1 and t_3 or they might arise because all elements of the Hamiltonian were averaged with the same averaging time. The latter might be a special problem for the fluctuating coupling. It might be possible to improve the agreement by optimizing the averaging procedure for problems where the coupling is fluctuating. The NISE calculation took 100 hours, while the TAA calculation took less than a minute on the same computer used for the trpzip2 simulations.

4.2.5 The stochastic Liouville equations

The time evolution of the density matrix describing the state of the two-mode system is described by the Liouville equation [99,133]

$$\frac{\partial}{\partial t}\rho(t) = -\frac{i}{\hbar}L(t)\rho(t) - \frac{i}{\hbar}L_{int}(t)\rho(t) \tag{4.42}$$

Here

$$L(t)\rho(t) = -\frac{i}{\hbar}[H_0(t), \rho(t)]$$

represents the isolated system, while

$$L_{int}(t)\rho(t) = -\frac{i}{\hbar}[H_{int}(t), \rho(t)].$$

represents the coupling with the radiation field.

The probability distributions $P(Q,t)$ of stochastic variables are modeled by the Markovian master equation:

$$\frac{\partial P(Q,t)}{\partial t} = -\Gamma(Q)P(Q,t) \tag{4.43}$$

where $\Gamma(Q)$ has the Smoluchowski overdamped Brownian Oscillator form

$$\frac{\partial P(Q,t)}{\partial t} = -\sum_{j=1}^{4}\gamma_j\frac{\partial}{\partial Q_j}\left(Q_j + \Delta_j^2\frac{\partial}{\partial Q_j}\right)P(Q,t) \tag{4.44}$$

The SLE is constructed by combining the Liouville equation for the exciton system and the Markovian master equation for the four collective Brownian oscillator coordinates.

$$\dot{\rho}(Q,t) = -\frac{i}{\hbar}L(Q)\rho(Q,t) - \Gamma(Q)\rho(Q,t) \tag{4.45}$$

SLE may be solved using a matrix continued-fraction representation [134] of the Green's functions, in the frequency domain. The 2DIR rephasing signal $S^{kI}(\omega_3, \tau_2, \omega_1)$, was computed by transforming the frequency ω_2 back to the time domain.

$$S^{kI}(\omega_3, \tau_2, \omega_1) = Im \int_{-\infty}^{\infty} d\omega_2 S^{kI}(\omega_3, \omega_2, \omega_1) e^{-i\omega_2\tau_2} \tag{4.46}$$

For example, trialanine has two amide units contributing to its amide I band. The Hamiltonian includes the frequencies ω_a and ω_b, anharmonicities K_a and K_b, and the coupling constant J of the two local modes. The simulation therefore includes six vibrational energy levels: the ground state (g), two single excited levels (e_1 and e_2), and three doubly excited levels (f_1, f_2, and f_3). The frequency fluctuations of the two modes ($\delta\omega_a$) and ($\delta\omega_b$) are treated as independent stochastic variables dominated by the interaction with the solvent water molecules. The Brownian oscillator parameters include the relaxation times ${\gamma_a}^{-1} = {\gamma_b}^{-1} = 220$ fs and the magnitudes $\Delta_a = \Delta_b = 16.1\ \text{cm}^{-1}$, which reproduce the experimental line shape for the isolated amide I mode in NMA. Transition dipole fluctuations of the local modes were neglected as well and their magnitude were set to unity. The stochastic variables of trialanine case are $Q_1 = \delta\omega_a$, $Q_2 = \delta\omega_b$, $Q_3 = \delta\varphi$, and $Q_4 = \delta\Psi$.

The Green's function for the t_2 interval may also be computed in the time domain by a direct time integration of the SLE. Different levels of simulation of the parallel ZZZZ signal were compared [127]. The highest level includes fluctuations of all four collective bath coordinates. The Liouville operator is constructed in the local basis and the coupling between the two local modes fluctuates with the Ramachandran angles. The local mode frequencies fluctuate as well. Satisfactory agreement with experimentation is obtained, as shown in Fig. 4.7. In summary, four collective coordinates can account for the effect of fluctuations on the two amide I modes for trialanine. Ramachandran angle fluctuations have significant signatures on 2DIR line shapes in nonrigid peptides. The exchange between conformers in trialanine is slow and the signal is given by a sum of the contributions of the various conformers. Fast-exchange shows interesting signatures in 2D signals, as demonstrated in hydrogen-bonding and isomerization dynamics. These can be described by including a multistate jump model in the SLE.

The SLE can be used to describe many types of fluctuations of all elements of the Hamiltonian. The only requirement is that they may be represented by a few collective (discrete or continuous) coordinates which satisfy a Markovian equation of motion. These equations account for the effect of the fluctuations of collective bath coordinates on the nonlinear infrared spectra by describing the evolution in the joint system + bath space.

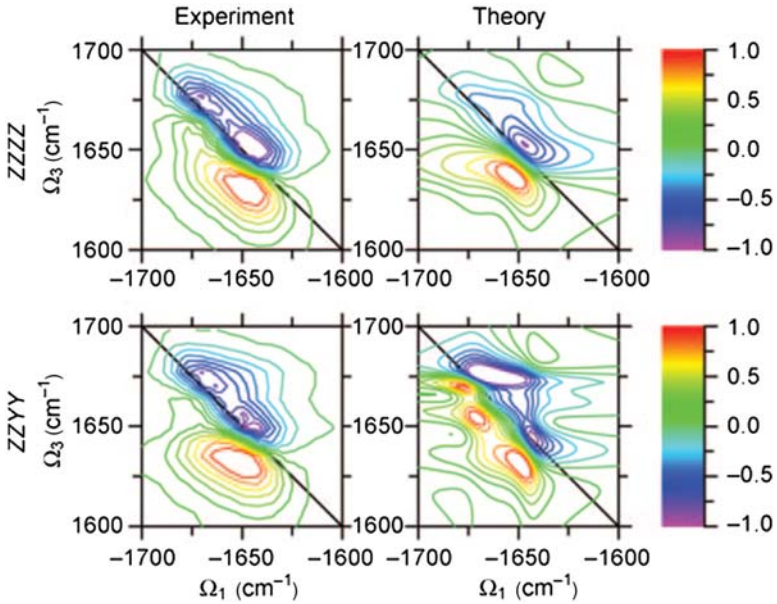


Figure 4.7

Top: The experimental k_I photon echo spectrum of trialanine [127] (*left*) and the simulated spectrum (*right*) for parallel polarized pulses. Bottom: Same comparison but for perpendicular polarized pulses. The spectra are normalized to the most intense peak. Source*: Reprinted from W. Zhuang, T. Hayashi, S. Mukamel, Coherent multidimensional vibrational spectroscopy of biomolecules: concepts, simulations, and challenges, Angew. Chem. Int. Ed. 48 (21) (2009), 3750–3781 [66].*

4.2.6 Applications of the statistical mechanic methods for longer dynamics or more comprehensive configuration ensembles

4.2.6.1 Simulating the peptide thermal unfolding 2DIR spectra using the integrated tempering sampling technique

When simulating the 2DIR spectra of a soft peptide molecules, a sufficient sampling of the peptide conformational ensemble becomes critical for rendering the accurate result. Sampling the peptide configuration distribution is still an extremely demanding task for MD simulation. Therefore enhanced sampling techniques have been developed to significantly improve the efficiency of configuration sampling over a large energy range in MD simulations. The thermal unfolding FTIR and 2DIR spectra of the polypeptides trpzip2 and trpzip4 have been modeled based on the configuration distributions generated by MD simulations using the implicit solvent model combined with the integrated tempering sampling (ITS) method [135]. The ITS method, developed by Gao [136], applies the Boltzmann distribution functions at multiple temperatures to sample the configurations of the system over a wide energy range. Details of this enhanced sampling technique have

been extensively discussed and reviewed elsewhere [137–140]. The melting temperature achieved in the site-specific changes of the calculated intensities are also very consistent with experimental results [141]. The characteristics of isotope-labeled 2DIR spectra signal, such as the intensity and ellipticity of the isotope-shifted peaks, are proposed to be a clearer and more sensitive tool to observe melting characters.

The differences of the folding landscape for trpzip2 and trpzip4 peptides are reflected in the relative thermal stabilities between the turn and the mid-strand segments. For trpzip2 the turn is easier to fold and the local structure around residue GLU-5 (isotope labeled in Turn-6) is thermally more stable compared with that around residues THR-3 (isotope labeled in Mid-3 and Mid-D) and THR-10 (isotope labeled in Mid-D). The ellipticity of the isotope-labeled peak in Turn-6 for trpzip2 and in the mid-strand region (Mid-3 and Mid-D) for trpzip4, as presented in Fig. 4.8B, shows much higher temperature-stability than other isotopologues. On the other hand, the frequency distribution full width at half maximum of M3 and M6 modes are rather invariant with temperature, which indicates that the inhomogeneous broadenings of isotope-labeled peaks are also temperature independent. Therefore the ellipticity is strongly correlated to variations in homogeneous broadenings, which are largely decided by the relaxation times of the frequency time correlation functions.

4.2.6.2 Simulating the temperature jump peptide two-dimensional infrared using the Markov state models

Temperature jump (T-jump)-triggered peptide unfolding processes usually take microseconds or even longer. Modeling of the related FTIR and 2DIR spectra therefore becomes difficult due to the demanding task of generating the lengthy unfolding trajectory ensemble using straightforward MD simulations. In the recent works, the T-jump spectra were simulated based on the peptide metastable states constructed by Markov state models (MSMs) [75,143]. The results have a much better convergence and agreement with the experiments than those of straightforward MD simulations. The unfolding-related spectroscopic observables are therefore largely determined by the efficient sampling of important relevant conformational states.

MSMs, in essence a time coarse graining method, provide a powerful means to handle timescale challenges. The configuration phase space is partitioned into a number of metastable states, and the interstate transition is designed to be much slower than the intrastate transition, so that the kinetics can be considered as Markovian and described as a memoryless master equation [144]:

$$P(n\Delta t) = [T(\Delta t)]^n P(0) \tag{4.47}$$

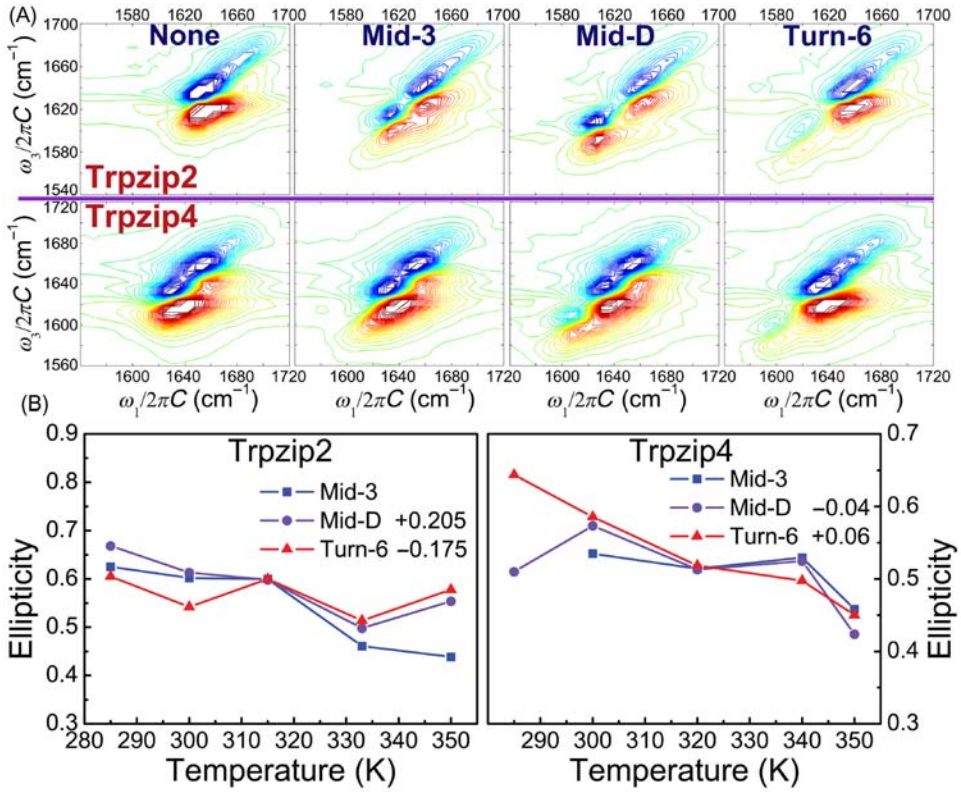


Figure 4.8

(A) Two-dimensional infrared (2DIR) spectra of each isotopologue for both trpzip2 (*above*) and trpzip4 (*below*) are taken at 285 K. (B) Temperature-dependent ellipticities of the isotope-shifted peaks in 2DIR spectra of trpzip2 (*left*) and trpzip4 (*right*) for different isotopologues. The curves of Mid-D and Turn-6 are shifted for better comparison. Source*: Reprinted from T.M. Wu, R.T. Zhang, H.H. Li, L.J. Yang, W. Zhuang, Discriminating trpzip2 and trpzip4 peptides' folding landscape using the two-dimensional infrared spectroscopy: a simulation study, J. Chem. Phys. 140 (5) (2014) 055101 [142].*

where $P(n\Delta t)$ is the state population at time $n\Delta t$, and T is the column stochastic transition probability matrix. Δt is the lag time representing the time intervals for transitions. T is calculated by normalizing the number of transitions between each pair of states after a lag time in the simulation database.

To construct MSMs to simulate the T-jump 2DIR spectra of the small trpzip2 hairpin peptide, two sets of conformations extracted from simulation data are grouped into 2000 microstates at 300 and 350K using a K-centers clustering algorithm. Then a superlevel-set hierarchical clustering (SHC) algorithm [145] was used to lump microstates together to construct macrostate MSMs. The key insight of the SHC algorithm is to generate a set of superlevels covering different density regions of phase space, then cluster each superlevel

separately, and finally recombine this information into a single MSM. Seventeen-macrostate and 13-macrostate MSMs at 300 and 350K are obtained, respectively. The relaxation dynamics after T-jump is simulated by calculating the evolution of metastable states in Eq. (4.47). The 2DIR spectra signals of each metastable state have been calculated from the structure ensemble obtained from MSM, and then the weighted sum of these signals of different states according to their instantaneous populations in order to achieve an overall signal.

The distribution of equilibrium populations for the 17-macrostates MSM is shown in Fig. 4.9A. The representative configuration of each macrostate at 300K, together with its population, is shown in Fig. 4.9B. The simulated linear absorption spectrum (dashed black curve) at 300K and the contributions from every macrostate with a population larger than 1% (solid colored curves) are plotted in Fig. 4.9C. The simulated linear absorption in Fig. 4.9C shows a stronger transition peak at $\sim 1640\ \mathrm{cm}^{-1}$ and a weaker peak at $\sim 1670\ \mathrm{cm}^{-1}$. This pattern agrees nicely with the experiments, and resembles the typical feature of a folded beta-hairpin peptide. The simulated FTIR signal of a typical unfolded configuration (macrostate 8), however, gives only a single peak at $\sim 1675\ \mathrm{cm}^{-1}$. The simulated 2DIR spectra using MSMs (Fig. 4.9D) present an asymmetric feature with the antidiagonal line width broadening at the red end, which agrees with the experiments [141]. The NMR structure-based result, used as the comparison, has symmetric peaks that are elongated along the parallel direction. The straightforward MD's unsuccessful reproduction of the diagonal elongation is due to the underestimation of the structural inhomogeneity, which is known to be the cause of the diagonal elongation (inhomogeneous broadening) in the 2D spectra.

4.3 Future perspective

In recent years 2DIR techniques have coalesced into a family of powerful spectroscopic tools for investigating the ultrafast peptide conformational dynamics. Although having relatively lower structural resolution with respect to the X-ray or NMR techniques, 2DIR's excellent time resolution makes it a unique tool for detecting subnanosecond events. Furthermore, a wider frequency window for 2DIR signals ranging from 900 to $4000\ \mathrm{cm}^{-1}$ have been achieved during the last decade, which significantly improves the structural resolution of this technique.

The simulated 2DIR signals using enhanced sampling methods, such as MSM, which provide much better statistical descriptions of the population in different parts of the phase space can provide much better agreement with the equilibrium experimental spectra compared with those generated from straightforward MD simulations. Based on MSMs, the T-jump-triggered long-time unfolding-related IR, 2DIR, and vibrationally resolved fluorescence of proteins can be simulated. Sufficient conformational sampling is crucial for

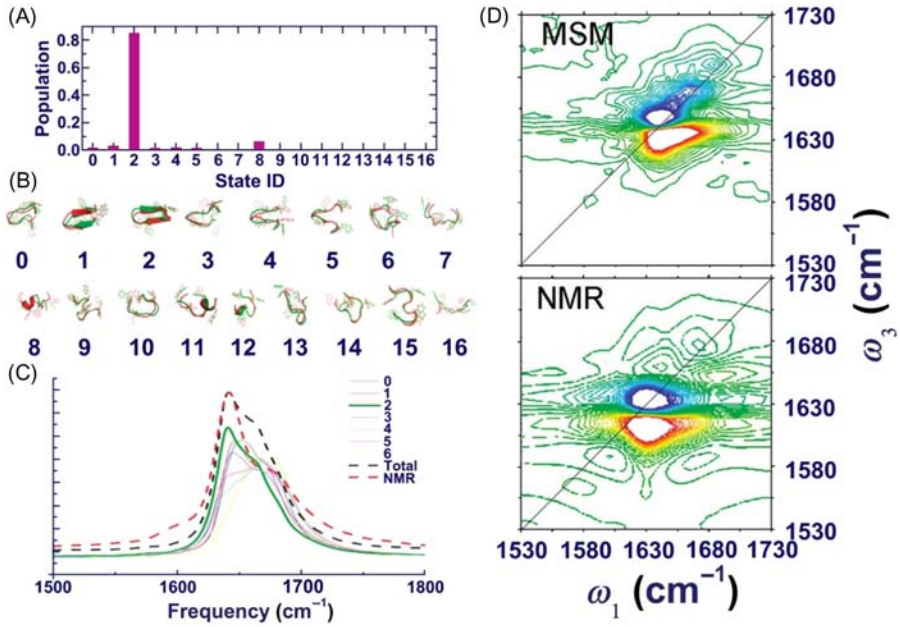


Figure 4.9

(A) Equilibrium populations for 17-macrostates and (B) the corresponding respective structures of each macrostates achieved by using Markov state models (MSMs) for the trpzip2 polypeptide at 300K. (C) The overall 300K equilibrium Fourier-transform infrared (FTIR) signal (*black dash*) and the corresponding contributions (*solid lines*) from macrostates with population >1%, simulated using MSMs data. For a better presentation, the overall signal is scaled by 1.2. The green solid line represents the contribution from the most significant macrostate (state 2), and the red dashed line gives the simulated linear absorption signal obtained from the NMR structure, and it is scaled to have the same value compared with the MSMs overall signals for a better comparison. (D) The simulated absorptive two-dimensional infrared (2DIR) spectra for trpzip2 at 300K, computed using MSM (*upper*). The same spectra calculated from the nanosecond molecular dynamics (MD) simulations on the basis of the NMR structure (*lower*). The blue 2D peak represents the vibrational transition from the ground state to the first excited states (0−1), and the red one corresponds to the transition from the first excited state to the second excited state (1−2). Source*: Reprint from W. Zhuang, R.Z. Cui, D.A. Silva, X.H. Huang, Simulating the T-jump-triggered unfolding dynamics of trpzip2 peptide and its time-resolved IR and two-dimensional IR signals using the Markov state model approach, J. Phys. Chem. B 115 (18) (2011), 5415−5424 [75].*

obtaining accurate spectroscopic observables. MSMs provide a good way to simulate the time-resolved spectroscopy from the relaxation of the metastable state populations. Evidently, sufficient accuracy of the spectroscopic signal simulation for peptides and proteins can push theoretical work to become more capable of shedding light for the interpretation of protein folding/unfolding problems. Correspondingly, this will also help to refine the theoretical model. The complexity of the atomistic processes contributing to the IR signal makes it still rather difficult to interpret IR absorption patterns in terms of local

structural organization and atomic motions. Due to computational demand, the limited conformational number certainly trims fluorescence spectra calculation accuracy. Obtaining spectroscopic signals of each state still requires an ensemble averaged over many protein conformations. There is every reason to believe that developing a high efficiency protocol that combines quantum mechanism with sampling algorithm is still the future trend for theorists.

Acknowledgments

This material is based upon work supported by the National Key Research and Development Program of China (2017YFA0206801), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB20000000 and XDB10040304), and National Natural Science Foundation of China Grants (21433014, 21733007, 21733007, 21803071), and the National 1000 Youth Talents Program.

References

[1] J. Venkatraman, S.C. Shankaramma, P. Balaram, Design of folded peptides, Chem. Rev. 101 (10) (2001) 3131–3152.
[2] D.J. Hill, et al., A field guide to foldamers, Chem. Rev. 101 (12) (2001) 3893–4011.
[3] Z. Shi, et al., Conformation of the backbone in unfolded proteins, Chem. Rev. 106 (5) (2006) 1877–1897.
[4] H. Susi, Sn Timashef, L. Stevens, Infrared spectra and protein conformations in aqueous solutions, J. Biol. Chem. 242 (23) (1967) 5467–5473.
[5] Y. Kusumoto, et al., Temperature dependence of amyloid beta-protein fibrillization, Proc. Natl Acad. Sci. U.S.A. 95 (21) (1998) 12277–12282.
[6] Y. Zhu, et al., Ultrafast folding of alpha(3): a de novo designed three-helix bundle protein, Proc. Natl Acad. Sci. U.S.A. 100 (26) (2003) 15486–15491.
[7] C. Cecconi, et al., Direct observation of the three-state folding of a single protein molecule, Science 309 (5743) (2005) 2057–2060.
[8] K. Inomata, et al., High-resolution multi-dimensional NMR spectroscopy of proteins in human cells, Nature 458 (7234) (2009) 106–U11.
[9] H.S. Chung, et al., Single-molecule fluorescence experiments determine protein folding transition path times, Science 335 (6071) (2012) 981–984.
[10] K. Neupane, et al., Direct observation of transition paths during the folding of proteins and nucleic acids, Science 352 (6282) (2016) 239–242.
[11] R.H. Zhou, B.J. Berne, R. Germain, The free energy landscape for beta hairpin folding in explicit water, Proc. Natl Acad. Sci. U.S.A. 98 (26) (2001) 14931–14936.
[12] C.Y. Huang, et al., Helix formation via conformation diffusion search, Proc. Natl Acad. Sci. U.S.A. 99 (5) (2002) 2788–2793.
[13] R.D. Gorbunov, D.S. Kosov, G. Stock, Ab initio-based exciton model of amide I vibrations in peptides: definition, conformational dependence, and transferability, J. Chem. Phys. 122 (22) (2005) 114507.
[14] A.W. Smith, A. Tokmakoff, Probing local structural events in beta-hairpin unfolding with transient nonlinear infrared spectroscopy, Angew. Chem. Int. Ed. 46 (42) (2007) 7984–7987.
[15] D.S. Cerutti, et al., Simulations of a protein crystal with a high resolution X-ray structure: evaluation of force fields and water models, J. Phys. Chem. B 114 (40) (2010) 12811–12824.
[16] K. Maenaka, et al., Dissection of protein-carbohydrate interactions in mutant hen egg-white lysozyme complexes and their hydrolytic activity, J. Mol. Biol. 247 (2) (1995) 281–293.

[17] W.A. Hendrickson, Determination of macromolecular structures from anomalous diffraction of synchrotron radiation, Science 254 (5028) (1991) 51−58.
[18] S. Doniach, Changes in biomolecular conformation seen by small angle X-ray scattering, Chem. Rev. 101 (6) (2001) 1763−1778.
[19] D.P. Zhong, S.K. Pal, A.H. Zewail, Biological water: a critique, Chem. Phys. Lett. 503 (1−3) (2011) 1−11.
[20] C.P. Garnham, R.L. Campbell, P.L. Davies, Anchored clathrate waters bind antifreeze proteins to ice, Proc. Natl Acad. Sci. U.S.A. 108 (18) (2011) 7363−7367.
[21] S. Li, et al., Water − protein interactions of an arginine-rich membrane peptide in lipid bilayers investigated by solid-state nuclear magnetic resonance spectroscopy, J. Phys. Chem. B 114 (11) (2010) 4063−4069.
[22] H. Wendt, et al., Kinetics of folding of leucine-zipper domains, Biochemistry 34 (12) (1995) 4097−4107.
[23] H.D. Burrows, et al., Aggregation of the hairy rod conjugated polyelectrolyte poly{1,4-phenylene- 9,9-bis (4-phenoxybutylsulfonate) fluorene-2,7-diyl} in aqueous solution: an experimental and molecular modelling study, Phys. Chem. Chem. Phys. 10 (30) (2008) 4420−4428.
[24] L. Nilsson, B. Halle, Molecular origin of time-dependent fluorescence shifts in proteins, Proc. Natl Acad. Sci. U.S.A. 102 (39) (2005) 13867−13872.
[25] A. Elliott, E.J. Ambrose, C. Robinson, Chain configurations in natured and denatured insulin—evidence from infra-red spectra, Nature 166 (4213) (1950). 194-194.
[26] E.J. Ambrose, A. Elliott, Infra-red spectroscopic studies of globular protein structure, Proc. R. Soc. Lond. Ser. Math. Phys. Sci. 208 (1092) (1951) 75−90.
[27] R.R. Wiederkehr, H.G. Drickamer, Effect of pressure on cyanide and carbonyl spectra in solution, J. Chem. Phys. 28 (2) (1958) 311−316.
[28] S. Krimm, J. Bandekar, Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins, in: C.B. Anfinsen, J.T. Edsall, F.M. Richards (Eds.), Advances in Protein Chemistry, Academic Press, 1986, pp. 181−364.
[29] H. Torii, M. Tasumi, Application of the 3-dimensional doorway-state theory to analyses of the amide-I infrared bands of globular-proteins, J. Chem. Phys. 97 (1) (1992) 92−98.
[30] L. Genzel, et al., Low-frequency raman-spectra of lysozyme, Biopolymers 15 (1) (1976) 219−225.
[31] Y. Tomimatsu, et al., Raman spectra of a solid antifreeze glycoprotein and its liquid and frozen aqueous solutions, J. Biol. Chem. 251 (8) (1976) 2290−2298.
[32] T.C. Cheam, S. Krimm, Vibrational analysis of crystalline diketopiperazine-I. Raman and i.r. spectra, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 40 (6) (1984) 481−501.
[33] J.A. Drewes, K.L. Rowlen, Raman-spectroscopic study of the secondary structure of an antifreeze glycopeptide, Abstr. Pap. Am. Chem. Soc. 205 (1993). p. 28-Carb.
[34] I.K. Lednev, et al., Nanosecond UV resonance Raman examination of initial steps in alpha-helix secondary structure evolution, J. Am. Chem. Soc. 121 (16) (1999) 4076−4077.
[35] J.A. Zitzewitz, et al., Probing the folding mechanism of a leucine-zipper peptide by stopped-flow circular-dichroism spectroscopy, Biochemistry 34 (39) (1995) 12812−12819.
[36] T.A. Keiderling, et al., Site-specific conformational chirality in peptides as studied with vibrational circular dichroism using isotopic labeling, Abstr. Pap. Am. Chem. Soc. 219 (2000). U270.
[37] P. Manavalan, W.C. Johnson, Sensitivity of circular-dichroism to protein tertiary structure class, Nature 305 (5937) (1983) 831−832.
[38] C.D. Snow, et al., Absolute comparison of simulated and experimental protein-folding dynamics, Nature 420 (6911) (2002) 102−106.
[39] J. Bailey, et al., Circular polarization in star-formation regions: implications for biomolecular homochirality, Science 281 (5377) (1998) 672−674.
[40] A.D. Mehta, et al., Single-molecule biomechanics with optical methods, Science 283 (5408) (1999) 1689−1695.

[41] A.M. van Oijen, et al., Unraveling the electronic structure of individual photosynthetic pigment-protein complexes, Science 285 (5426) (1999) 400–402.
[42] S. Weiss, Fluorescence spectroscopy of single biomolecules, Science 283 (5408) (1999) 1676–1683.
[43] T.E.T. Hughes, et al., Structural basis of TRPV5 channel inhibition by econazole revealed by cryo-EM, Nat. Struct. Mol. Biol. 25 (1) (2018) 53–60.
[44] J. Frank, Advances in the field of single-particle cryo-electron microscopy over the last decade, Nat. Protoc. 12 (2) (2017) 209–212.
[45] H.Y. Liao, Y. Hashem, J. Frank, Efficient estimation of three-dimensional covariance and its application in the analysis of heterogeneous samples in cryo-electron microscopy, Structure 23 (6) (2015) 1129–1137.
[46] S. Mukamel, Multidimensional femtosecond correlation spectroscopies of electronic and vibrational excitations, Annu. Rev. Phys. Chem. 51 (2000) 691–729.
[47] E.T.J. Nibbering, H. Fidder, E. Pines, *Ultrafast chemistry: using time-resolved vibrational spectroscopy for interrogation of structural dynamics*, in, Annu. Rev. Phys. Chem. (2005) 337–367.
[48] M. Kowalewski, et al., Simulating coherent multidimensional spectroscopy of nonadiabatic molecular processes: from the infrared to the X-ray regime, Chem. Rev. 117 (19) (2017) 12165–12226.
[49] N. Demirdoven, M. Khalil, A. Tokmakoff, Correlated vibrational dynamics revealed by two-dimensional infrared spectroscopy, Phys. Rev. Lett. 89 (23) (2002).
[50] D. Laage, T. Elsaesser, J.T. Hynes, Water dynamics in the hydration shells of biomolecules, Chem. Rev. 117 (16) (2017) 10694–10725.
[51] J.R. Schmidt, S.A. Corcelli, J.L. Skinner, Ultrafast vibrational spectroscopy of water and aqueous N-methylacetamide: comparison of different electronic structure/molecular dynamics approaches, J. Chem. Phys. 121 (18) (2004) 8887–8896.
[52] T.M. Watson, J.D. Hirst, Theoretical studies of the amide I vibrational frequencies of Leu -enkephalin, Mol. Phys. 103 (11–12) (2005) 1531–1546.
[53] T. Hayashi, W. Zhuang, S. Mukamel, Electrostatic DFT map for the complete vibrational amide band of NMA, J. Phys. Chem. A 109 (43) (2005) 9747–9759.
[54] T.L. Jansen, J. Knoester, A transferable electrostatic map for solvation effects on amide I vibrations and its application to linear and two-dimensional spectroscopy, J. Chem. Phys. 124 (4) (2006) 044502.
[55] B. Blasiak, H. Lee, M. Cho, Vibrational solvatochromism: towards systematic approach to modeling solvation phenomena, J. Chem. Phys. 139 (4) (2013) 044111.
[56] S. Ham, et al., Correlation between electronic and molecular structure distortions and vibrational properties. II. Amide I modes of NMA-nD(2)O complexes, J. Chem. Phys. 118 (8) (2003) 3491–3498.
[57] M.H. Cho, Coherent two-dimensional optical spectroscopy, Chem. Rev. 108 (4) (2008) 1331–1418.
[58] I.W. Hamley, The amyloid beta peptide: a chemist's perspective. Role in Alzheimer's and fibrillization, Chem. Rev. 112 (10) (2012) 5147–5192.
[59] C.T. Middleton, et al., Two-dimensional infrared spectroscopy reveals the complex behaviour of an amyloid fibril inhibitor, Nat. Chem. 4 (5) (2012) 355–360.
[60] W. Kopec, et al., Direct knock-on of desolvated ions governs strict ion selectivity in K + channels, Nat. Chem. 10 (8) (2018) 813–820.
[61] W. Zhuang, D. Abramavicius, S. Mukamel, Dissecting coherent vibrational spectra of small proteins into secondary structural elements by sensitivity analysis, Proc. Natl Acad. Sci. U.S.A. 102 (21) (2005) 7443–7448.
[62] J.H. Choi, S.Y. Ham, M. Cho, Local amide I mode frequencies and coupling constants in polypeptides, J. Phys. Chem. B 107 (34) (2003) 9132–9138.
[63] S.Y. Cha, S.H. Ham, M.H. Cho, Amide I vibrational modes in glycine dipeptide analog: ab initio calculation studies, J. Chem. Phys. 117 (2) (2002) 740–750.
[64] J.H. Choi, S. Ham, M. Cho, Inter-peptide interaction and delocalization of amide I vibrational excitons in myoglobin and flavodoxin, J. Chem. Phys. 117 (14) (2002) 6821–6832.
[65] J.P. Wang, J.X. Chen, R.M. Hochstrasser, Local structure of beta-hairpin isotopomers by FTIR, 2D IR, and ab initio theory, J. Phys. Chem. B 110 (14) (2006) 7545–7555.

[66] W. Zhuang, T. Hayashi, S. Mukamel, Coherent multidimensional vibrational spectroscopy of biomolecules: concepts, simulations, and challenges, Angew. Chem. Int. Ed. 48 (21) (2009) 3750–3781.
[67] S. Krimm, Infrared spectra and chain conformation of proteins, J. Mol. Biol. 4 (6) (1962) 528–540.
[68] T. Miyazawa, E.R. Blout, Infrared spectra of polypeptides in various conformations—amide I and II bands, J. Am. Chem. Soc. 83 (3) (1961) 712–719.
[69] T. Miyazawa, Perturbation treatment of the characteristic vibrations of polypeptide chains in various configurations, J. Chem. Phys. 32 (6) (1960) 1647–1652.
[70] J.B. Bates, Fourier-transform infrared spectroscopy, Science 191 (4222) (1976) 31–37.
[71] K.J. Rothschild, W.A. Cantore, H. Marrero, Fourier-transform infrared difference spectra of intermediates in rhodopsin bleaching, Science 219 (4590) (1983) 1333–1335.
[72] Y.J. Chang, J. Edward, W. Castner, Fast responses from "slowly relaxing" liquids: a comparative study of the femtosecond dynamics of triacetin, ethylene glycol, and water, J. Chem. Phys. 99 (1993) 7289.
[73] Y. Tanimura, S. Mukamel, 2-Dimensional femtosecond vibrational spectroscopy of liquids, J. Chem. Phys. 99 (12) (1993) 9496–9511.
[74] P. Hamm, M.H. Lim, R.M. Hochstrasser, Structure of the amide I band of peptides measured by femtosecond nonlinear-infrared spectroscopy, J. Phys. Chem. B 102 (31) (1998) 6123–6138.
[75] W. Zhuang, et al., Simulating the T-jump-triggered unfolding dynamics of trpzip2 peptide and its time-resolved IR and two-dimensional IR signals using the Markov state model approach, J. Phys. Chem. B 115 (18) (2011) 5415–5424.
[76] D. Abramavicius, W. Zhuang, S. Mukamel, Peptide secondary structure determination by three-pulse coherent vibrational spectroscopies: a simulation study, J. Phys. Chem. B 108 (46) (2004) 18034–18045.
[77] D.C. Urbanek, et al., The two-dimensional vibrational echo of a nitrile probe of the villin HP35 protein, J. Phys. Chem. Lett. 1 (23) (2010) 3311–3315.
[78] M.C. Thielges, M.D. Fayer, Protein dynamics studied with ultrafast two-dimensional infrared vibrational echo spectroscopy, Acc. Chem. Res. 45 (11) (2012) 1866–1874.
[79] L. Wang, J.L. Skinner, Thermally induced protein unfolding probed by isotope-edited IR spectroscopy, J. Phys. Chem. B 116 (32) (2012) 9627–9634.
[80] T.L.C. Jansen, J. Knoester, Two-dimensional infrared population transfer spectroscopy for enhancing structural markers of proteins, Biophys. J. 94 (5) (2008) 1818–1825.
[81] C.W. Liang, T.L.C. Jansen, J. Knoester, Proton transport in biological systems can be probed by two-dimensional infrared spectroscopy, J. Chem. Phys. 134 (4):) (2011) 044502.
[82] C.W. Liang, J. Knoester, T.L.C. Jansen, Proton transport in a membrane protein channel: two-dimensional infrared spectrum modeling, J. Phys. Chem. B 116 (22) (2012) 6336–6345.
[83] H.T. Kratochvil, et al., Instantaneous ion configurations in the K^+ ion channel selectivity filter revealed by 2D IR spectroscopy, Science 353 (6303) (2016) 1040–1044.
[84] A. Ghosh, et al., 2D IR spectroscopy reveals the role of water in the binding of channel-blocking drugs to the influenza M2 channel, J. Chem. Phys. 140 (23) (2014) 235105.
[85] C.S. Peng, K.C. Jones, A. Tokmakoff, Anharmonic vibrational modes of nucleic acid bases revealed by 2D IR spectroscopy, J. Am. Chem. Soc. 133 (39) (2011) 15650–15660.
[86] G. Hithell, et al., Ultrafast 2D-IR and optical Kerr effect spectroscopy reveal the impact of duplex melting on the structural dynamics of DNA, Phys. Chem. Chem. Phys. 19 (16) (2017) 10333–10342.
[87] K.K. Lee, et al., Ultrafast internal rotational dynamics of the azido group in (4S)-azidoproline: chemical exchange 2DIR spectroscopic investigations, Chem. Phys. 396 (2012) 23–29.
[88] J.R. Zheng, et al., Ultrafast carbon-carbon single-bond rotational isomerization in room-temperature solution, Science 313 (5795) (2006) 1951–1955.
[89] Y.S. Kim, R.M. Hochstrasser, Chemical exchange 2D IR of hydrogen-bond making and breaking, Proc. Natl Acad. Sci. U.S.A. 102 (32) (2005) 11185–11190.
[90] T. Elsaesser, Two-dimensional infrared spectroscopy of intermolecular hydrogen bonds in the condensed phase, Acc. Chem. Res. 42 (9) (2009) 1220–1228.
[91] S. Park, M.D. Fayer, Hydrogen bond dynamics in aqueous NaBr solutions, Proc. Natl Acad. Sci. U.S.A. 104 (2007) 16731–16738.

[92] J.L. Skinner, B.M. Auer, Y.-S. Lin, Vibrational line shapes, spectral diffusion, and hydrogen bonding in liquid water, Advances in Chemical Physics, John Wiley & Sons, Inc, 2008, pp. 59–103.
[93] T.L. Jansen, et al., Stochastic Liouville equations for hydrogen-bonding fluctuations and their signatures in two-dimensional vibrational spectroscopy of water, J. Chem. Phys. 123 (11) (2005) 114504.
[94] H.M. Muller-Werkmeister, et al., Ultrafast hopping from band to band: assigning infrared spectra based on vibrational energy transfer, Angew. Chem. Int. Ed. 52 (24) (2013) 6214–6217.
[95] D.V. Kurochkin, S.R.G. Naraharisetty, I.V. Rubtsov, A relaxation-assisted 2D IR spectroscopy method, Proc. Natl Acad. Sci. U.S.A. 104 (36) (2007) 14209–14214.
[96] C.S. Peng, C.R. Baiz, A. Tokmakoff, Direct observation of ground-state lactam–lactim tautomerization using temperature-jump transient 2D IR spectroscopy, Proc. Natl Acad. Sci. U.S.A. 110 (23) (2013) 9243.
[97] Baiz, C.R., et al., Tracking ultrafast chemical reaction dynamics using transient 2DIR spectroscopy, in: International Conference on Ultrafast Phenomena, 2010, Optical Society of America, Snowmass, CO.
[98] J. Frenkel, On the transformation of light into heat in solids. I, Phys. Rev. 37 (1) (1931) 17–44.
[99] S. Mukamel, Principles of Nonlinear Optical Spectroscopy, Oxford University Press, 1995.
[100] T. Hayashi, H. Hamaguchi, Solvent induced dynamic polarization, vibrational dephasing, and infrared band shape of the C = O stretch mode of acetone in acetonitrile: a new theoretical approach, Chem. Phys. Lett. 326 (1–2) (2000) 115–122.
[101] C.J. Fecko, et al., Ultrafast hydrogen-bond dynamics in the infrared spectroscopy of water, Science 301 (5640) (2003) 1698–1702.
[102] J.R. Hill, et al., Vibrational dynamics of carbon monoxide at the active site of myoglobin: picosecond infrared free-electron laser pump-probe experiments, J. Phys. Chem. 98 (43) (1994) 11213–11219.
[103] K.A. Merchant, et al., Myoglobin-CO substate structures and dynamics: multidimensional vibrational echoes and molecular dynamics simulations, J. Am. Chem. Soc. 125 (45) (2003) 13804–13818.
[104] S.F. Boys, F. Bernardi, Calculation of small molecular interactions by differences of separate total energies—some procedures with reduced errors, Mol. Phys. 19 (4) (1970) 553–566.
[105] J.H. Jensen, M.S. Gordon, An approximate formula for the intermolecular Pauli repulsion between closed shell molecules, Mol. Phys. 89 (5) (1996) 1313–1325.
[106] J.H. Jensen, M.S. Gordon, An approximate formula for the intermolecular Pauli repulsion between closed shell molecules. II. Application effective fragment potential method, J. Chem. Phys. 108 (12) (1998) 4772–4782.
[107] C. Reichardt, Solvatochromic dyes as solvent polarity indicators, Chem. Rev. 94 (8) (1994) 2319–2358.
[108] D. Ben-Amotz, M.R. Lee, S.Y. Cho, D.J. List, Solvent and pressure-induced perturbations of the vibrational potential surface of acetonitrile, J Phys Chem. 96 (12) (1992) 8781–8792.
[109] L.R. Pratt, D. Chandler, Effective intramolecular potentials for molecular Bromine in Argon, J Phys Chem. 72 (7) (1980) 4045–4048.
[110] K.S. Schweizer, D. Chandler, Vibrational dephasing and frequency shifts of polyatomic molecules in solution, J Phys Chem. 76 (5) (1982) 2296–2314.
[111] R.R. Wiederkehr, H.G. Drickamer, Effect of Pressure on Cyanide and Carbonyl Spectra in Solution, J. Phys Chem. 28 (2) (1958) 311–316.
[112] M.R. Zakin, D.R. Herschbach, Density dependence of attractive forces for hydrogen stretching vibrations of molecules in compressed liquids, J Phys Chem. 89 (4) (1988) 2380–2387.
[113] G. Eaton, M.C.R. Symons, P.P. Rastogi, *Spectroscopic studies of the solvation of amides with N-H groups. 1. The carbonyl group*, J. Chem. Soc. Faraday Trans. I 85 (1989) 3257–3271.
[114] B. Blasiak, M. Cho, Vibrational solvatochromism. II. A first-principle theory of solvation-induced vibrational frequency shift based on effective fragment potential method, J. Chem. Phys. 140 (16) (2014) 164107.
[115] W.A. Sokalski, R.A. Poirier, Cumulative atomic multipole representation of the molecular charge-distribution and its basis set dependence, Chem. Phys. Lett. 98 (1) (1983) 86–92.

[116] H. Torii, M. Tasumi, Ab initio molecular orbital study of the amide I vibrational interactions between the peptide groups in di- and tripeptides and considerations on the conformation of the extended helix, J. Raman Spectrosc. 29 (1) (1998) 81–86.
[117] T.L. Jansen, et al., Modeling the amide I bands of small peptides (vol 125, 044312, 2006), J. Chem. Phys. 136 (20) (2012) 209901.
[118] T.L. Jansen, et al., Modeling the amide I bands of small peptides, J. Chem. Phys. 125 (4) (2006) 044312.
[119] T. Hayashi, S. Mukamel, Vibrational-exciton couplings for the amide I, II, III, and a modes of peptides, J. Phys. Chem. B 111 (37) (2007) 11032–11046.
[120] P. Hamm, S. Woutersen, Coupling of the amide I modes of the glycine dipeptide, Bull. Chem. Soc. Jpn. 75 (5) (2002) 985–988.
[121] J.W. Brauner, C.R. Flach, R. Mendelsohn, Quantitative reconstruction of the amide I contour in the IR spectra of globular proteins: from structure to spectrum, J. Am. Chem. Soc. 127 (1) (2005) 100–109.
[122] J.W. Brauner, C. Dugan, R. Mendelsohn, C-13 isotope labeling of hydrophobic peptides. Origin of the anomalous intensity distribution in the infrared amide I spectral region of beta-sheet structures, J. Am. Chem. Soc. 122 (4) (2000) 677–683.
[123] S. Mukamel, D. Abramavicius, Many-body approaches for simulating coherent nonlinear spectroscopies of electronic and vibrational excitons, Chem. Rev. 104 (4) (2004) 2073–2098.
[124] M. Cho, Two-Dimensional Optical Spectroscopy, CRC Press, 2009.
[125] P. Hamm, M. Zanni, Concepts and Methods of 2D Infrared Spectroscopy, Cambridge University Press, 2011.
[126] N. Simpson, N.T. Hunt, Ultrafast 2D-IR spectroscopy of haemoproteins, Int. Rev. Phys. Chem. 34 (3) (2015) 361–383.
[127] T.L. Jansen, W. Zhuang, S. Mukamel, Stochastic Liouville equation simulation of multidimensional vibrational line shapes of trialanine, J. Chem. Phys. 121 (21) (2004) 10577–10598.
[128] R. Kubo, Stochastic Liouville equations, J. Math. Phys. 4 (2) (1963) 174–183.
[129] W. Zhuang, et al., Simulation protocols for coherent femtosecond vibrational spectra of peptides, J. Phys. Chem. B 110 (7) (2006) 3362–3374.
[130] T.L. Jansen, J. Knoester, Nonadiabatic effects in the two-dimensional infrared spectra of peptides: application to alanine dipeptide, J. Phys. Chem. B 110 (45) (2006) 22910–22916.
[131] B.M. Auer, J.L. Skinner, Dynamical effects in line shapes for coupled chromophores: time-averaging approximation, J. Chem. Phys. 127 (10) (2007) 104105.
[132] T.L.C. Jansen, W.M. Ruszel, Motional narrowing in the time-averaging approximation for simulating two-dimensional nonlinear infrared spectra, J. Chem. Phys. 128 (21) (2008) 214501.
[133] Y. Tanimura, R. Kubo, Time evolution of a quantum system in contact with a nearly Gaussian-Markoffian noise bath, J. Phys. Soc. Jpn. 58 (1) (1989) 101–114.
[134] F. Sanda, S. Mukamel, Stochastic simulation of chemical exchange in two dimensional infrared spectroscopy, J. Chem. Phys. 125 (1) (2006) 014507.
[135] Y.Q. Gao, Simple theoretical model for ion cooperativity in aqueous solutions of simple inorganic salts and its effect on water surface tension, J. Phys. Chem. B 115 (2011) 12466–12472.
[136] Y.Q. Gao, An integrate-over-temperature approach for enhanced sampling, J. Chem. Phys. 128 (6) (2008) 064105.
[137] Q. Shao, L.J. Yang, Y.Q. Gao, A test of implicit solvent models on the folding simulation of the GB1 peptide, J. Chem. Phys. 130 (19) (2009) 195104.
[138] L.J. Yang, Q. Shao, Y.Q. Gao, Comparison between integrated and parallel tempering methods in enhanced sampling simulations, J. Chem. Phys. 130 (12) (2009) 124111.
[139] L.J. Yang, Q. Shao, Y.Q. Gao, Thermodynamics and folding pathways of Trpzip2: an accelerated molecular dynamics simulation study, J. Phys. Chem. B 113 (3) (2009) 803–808.
[140] Q.A. Shao, H.Y. Wei, Y.Q. Gao, Effects of turn stability and side-chain hydrophobicity on the folding of beta-structures, J. Mol. Biol. 402 (3) (2010) 595–609.

[141] A.W. Smith, et al., Melting of a beta-hairpin peptide using isotope-edited 2D IR spectroscopy and simulations, J. Phys. Chem. B 114 (34) (2010) 10913–10924.
[142] T.M. Wu, et al., Discriminating trpzip2 and trpzip4 peptides' folding landscape using the two-dimensional infrared spectroscopy: a simulation study, J. Chem. Phys. 140 (5) (2014) 055101.
[143] J. Song, et al., Investigating the structural origin of Trpzip2 temperature dependent unfolding fluorescence line shape based on a Markov state model simulation, J. Phys. Chem. B 116 (42) (2012) 12669–12676.
[144] J.D. Chodera, et al., Automatic discovery of metastable states for the construction of Markov models of macromolecular conformational dynamics, J. Chem. Phys. 126 (15) (2007) 155102.
[145] W.C. Swope, J.W. Pitera, F. Suits, Describing protein folding kinetics by molecular dynamics simulations. 1. Theory, J. Phys. Chem. B 108 (21) (2004) 6571–6581.

CHAPTER 5

Infrared spectroscopy and imaging for understanding neurodegenerative protein-misfolding diseases

Lisa M. Miller

National Synchrotron Light Source II, Brookhaven National Laboratory, Upton, NY, United States

5.1 Introduction to Fourier transform infrared spectroscopy and protein misfolding

5.1.1 In vitro studies

It is well established that Fourier transform infrared spectroscopy (FTIR) can be used to examine the secondary structure of proteins [1,2]. The vibrations of the protein backbone, most notably the C = O stretch and C−N bend, lead to the amide I and amide II absorption peaks, respectively. Moreover, based on the environment of these backbone modes, their absorption frequencies can change. This is best characterized in the amide I band, where the different secondary structural components, for example, α-helix, β-sheet, and unordered, absorb at different frequencies (Fig. 5.1). Hence, the secondary structural composition of a protein can be determined by deconvolution of the amide I band.

For protein-misfolding diseases, the structural change associated with the diseased state is frequently a change from α-helical and/or unordered structure to β-sheet, which involves a shift in the amide I band to lower frequency due to the increased β-sheet contribution at $\sim 1630\ \mathrm{cm}^{-1}$. Attenuated total reflection-Fourier transform infrared spectroscopy (ATR-FTIR) has been used to show that a number of proteins that form aggregates acquire this new β-sheet band following the aggregation process, which involves the formation of new intermolecular interactions [3]. In addition to steady state, in vitro studies, time-resolved measurement can be performed in order to assess folding intermediates. For example, the structural behavior of a model amyloid beta (Aβ) peptide and a series of mutants was examined during solvent exchange and showed that aggregation of the peptide proceeds via a helical intermediate, preceded by creation of 3_{10}-helix/3_{10}-turnstructure [4].

Vibrational Spectroscopy in Protein Research.
DOI: https://doi.org/10.1016/B978-0-12-818610-7.00005-0

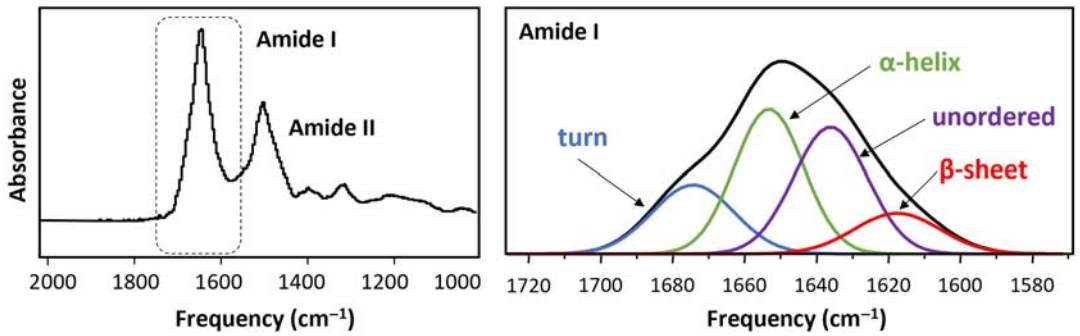


Figure 5.1
(*Left*) Fourier transform infrared spectroscopy (FTIR) spectrum of myoglobin protein in solution where the amide I and amide II bands are labeled. (*Right*) Deconvolved amide I band illustrating the peak frequencies of the secondary structure components.

Beyond the observation of β-sheet formation, FTIR spectroscopy can go further to examine the parallel versus antiparallel nature of the β-sheet [5]. In antiparallel β-sheet structures, the amide I region shows a major component at $\sim 1630\ \mathrm{cm}^{-1}$, whereas the minor component appears at $1695\ \mathrm{cm}^{-1}$. For parallel β-sheet structures, the amide I region exhibits only the major component at $1630\ \mathrm{cm}^{-1}$. Thus a 1695/1630 intensity ratio has been suggested to be proportional to the percentage of antiparallel arrangement of the β-strands in a β-sheet. This has proven important in distinguishing Aβ oligomers versus fibrils in Alzheimer's disease (AD) [5] and it has also recently been shown that Aβ deposits in AD brain vessels have a different β-sheet structure than extracellular parenchymal plaques [6].

5.1.2 Isotopic labeling

The use of isotopes has also been an important part of studying protein folding in solution. Firstly, it should be noted that most in vitro FTIR studies of protein structure in solution are performed in D_2O. Water has a broad O−H bending mode that overlaps the amide 1 band, which is shifted to lower frequency by deuterium (O−D) substitution [7]. Secondly, ^{13}C or ^{15}N labeling of the peptide backbone can be used to detect vibrational dipolar coupling, implying close physical proximity. This approach helps to tease out the particular backbone residues involved in protein misfolding, providing even more structural details than unlabeled peptide [8]. For Aβ, ^{13}C labeling of a series of backbone residues showed that the fibrils can assume both a parallel and antiparallel configuration based on the amino acid sequence and environment [9]. A similar approach was used to show that seeding of Aβ fibrils in vitro with vascular amyloid seeds from tissue can template the formation of antiparallel fibrils (Fig. 5.2) [6].

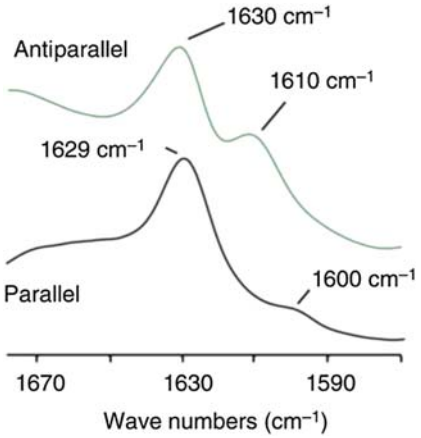


Figure 5.2
Fourier transform infrared spectroscopy (FTIR) spectra of the amyloid beta 40-Dutch Iowa (Aβ40-DI) mutant in the antiparallel (*green*) and parallel (*black*) β-sheet configuration. The Aβ40-DI peptide contained a $^{13}C=O$ label at Gly33, a residue within the hydrophobic C-terminus that typically forms β-sheet in Aβ fibrils. The ^{13}C substitution lowers the frequency of the C = O stretching mode, resulting in isotope-shifted peaks at 1610 and 1600 cm^{-1}, which differ in intensity depending on whether the β-strands hydrogen-bond in a parallel or antiparallel orientation. The increase of intensity at 1610 cm^{-1} relative to 1600 cm^{-1} is a signature of antiparallel β-sheet. Source*: From F. Xu, Z. Fu, S. Dass, A.E. Kotarba, J. Davis, S.O. Smith, et al., Cerebral vascular amyloid seeds drive amyloid beta-protein fibril assembly with a distinct anti-parallel structure, Nat. Commun. 7 (2016) 13527 [6].*

5.1.3 Infrared microspectroscopy

Beyond in vitro studies, FTIR spectroscopic microscopy can be used to examine protein structure and misfolding directly within cells and tissues. Traditionally, FTIR microscopes can achieve a spatial resolution of ~20–50 microns. This is sufficient for examining large misfolded aggregates in human tissue. For example, the secondary structure of Aβ plaques in human autopsy tissue was compared to the in vitro structure of Aβ and found to be significantly different, where the β-sheet peak in the tissue was downshifted to lower frequency, suggesting a more hydrophobic environment than Aβ fibrils in solution [10].

While some aggregates are large enough to study with conventional infrared sources, most aggregates are considerably smaller—frequently 1 μm or less in size. In this case, the advent of the synchrotron-powered infrared microscope has dramatically improved the spatial resolution of the data to the diffraction limit of ~2–5 microns, with a dramatically improved signal to noise [2,7,11]. These measurements have been further improved by coupling the synchrotron source to full-field infrared microscopes, that is, those equipped with focal plane array (FPA) detectors [12,13]. With the high brightness of the synchrotron

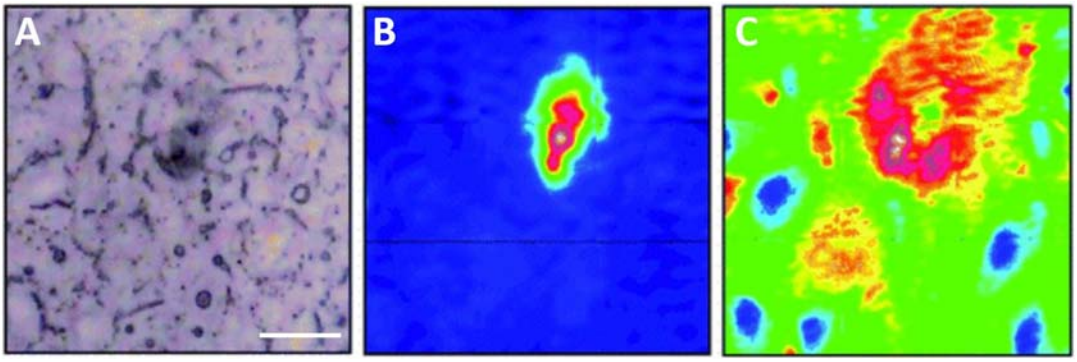


Figure 5.3
Visible and Fourier transform infrared spectroscopy (FTIR) images of a dense core plaque from a 15-month-old CRND8 mouse model of Alzheimer's disease (AD) at 15 months old. (A) Dense core appears darker than surrounding unstained tissue under white light. (B) Aggregated plaque is shown from the intensity of the β-sheet shoulder. (C) Plaques are surrounded by a lipid membrane-like signature that infiltrates to the core of all plaques. Scale bar is 30 μm. Source: *From C.R. Liao, M. Rak, J. Lund, M. Unger, E. Platt, B.C. Albensi, et al., Synchrotron FTIR reveals lipid around and within amyloid plaques in transgenic mice and Alzheimer's disease brain, Analyst 138 (14) (2013) 3991–3997 [15].*

source and the multipixels of the FPA, high-resolution, high-quality spectra can be obtained quickly [14]. Moreover, since the entire FTIR spectrum is obtained simultaneously, correlative information on other chemical components can be obtained simultaneously. For example, this approach has been used to study whole hippocampal and cortical regions of the brain in AD, where results showed not only the structure of the dense core plaques but also a lipid membrane-like spectral signature surrounding and infiltrating into the plaques (Fig. 5.3) [15].

The full-field nature of the FPA detector also enables time-resolved measurements to be performed. In familial amyotrophic lateral sclerosis (ALS), mutant copper–zinc superoxide dismutase (SOD1) has been shown to form aggregates that lead to neuronal death. Time-resolved, full-field synchrotron FTIR imaging of a cell culture model of SOD1-ALS demonstrated the misfolding pathway of the protein directly within the cells, which was found to vary based on the SOD1 mutation [16].

5.1.4 Infrared nanospectroscopy

Protein misfolding in neurodegenerative diseases always begins with the formation of nanoscale insoluble structures that aggregate and/or fibrillize into a supramolecular structure. It has been widely hypothesized that these oligomers and protofibrils are more toxic than the macroscopic deposits [17]. Hence there continues to be the desire to

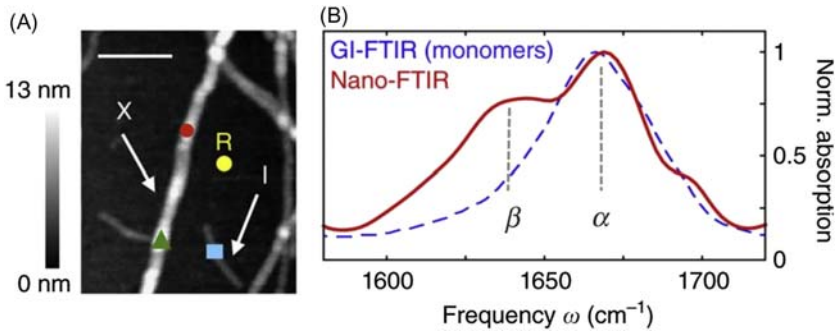


Figure 5.4

Infrared nanospectroscopy (nano-FTIR) of the secondary structure in individual insulin fibrils. (A) Topography of insulin fibrils on a silicon substrate. The arrows indicate a type I fibril (I) and a 9-nm-thick fibril composed of several protofilaments (X), respectively. Scale bar: 200 nm. (B) Nano- Fourier transform infrared spectroscopy (FTIR) spectrum of a 9-nm-thick insulin fibril (*red*) recorded at the position marked by the red dot in (A). The position marked by the yellow dot and R indicates where the reference spectrum was recorded. The dashed blue line shows a grazing incidence Fourier transform infrared spectroscopy (FTIR) spectrum of insulin monomers on a gold substrate. Source*: From I. Amenabar, S. Poly, W. Nuansing, E.H. Hubrich, A.A. Govyadinov, F. Huth, et al., Structural analysis and mapping of individual protein complexes by infrared nanospectroscopy, Nat. Commun. 4 (2013) 2890 [20].*

understand the misfolding process on an even smaller size scale. One drawback of FTIR microspectroscopy is that the diffraction limit prevents nanoscale imaging of protein structure. However, recent instrumentation developments that couple atomic force microscopy (AFM) with FTIR spectroscopy now provide a way to examine individual proteins at the nanoscale, that is, infrared nanospectroscopy (nano-FTIR) [18]. This technique is based on scattering-type scanning near-field optical microscopy where infrared images are obtained by recording the infrared light scattered from a scanning probe tip, which is typically a metalized AFM tip.

As an example, nano-IR studies on individual silk protein fibers have shown their high β-sheet crystal content, which gives them their well-known strength and extensibility [19]. In the field of neurodegeneration, studies on 3-nm insulin fibrils, which are often a model system for neurodegenerative diseases research, revealed that they contained a surprisingly high degree of α-helical structure [20]. As can be seen in Fig. 5.4, band decomposition of the nano-FTIR spectrum of insulin fibrils reveals two major bands at 1639 and at 1671 cm^{-1}, confirming that β-sheet and α-helical structures are the predominant contributions. Current models have suggested that the core of the insulin fibrils is purely β-sheets, suggesting that the fibril shell is mainly α-helical structure and not randomly organized.

5.2 Applications of Fourier transform infrared spectroscopy to neurodegenerative diseases

5.2.1 Alzheimer's disease

Alzheimer's disease is the most common age-related neurodegenerative disease, affecting approximately 2% of the population in industrialized countries [21]. Clinically, AD is characterized by memory loss and diminished cognitive functioning. Pathologically, amyloid plaques and neurofibrillary tangles (NFT) are found in the AD brain [22]. The amyloid plaques consist primarily of aggregated Aβ peptide, which is produced by the cleavage of the amyloid precursor protein (APP). The functions of both Aβ and APP are unclear, but there is evidence that APP may regulate neuronal survival, neurite outgrowth, synaptic plasticity, and cell adhesion [23]. NFTs are formed from paired helical filaments (PHFs) of hyperphosphorylated tau, a microtubule-associated protein, and are found in other neurodegenerative diseases such as frontotemporal dementia [24].

Cleavage of APP results in the Aβ peptide with lengths from 37 to 43 amino acids. The most abundant cleavage product is Aβ40, but the primary component in parenchymal plaques in AD is Aβ42. Fibril formation has been thoroughly studied in vitro using synthetic Aβ peptides and is described as a nucleation-dependent polymerization process. During this process, where there is a slow nucleation phase followed by a rapid elongation reaction, a range of aggregated species precede fibril formation. It is thought that these soluble oligomers and protofibrils are the toxic species related to the onset and development of AD [17].

The Aβ peptide has been studied with FTIR spectroscopy for almost 30 years. The secondary structure has been examined as a function of peptide length, pH, mutation, and under denaturing conditions [25,26]. The general consensus is that the secondary structure of Aβ is widely variable in α-helical, β-sheet, and unordered content as the environment varies; however, under physiological conditions, two processes—random/helix conversion to β-sheet and aggregation of β-sheets—in multiple conformational states were observed before the most stable form was attained [27].

A combination of nuclear magnetic resonance (NMR) and FTIR, along with single-touch AFM, was used to establish the structural transitions involved in fibril formation [8,28,29]. Under conditions favorable for the nucleated conformation conversion, the Aβ42 peptide aggregates into largely unstructured low-molecular weight oligomers that stack to form high-molecular weight oligomers and to laterally associate to form protofibrils [28]. During protofibril formation, the low-molecular weight oligomers do not contain significant β-sheet structure. Rather, FTIR spectra of Aβ42 containing a ^{13}C-labeled Gly33 exhibit the transient appearance of an IR band at 1610 cm^{-1} revealing a transient antiparallel β-hairpin structure

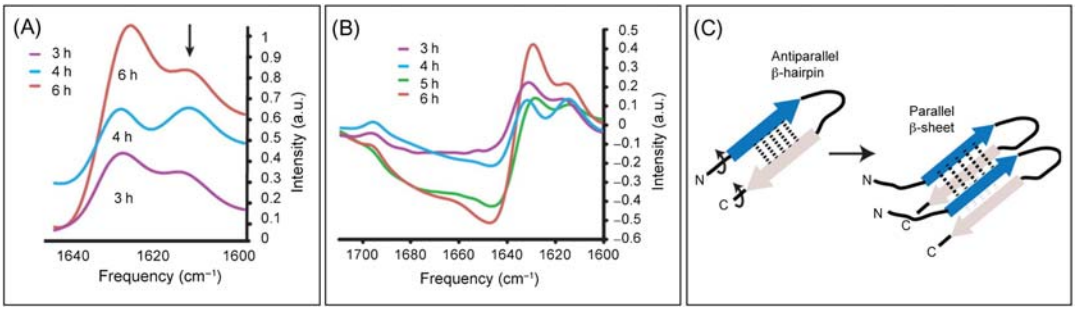


Figure 5.5

(A) The Fourier transform infrared spectroscopy (FTIR) spectra of Aβ42 containing ^{13}C-Gly33 during fibril formation showing the appearance of a peak at 1610 cm^{-1}, attributed to a transient antiparallel β-hairpin structure prior to fibril formation. (B) FTIR difference spectra obtained by subtracting the spectrum at time zero from spectra obtained at each incubation time point showing the loss of unordered structure and the appearance of β-sheet structure. (C) Conversion of antiparallel β-hairpin secondary structure to β-sheet secondary structure. Intramolecular hydrogen bonding stabilizes the antiparallel β-strands in the β-hairpin. Rotation of these β-strands allows the formation of intermolecular hydrogen bonds, which nucleates β-sheet secondary structure. Source*: From Z. Fu, D. Aucoin, J. Davis, W.E. Van Nostrand, S.O. Smith, Mechanism of nucleated conformational conversion of Abeta42, Biochemistry 54 (27) (2015) 4197–4207.*

prior to fibril formation (Fig. 5.5) [29]. This is the first step in fibril formation, where the antiparallel β-hairpin is stabilized by intramonomer hydrogen bonding [8]. As Aβ fibrils begin to form, the FTIR spectra of Aβ42 exhibit a loss of unordered structure between 1620 and 1700 cm^{-1} and a shift of the amide I peak from 1640 to 1630 cm^{-1} [29]. The antiparallel β-hairpins then associate into a cross β-sheet structure with parallel and in-register beta-strands having intermonomer hydrogen bonding. Once assembled, Aβ fibrils exhibit a cross-β molecular structures, where ribbon-like β-sheets have β-strands that run perpendicular to the long fibril axis and interstrand hydrogen bonds run parallel to the long fibril axis [28].

While wild-type (WT), full-length (40- and 42-residue) Aβ fibrils have been shown to contain cross-β structures in which the β-sheets have an in-register parallel supramolecular organization, certain mutations of the Aβ peptide can yield either a parallel and antiparallel cross-β structure [30]. For the Iowa-type mutation, D23N-Aβ40, antiparallel fibrils are thermodynamically metastable with respect to conversion to parallel structures, propagate less efficiently than parallel fibrils in seeded fibril growth, and therefore must nucleate more efficiently than parallel fibrils in order to be observable. Experiments in neuronal cell cultures indicate that both antiparallel and parallel D23N-Aβ40 fibrils are cytotoxic.

Not only has vibrational spectroscopy been used to study isolated Aβ peptides and fibrils, infrared microspectroscopy has been used to examine amyloid structure within

Alzheimer's tissue. The first studies, which used a synchrotron infrared microscope, confirmed that the in situ structure of dense core plaques in the parenchyma of human brain tissue were primarily β-sheet in structure [31]. Later studies took this further to show that the β-sheet peak in the brain tissue was shifted to lower frequency compared to in vitro Aβ aggregates, suggesting that plaques have a more hydrophobic core in the tissue [10]. In contrast, the spectra of diffuse plaques in the brain showed an amorphous secondary structure [32].

One advantage to performing infrared microspectroscopy on AD tissue is that the full spectrum provides the ability to correlate other chemical features with the Aβ plaques. For example, studies showed a lipid membrane-like spectral signature surrounding and infiltrating the dense core plaques [15,33] (Fig. 5.3), which have been attributed to astrogliosis—an abnormal increase in the number of astrocytes due to the destruction of nearby neurons [34]. A decrease in the unsaturated lipid levels was also observed by FTIR imaging in a double trangenic mouse model of AD, suggesting evidence of lipid peroxidation in the white matter [35]. And microcrystalline deposits of creatine have also been observed in the hippocampus of human AD brain biopsies [36].

5.2.1.1 Amyloid precursor protein

In addition to Aβ, FTIR has also been used to examine the structure of the APP. The Aβ peptide is generated by the proteolysis of the beta C-terminal fragment (beta-CTF) of APP. Familial mutations in the beta-CTF, such as the A21G Flemish mutation, can increase Aβ production. FTIR spectroscopy combined with solid-state NMR spectroscopy showed that the Flemish mutation alters the structure of the first 55 residues of the beta-CTF [37]. The mutation increases the α-helical structure from Gly25 to Gly29, which is a region near the membrane surface and thought to interact with cholesterol. Incorporation of cholesterol into model membranes enhanced these structural changes, suggesting a common link between familial mutations and the cellular environment.

5.2.1.2 Tau and neurofibrillary tangles

A second pathological feature of AD is the formation of NFTs composed of the tau protein. Tau stabilizes neuronal microtubules and aids in neuronal development and plasticity. In AD, tau becomes hyperphosphorylated, forming PHFs that can become NFTs. Electron microscope images of PHFs clearly show pairs of twisted ribbons with 80-nm periodicity [38]. However, while it is commonly believed that β-sheet formation is the general mechanism for aggregate formation in neurodegenerative diseases, there has been much debate over the aggregation pathway and structure of the PHFs.

In solution, the soluble tau protein is one of the longest natively unfolded proteins, having very little secondary structure over a sequence of 441 amino acids [39]. The physiological function of tau is to stabilize microtubules by binding to the surface of the well-folded

protein, tubulin. But due to its intrinsically unordered structure, aggregation of the protein is highly dependent upon its environment, including temperature, pH, ionic strength, and hydrophobicity of the environment [40]. FTIR and circular dichroism (CD) spectroscopy of PHF formation under different conditions have shown the formation of β-sheet structures in some cases [39,41] and α-helical structures in others [38,42]. Moreover, the different secondary structures led to different filament morphologies, where the traditional phytosphingosine filaments had a higher α-helix structure, whereas environmental conditions that promoted higher β-sheet content resulted in more straight filaments [40]. Hence, the consensus finding is that the tau environment, including phosphorylation or nonproteinaceous components, is able to induce a heterogeneous distribution of tau conformations due to the unordered character of the soluble protein.

5.2.2 Cerebral amyloid angiopathy

Cerebral amyloid angiopathy (CAA) is a prominent disorder affecting the blood vessels of the brain that can cause cognitive impairment and dementia. The most common form of CAA, which is present in >80% of elderly individuals, results from the accumulation of the Aβ protein in the vessels of the brain [6].

While accumulation of misfolded Aβ is most recognized in extracellular parenchymal plaques in AD, deposition in the brain vasculature also occurs. Interestingly, however, parenchymal plaques are largely composed of Aβ42, whereas cerebral vascular amyloid is mainly composed of the shorter Aβ40 species. While spontaneous CAA forms with WT Aβ40, several familial forms of the disease result from specific mutations in the Aβ40 peptide. For example, the Dutch- and Iowa-types of familial CAA are caused by the E22Q and D23N mutation in Aβ40, respectively, both leading to early-onset deposition of cerebral vascular amyloid, brain hemorrhages in mid-life, and early cognitive impairment.

The neurotoxic effects of these CAA mutants were examined in PC12 cell culture and results showed a good correlation between cell toxicity and the aggregation rates of the Aβ40 and Aβ42 mutants at positions 22 and 23 [43]. ATR-FTIR analysis of the same aggregates showed Aβ42 mutants at positions 22 and 23 had a higher β-sheet structure than WT Aβ42, indicating a correlation between neurotoxicity, aggregate formation, and β-sheet structure. As mentioned earlier, NMR and isotope-edited FTIR showed that the predominant fibril form of Aβ contains a cross β-sheet structure in which the individual β-strands have a parallel, in-register orientation [30,44]. In contrast, the CAA mutant Aβ40-Iowa can form fibrils with both parallel and antiparallel structure, where the antiparallel form of Aβ40-Iowa was thermodynamically metastable with respect to the parallel form. In addition, fibril seeds from both structures can promote rapid fibril assembly of WT Aβ peptides with the corresponding fibrillar signatures [6].

Together, these studies suggest that structural differences might exist between cerebrovascular Aβ deposits and parenchymal plaque Aβ and that perhaps these differences account for the specific developmental aspects of CAA and some of the unique pathological responses to CAA that result in cognitive impairment, brain hemorrhage, and dementia.

5.2.3 Parkinson's disease

Parkinson's disease (PD) is a neurodegenerative disorder attributed to the loss of dopaminergic neurons from the substantia nigra in the brain [45]. PD is characterized by the formation of intraneuronal, cytoplasmic Lewy bodies (LBs), which contain fibrillar α-synuclein protein in a form that resembles Aβ fibrils in AD plaques.

In solution, α-synuclein is an intrinsically disordered protein, with three Tyr residues in the C-terminal region and one in the N-terminus. However, it can also exist as nonamyloid oligomers, ordered amyloid oligomers, and fibrils. As such, it is important to understand which conformer(s) contribute to specific PD phenotypes [46]. Structurally, it is thought that interactions between the C-terminus and the central portion of the molecule may prevent or minimize aggregation. Studies have shown that fibril formation was completely inhibited by replacing the three C-terminal Tyr residues with Ala [47]. FTIR spectra of WT α-synuclein show largely unordered structure, with some β-sheet components.

In contrast, examination of the oligomers and fibrils using FTIR showed that the fibrillar species are primarily composed of parallel β-sheets (band at 1620−1630 cm^{-1}) and that the oligomeric species are primarily antiparallel (band at 1620−1630 cm^{-1} and shoulder at 1695 cm^{-1}) [46].

Two familial forms of PD that lead to early-onset disease are caused by point mutations (A53T, A30P) in the gene encoding for α-synuclein. Oligomeric forms of these mutants accumulate more rapidly and/or persist for longer periods of time than oligomeric WT α-synuclein, suggesting a link between oligomerization and cell death [45]. The β-sheet secondary structure of fibrillar α-synuclein, as determined by FTIR and CD, is consistent with the cross-β structure suggested by X-ray and electron diffraction analysis of fibrils comprising human, rat, and zebra finch α-synuclein [45].

Injection of the β-sheet oligomers into the mouse brain produced a small but significant loss of dopamine neurons in the substantia nigra [46]. However, injection of small β-sheet fibril fragments produced the most dramatic phenotypes, including reduction of striatal dopamine terminals, substantia nigra pars compacta loss of dopamine neurons, and motor-behavior defects. Thus it appears that the potent effects of the short fibrillar fragments can be attributed to their ability to recruit monomeric α-synuclein and spread in vivo. It is also possible that the fibrils themselves are inert and possibly even protective, sequestering the toxic, α-synuclein-derived oligomeric species [46].

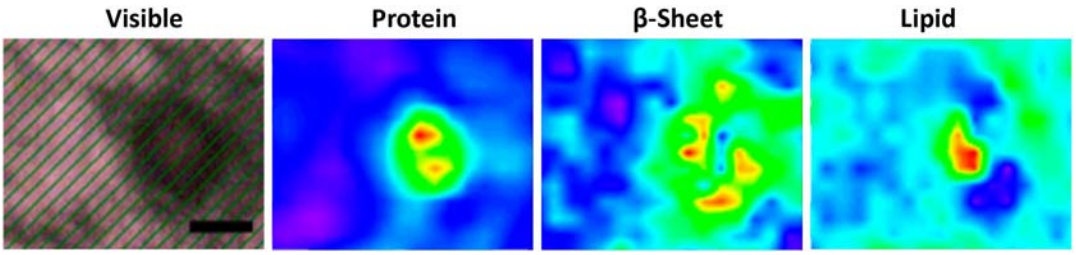


Figure 5.6
Visible and Fourier transform infrared spectroscopy (FTIR) images of a Lewy body (LB) in the substantia nigra of the midbrain derived from a Parkinson's disease (PD) patient. Shown from left to right are a visible microscope image, the total protein distribution, the proportion of β-sheet structures, and lipid distribution. The color bar indicates low (*blue*) to high (*red*) contents. Scale bar: 10 μm. Source*: From K. Araki, N. Yagi, Y. Ikemoto, H. Yagi, C.J. Choong, H. Hayakawa, et al., Synchrotron FTIR micro-spectroscopy for structural analysis of Lewy bodies in the brain of Parkinson's disease patients, Sci. Rep. 5 (2015) 17625.*

The α-synuclein structure in LBs has also been imaged in human autopsy tissue [48]. Synchrotron Fourier transform infrared microspectroscopy revealed a shift in the infrared spectrum within the LBs that indicates abundance of a β-sheet-rich structures. Also, two-dimensionl (2D) infrared mapping of LBs revealed that the content of the β-sheet structure is higher in the halo than in the core, and the core contains a large amount of proteins and lipids (Fig. 5.6). This is consistent with earlier studies that showed LBs are composed primarily of α-synuclein, but they also include other proteins such as p62 and ubiquitin [46]. But it is contrasted with parenchymal Aβ plaques in AD, where the dense core plaques are Aβ42 with a lipid membrane-like spectral signature in the periphery [15,33]

5.2.4 Amyotrophic lateral sclerosis

ALS is the most common motor neuron disease, with a lifetime risk of 1 in 2000 [49]. It is characterized by motor neuron degeneration that leads to muscle weakness followed by progressive paralysis and eventually respiratory failure. ALS is generally fatal within 3–5 years from the onset of symptoms [50,51]. The disease primarily affects the spinal cord and, at later stages, the brainstem. Approximately 90% of ALS cases are sporadic with no known cause, but the remaining 10% have a genetic link.

One familial form of the disease comes from mutations in the protein, copper–zinc superoxide (SOD1), an antioxidant protein responsible for the elimination of superoxide (O_2^-) from cells in the body. While aggregates of mutant SOD1 are known to be toxic to the motor neurons of patients with the disease, the mechanism behind the aggregation process and the cause of cell death is not clearly understood. In vitro studies have produced conflicting evidence for proposed SOD1 aggregation pathways. One explanation proposes a

direct aggregation pathway from nascent, unfolded protein [52], while others show evidence for an unfolding process from fully folded or partially folded SOD1 [53].

SOD1 is a is a β-barrel homodimeric protein composed of eight antiparallel β-sheets, where each monomer contains a disulfide bridge as well as a copper/zinc binuclear site. SOD1 aggregates from transgenic mouse models tend to be metal-deficient and/or lack the disulfide bond [54]. Changes in tertiary structure during unfolding and aggregation events are reflected in IR vibrational modes at 1695 cm^{-1} for antiparallel β-sheets, at 1645 cm^{-1} for unordered structure, and at 1630 cm^{-1} for a combination band from both parallel and antiparallel β-sheets [5].

Since the cellular environment can affect protein folding and aggregation differently than in vitro models, obtaining in vivo evidence can greatly clarify the true pathway. A time-resolved synchrotron FTIR imaging study was performed on a cell culture model of SOD1-ALS, where yellow fluorescent protein-tagged mutant SOD1 was transiently transfected in Chinese Hamster ovary (CHO) cells and the secondary structure of the SOD1 was probed in space and time [16]. For the D125H mutant, results showed a decrease in antiparallel β-sheet content, an increase in the combination band, and an increase in unordered structural motifs, suggesting the formation of protein fibrils connected by parallel β-sheet linkages (Fig. 5.7). Interestingly, however, the misfolding pathway was different for several other mutations.

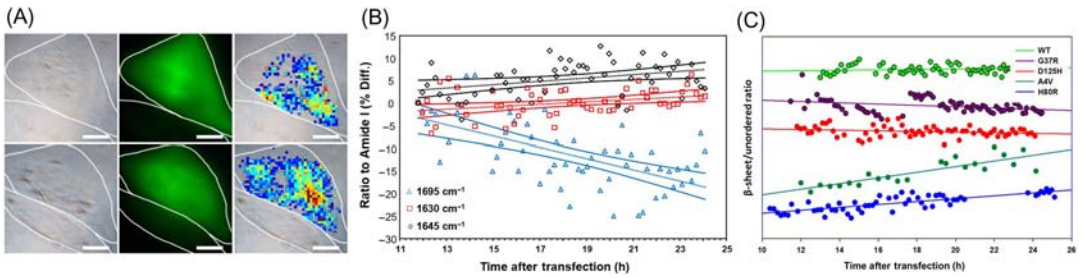


Figure 5.7

(A) D125H aggregation process at 11.5 and 24.5 h (*top* and *bottom rows*, respectively). The visible, fluorescence, and antiparallel/parallel β-sheet ratio (1695/1630 cm^{-1}) map are shown left to right. Scale bar is 10 μm. (B) Ratios of the 1695, 1645, and 1630 cm^{-1} bands to the amide I area were generated and percent differences from the original time point were plotted as a function of time for the D125H mutant. Results show a decrease in antiparallel β-sheet content, an increase in the combination band, and an increase in unordered structural motifs suggesting the formation of protein fibrils connected by parallel β-sheet linkages for the D125H mutant. (C) Ratio of β-sheet/unordered structure as a function of time for wild-type (WT) and four superoxide dismutase (SOD1) mutants, demonstrating that the misfolding pathway is not the same for all mutations. Source*: From P. Gelfand, R.J. Smith, E. Stavitski, D.R. Borchelt, L.M. Miller, Characterization of protein structural changes in living cells using time-lapsed FTIR imaging, Anal. Chem. 87 (12) (2015) 6025–6031 [16].*

FTIR spectroscopy was also used to demonstrate that the aggregation pathway for SOD1 is also dramatically affected by the presence of Ca^{2+}, where Ca^{2+} diverts SOD1 aggregation from fibrils toward amorphous aggregates [55]. The FTIR spectrum of SOD1 in the presence of Ca^{2+} showed a high content of intermolecular β-sheets (band at 1612 cm^{-1}) as well as β-turns (bands at 1678 and 1664 cm^{-1}) and a decrease in random structures (1655 cm^{-1}). This morphology has been observed in ALS-derived protein inclusions, suggesting that calcium overload may be a toxic pathway for SOD1 misfolding in the disease [55].

Aggregate formation has also been shown to be altered with curcumin, where fibrillation of SOD1 is inhibited in favor of the formation of smaller and disordered aggregates of SOD1 [56]. Docking studies show that curcumin binds to the amyloidogenic regions of SOD1 and also reduces SOD1-mediated toxicity, suggesting a possible mechanism for treatment of ALS.

Evidence of other possible metabolic imbalances have also been observed with FTIR in ALS. Specifically, small deposits of crystalline creatine have been observed with synchrotron FTIR microspectroscopy in the spinal cord, brainstem, and motor neuron cortex in human autopsy cases [57]. Creatine is used as an energy buffer in the muscle and the brain, and presence of creatine crystals in the neuronal tissue may be an indicator of dysfunctional oxidative processes in some ALS cases.

Further evidence of metabolism imbalance in ALS comes from another synchrotron FTIR imaging study that examined brainstem cross sections from the transgenic SOD1-G93A rat model of ALS. Specifically, an increase in antiparallel β-sheet was observed, most notably in the nucleus of the brainstem, consistent with the formation of SOD1 aggregates in the tissue [58]. Additional compositional changes were observed, including decreases in the lipids, choline, and phosphates in the same anatomical region. The decrease in lipids may arise from the high rate of lipid metabolism in ALS, which leads to a high cell membrane turnover. Reductions in choline and phosphate are consistent with the alterations in energy metabolism as observed in the human studies [57].

5.2.5 Prion diseases

Prion diseases are fatal neurodegenerative diseases in which the normal prion protein, cellular prion protein (PrPC), undergoes a conformational change to the misfolded, pathogenic form, scrapie isoform of the prion protein (PrPSc). The physiological role of PrPC is unknown, but it has been suggested to have many functions [59]. For example, it can bind multiple copper ions through its histidine and glycine residues in the octarepeat suggesting it may be a copper transport protein. There is also evidence that PrPC plays numerous roles in other cellular processes such as being a proapoptotic protein or

antioxidant. PrPC is also likely to be involved in transmembrane signaling for neuronal survival and neurite outgrowth.

The human prion protein consists of a globular domain with three α-helices and a two-strand antiparallel β-sheet [60]. A glycophosphatidylinositol (GPI) membrane anchor at the C-terminus tethers prion protein (PrP) to cell membranes, which is integral for conformational change. FTIR spectroscopy has demonstrated that PrPC has a high α-helical content (42%) and no β-sheet (3%). In contrast, the β-sheet and α-helix content of PrPSc is 43% and 30%, respectively [61].

The conformational transition from PrPC to the pathogenic form involves the unfolding of one or more α-helices of PrPC followed by refolding into β-sheets. FTIR spectroscopy was used to examine the secondary structure of a series of fragments, showing that PrP residues 109−122 (H1), which are predicted to be α-helical, form β-sheets in aqueous buffers, where the longer peptide 104−122 (104H1) and also peptide 129−141 (H2) have coil or α-helical structures in solution. Both 104H1 and H2 convert into β-sheets upon interaction with H1 [62]. Further, ^{13}C isotope-edited FTIR studies of the H1 peptide showed that all strands form a β-sheet from the hydrophobic core (112−122), where residue 117 aligns in all strands [63]. Kinetic studies with a ^{13}C-labeled carbonyl at residue 117 reveal nonexponential kinetics with a concentration dependence [64,65]. This alignment initiates the transition from a disordered sheet by the repeated detachment and annealing of strands in solution to promote aggregate nucleation and propagation.

Most structural studies examining the PrPC to PrPSc conversion involve the recombinant PrPC expressed in *Escherichia coli*. However, eukaryotic PrPC is posttranslationally N-glycosylated with a GPI anchor to attach PrPC to the cell membrane. Time-resolved FTIR studies of the eukaryotic PrPC tethered to a raft-like lipid membrane showed that excessive membrane anchoring induces misfolding of the PrP to intermolecular β-sheets (Fig. 5.8) [66]. Once PrPSc oligomers are formed, they can disrupt membrane structural integrity, which is a first indicator of toxic function. Soluble oligomers have been shown to form various sizes, where 36-mers form fibrils whereas 12-mers remain in solution with monomers [67]. The on-pathway (36-mer) oligomers become inserted within the lipid bilayer and ATR-FTIR showed that they have a completely different secondary structure, with a higher content in intermolecular β-sheets, showing the multimeric state. In contrast, the off-pathway (12-mer) oligomers remain at the surface of the bilayer and a calcein-leakage assay showed that these soluble oligomers are the most potent in producing leakage with negatively charged vesicles [68]. Hence PrPSc can interact with the membrane in different ways where one operates as a surface-acting agent more than a pore-forming species, whereas the other inserts in the lipid bilayer, both of them having a presumed role in neurotoxicity.

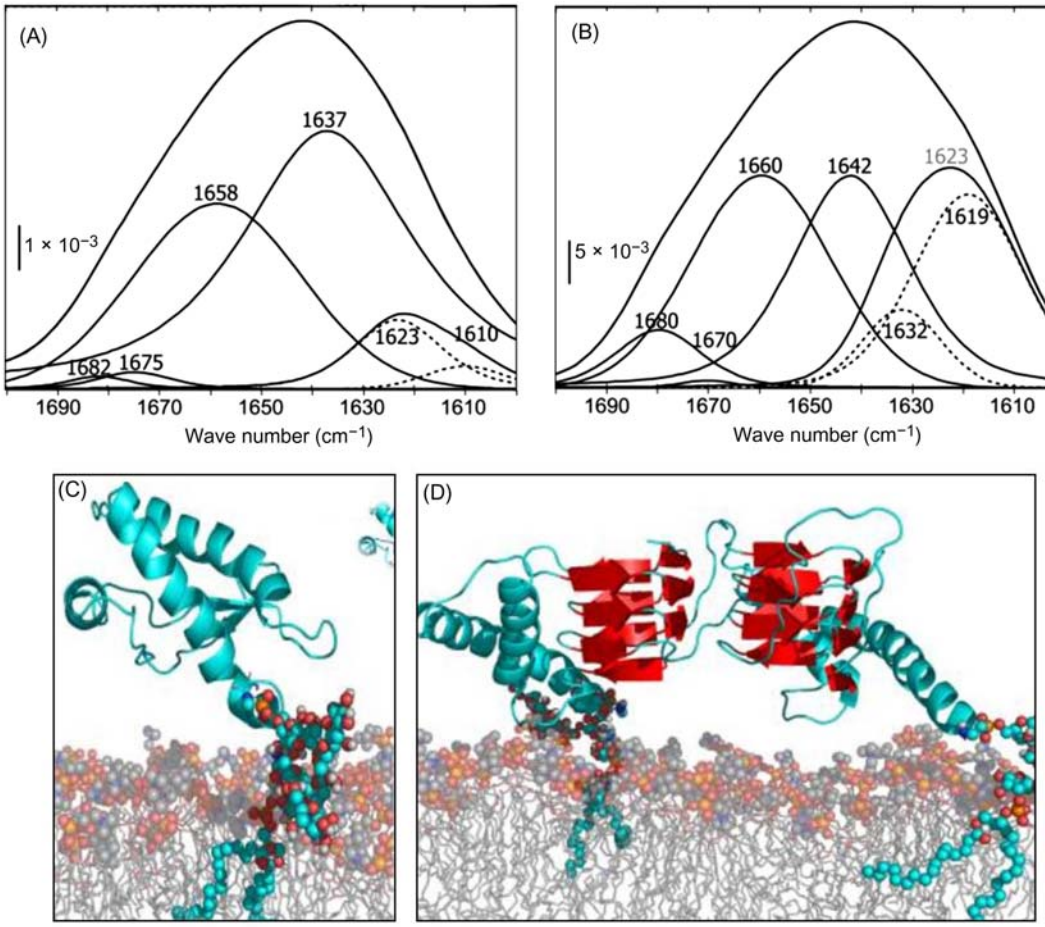


Figure 5.8

The amide I bands of cellular prion protein (PrPC) shift upon membrane anchoring. (A) The spectral deconvolution of the amide I band at 1.0 μM PrPC after 3 min incubation with the lipid. (B) After 34 min incubation time, substantial band shifts toward oligomeric β-structure were evident. Bands were assigned to antiparallel β-sheets, 1676–1663 cm^{-1} to β-turns, 1662–1645 cm^{-1} to α-helices, 1644–1635 cm^{-1} to random coil, and 1634–1610 cm^{-1} to β-sheets. This last region was subdivided (*dashed bands*) at 1620 cm^{-1}, with the low-frequency component representing intermolecular β-structure. (C–D) Illustration of the concentration-dependent secondary structure changes upon membrane anchoring where PrPC bound to the raft-like lipid bilayer exhibits the same secondary structure (at lower concentrations) as anchorless PrPC in solution (C), whereas an increased concentration of PrPC at the membrane leads to a structural transition toward intermolecular β-sheet (D). This dimerization could be the initial step on the pathway of the conversion into scrapie isoform of the prion protein (PrPSc). Source*: From K. Elfrink, J. Ollesch, J. Stohr, D. Willbold, D. Riesner, K. Gerwert, Structural changes of membrane-anchored native PrP(C), Proc. Natl Acad. Sci. U.S.A. 105 (31) (2008) 10815–10819 [66].*

While most FTIR studies on PrPC have been done in vitro, FTIR imaging has also been used to examine PrPSc deposition in terminal disease hamster dorsal root ganglia [69]. Results showed that some neurons exhibited an increased amount of β-sheet at narrow regions in or around the infected neuronal cell membranes. Based on these FTIR findings, and by comparison with immunostained adjacent sections, it was concluded that the spectral signals indicating elevated β-sheet content are due to PrPSc deposition.

5.3 Clinical imaging and diagnosis

Most research involving FTIR spectroscopy to examine neurodegenerative diseases involves gaining an understanding of the fundamental chemical and structural changes that occur in the disease. However, given the sensitivity of FTIR to chemical composition, it has the potential to be used as an analytical assay for disease screening and diagnosis. Advantages of the technique include low cost, minimal sample preparation, and the nondestructive nature of the analysis. However, challenges exist with detection sensitivity, repeatability, and specificity in order to validate the methods in large-scale studies and across research groups. With the development of fast and sensitive instruments, along with standardized sample preparation and sophisticated machine learning approaches, the method is moving closer to clinical implementation [70,71].

Both FTIR spectroscopy and microspectroscopy have been used widely for the diagnosis of diseases. Perhaps the most widely used application involves FTIR imaging of biopsied tissue for the diagnosis of cancer [72–76]. But traditional FTIR spectroscopy has also been used on bodily fluids for diagnostic purposes, such as blood tests for cancer diagnosis [77], blood cultures for pathogen identification [78], cerebrospinal fluid for multiple sclerosis and ALS [79] and PD [80], and blood plasma and white blood cells (WBCs) for AD and other dementias [70,81–84]. In all cases, proof-of-principle studies have been performed but commercialized tests do not yet exist.

In the field of neurodegeneration, analysis of bodily fluid has shown some success. For example, blood plasma samples were collected from ~500 individuals with various neurodegenerative diseases and age-matched controls. AD ($n = 164$) was identified with 70% sensitivity and specificity. Early AD cases were identified with 80% sensitivity and 74% specificity and segregation of AD from dementia with LBs was achieved with 90% sensitivity and specificity [83]. Another FTIR study using WBCs and plasma from AD patients showed it was possible to differentiate among mild, moderate, and severe AD, and the controls, with 85% accuracy when using the WBC spectra and about 77% when using the plasma spectra. When only the moderate and severe stages were included, an 83% accuracy was obtained using the WBC spectra and about 89% when using the plasma spectra [84]. And an FTIR study of blood serum from patients with multiple sclerosis and ALS showed high specificity for multiple sclerosis (100%) and ALS (98%) when compared to healthy controls [79].

Clearly FTIR has shown promise in diagnosing disease through tissue or bodily fluid analysis. However, in order to be most beneficial, it is important to assign and understand the biomarkers leading to spectral differences that distinguish the disease state [70]. For neurodegenerative diseases, diagnosis is typically done based on cognitive measures of impairment and not laboratory testing, and an autopsy is necessary to confirm the diagnosis. The pathogenesis of AD and other neurodegenerative diseases involve many chemical changes in cells and tissue that can potentially be detected by FTIR, including amyloid formation, oxidative stress, inflammation, and lipid dysregulation [85]. Studies on cerebral spinal fluid have shown a significant downshift at the amide I region in AD [86]. Since cerebrospinal fluid sampling can be an invasive procedure, more studies have been performed on human blood plasma. Here, results have suggested that an increase in oxidative stress leads to an increase in R−OH bonds and a decrease in both CH and NH bonds [87]. More specifically, a recent study has shown a decrease in lipids in AD, which was attributed to damaged cell membranes caused by free radicals. A decrease in the DNA−RNA region was also observed, which was attributed to the loss of phosphoric acid and strand breaks also caused by hydroxyl radicals [83].

In the end, it is important to emphasize that spectral changes arising from neurodegeneration are very small and multifactorial. Hence, it is necessary to standardize protocols for sample preparation, instrument calibration, and data collection. With this in hand, different sources of variation can be normalized into a single model using computational-based methods, such as direct standardization and piecewise direct standardization [71]. Once achieved, the true value of the spectral data for harvesting changes in chemical composition associated with disease can be harvested using a wide range of multivariate techniques in order to successfully adopt FTIR spectroscopy for clinical diagnosis.

References

[1] M. Jackson, H.H. Mantsch, The use and misuse of FTIR spectroscopy in the determination of protein structure, Crit. Rev. Biochem. Mol. Biol. 30 (2) (1995) 95−120.
[2] L.M. Miller, P. Dumas, Chemical imaging of biological tissue with synchrotron infrared light, Biochim. Biophys. Acta 1758 (7) (2006) 846−857.
[3] B. Shivu, S. Seshadri, J. Li, K.A. Oberg, V.N. Uversky, A.L. Fink, Distinct beta-sheet structure in protein aggregates determined by ATR-FTIR spectroscopy, Biochemistry 52 (31) (2013) 5176−5183.
[4] P. Juszczyk, A.S. Kolodziejczyk, Z. Grzonka, FTIR spectroscopic studies on aggregation process of the beta-amyloid 11-28 fragment and its variants, J. Pept. Sci. 15 (1) (2009) 23−29.
[5] E. Cerf, R. Sarroukh, S. Tamamizu-Kato, L. Breydo, S. Derclaye, Y.F. Dufrene, et al., Antiparallel beta-sheet: a signature structure of the oligomeric amyloid beta-peptide, Biochem. J. 421 (3) (2009) 415−423.
[6] F. Xu, Z. Fu, S. Dass, A.E. Kotarba, J. Davis, S.O. Smith, et al., Cerebral vascular amyloid seeds drive amyloid beta-protein fibril assembly with a distinct anti-parallel structure, Nat. Commun. 7 (2016) 13527.
[7] L.M. Miller, M.W. Bourassa, R.J. Smith, FTIR spectroscopic imaging of protein aggregation in living cells, Biochim. Biophys. Acta 1828 (10) (2013) 2339−2346.

[8] R. Tycko, Progress towards a molecular-level structural understanding of amyloid fibrils, Curr. Opin. Struct. Biol. 14 (1) (2004) 96–103.
[9] C. Paul, P.H. Axelsen, Beta Sheet structure in amyloid beta fibrils and vibrational dipolar coupling, J. Am. Chem. Soc. 127 (16) (2005) 5754–5755.
[10] L.M. Miller, Q. Wang, T.P. Telivala, R.J. Smith, A. Lanzirotti, J. Miklossy, Synchrotron-based infrared and X-ray imaging shows focalized accumulation of Cu and Zn co-localized with beta-amyloid deposits in Alzheimer's disease, J. Struct. Biol. 155 (1) (2006) 30–37.
[11] L.M. Miller, G.D. Smith, G.L. Carr, Synchrotron-based biological microspectroscopy: from the mid-infrared through the far-infrared regimes, J. Biol. Phys. 29 (2–3) (2003) 219–230.
[12] M.J. Nasse, M.J. Walsh, E.C. Mattson, R. Reininger, A. Kajdacsy-Balla, V. Macias, et al., High-resolution Fourier-transform infrared chemical imaging with multiple synchrotron beams, Nat. Methods 8 (5) (2011) 413–416.
[13] E. Stavitski, R.J. Smith, M.W. Bourassa, A.S. Acerbo, G.L. Carr, L.M. Miller, Dynamic full-field infrared imaging with multiple synchrotron beams, Anal. Chem. 85 (7) (2013) 3599–3605.
[14] M.Z. Kastyak-Ibrahim, M.J. Nasse, M. Rak, C. Hirschmugl, M.R. Del Bigio, B.C. Albensi, et al., Biochemical label-free tissue imaging with subcellular-resolution synchrotron FTIR with focal plane array detector, Neuroimage 60 (1) (2012) 376–383.
[15] C.R. Liao, M. Rak, J. Lund, M. Unger, E. Platt, B.C. Albensi, et al., Synchrotron FTIR reveals lipid around and within amyloid plaques in transgenic mice and Alzheimer's disease brain, Analyst 138 (14) (2013) 3991–3997.
[16] P. Gelfand, R.J. Smith, E. Stavitski, D.R. Borchelt, L.M. Miller, Characterization of protein structural changes in living cells using time-lapsed FTIR imaging, Anal. Chem. 87 (12) (2015) 6025–6031.
[17] B. Caughey, P.T. Lansbury, PROTOFIBRILS, PORES, FIBRILS, ANDNEURODEGENERATION: separating the responsible protein aggregates from the innocent bystanders, Annu. Rev. Neurosci. 26 (1) (2003) 267–298.
[18] Y.C. Yong, Y.Z. Wang, J.J. Zhong, Nano-spectroscopic imaging of proteins with near-field scanning optical microscopy (NSOM), Curr. Opin. Biotechnol. 54 (2018) 106–113.
[19] N. Qin, S. Zhang, J. Jiang, S.G. Corder, Z. Qian, Z. Zhou, et al., Nanoscale probing of electron-regulated structural transitions in silk proteins by near-field IR imaging and nano-spectroscopy, Nat. Commun. 7 (2016) 13079.
[20] I. Amenabar, S. Poly, W. Nuansing, E.H. Hubrich, A.A. Govyadinov, F. Huth, et al., Structural analysis and mapping of individual protein complexes by infrared nanospectroscopy, Nat. Commun. 4 (2013) 2890.
[21] M.P. Mattson, Pathways towards and away from Alzheimer's disease, Nature 430 (7000) (2004) 631–639.
[22] A. Alzheimer, A new disease of the cortex (Ger), Allg. Z. Psychiatr. 64 (1907) 146–148.
[23] M.P. Mattson, Cellular actions of beta-amyloid precursor protein and its soluble and fibrillogenic derivatives, Physiol. Rev. 77 (4) (1997) 1081–1132.
[24] M. Obulesu, R. Venu, R. Somashekhar, Tau mediated neurodegeneration: an insight into Alzheimer's disease pathology, Neurochem. Res. 36 (8) (2011) 1329–1335.
[25] P.E. Fraser, J.T. Nguyen, H. Inouye, W.K. Surewicz, D.J. Selkoe, M.B. Podlisny, et al., Fibril formation by primate, rodent, and Dutch-hemorrhagic analogues of Alzheimer amyloid beta-protein, Biochemistry 31 (44) (1992) 10716–10723.
[26] P.E. Fraser, J.T. Nguyen, W.K. Surewicz, D.A. Kirschner, pH-dependent structural transitions of Alzheimer amyloid peptides, Biophys. J. 60 (5) (1991) 1190–1201.
[27] Z. Szabo, E. Klement, K. Jost, M. Zarandi, K. Soos, B. Penke, An FT-IR study of the beta-amyloid conformation: standardization of aggregation grade, Biochem. Biophys. Res. Commun. 265 (2) (1999) 297–300.
[28] R. Tycko, Solid-state NMR studies of amyloid fibril structure, Annu. Rev. Phys. Chem. 62 (2011) 279–299.

[29] Z. Fu, D. Aucoin, J. Davis, W.E. Van Nostrand, S.O. Smith, Mechanism of nucleated conformational conversion of Abeta42, Biochemistry 54 (27) (2015) 4197−4207.
[30] W. Qiang, W.M. Yau, Y. Luo, M.P. Mattson, R. Tycko, Antiparallel beta-sheet architecture in Iowa-mutant beta-amyloid fibrils, Proc. Natl Acad. Sci. U.S.A. 109 (12) (2012) 4443−4448.
[31] L.P. Choo, D.L. Wetzel, W.C. Halliday, M. Jackson, S.M. LeVine, H.H. Mantsch, In situ characterization of beta-amyloid in Alzheimer's diseased tissue by synchrotron Fourier transform infrared microspectroscopy, Biophys. J. 71 (4) (1996) 1672−1679.
[32] M. Rak, M.R. Del Bigio, S. Mai, D. Westaway, K. Gough, Dense-core and diffuse Abeta plaques in TgCRND8 mice studied with synchrotron FTIR microspectroscopy, Biopolymers 87 (4) (2007) 207−217.
[33] A. Kuzyk, M. Kastyak, V. Agrawal, M. Gallant, G. Sivakumar, M. Rak, et al., Association among amyloid plaque, lipid, and creatine in hippocampus of TgCRND8 mouse model for Alzheimer disease, J. Biol. Chem. 285 (41) (2010) 31202−31207.
[34] F. Palombo, F. Tamagnini, J.C.G. Jeynes, S. Mattana, I. Swift, J. Nallala, et al., Detection of Abeta plaque-associated astrogliosis in Alzheimer's disease brain by spectroscopic imaging and immunohistochemistry, Analyst 143 (4) (2018) 850−857.
[35] A.C. Leskovjan, A. Kretlow, L.M. Miller, Fourier transform infrared imaging showing reduced unsaturated lipid content in the hippocampus of a mouse model of Alzheimer's disease, Anal. Chem. 82 (7) (2010) 2711−2716.
[36] M. Gallant, M. Rak, A. Szeghalmi, M.R. Del Bigio, D. Westaway, J. Yang, et al., Focally elevated creatine detected in amyloid precursor protein (APP) transgenic mice and Alzheimer disease brain tissue, J. Biol. Chem. 281 (1) (2006) 5−8.
[37] T.C. Tang, Y. Hu, P. Kienlen-Campard, L. El Haylani, M. Decock, J. Van Hees, et al., Conformational changes induced by the A21G Flemish mutation in the amyloid precursor protein lead to increased Abeta production, Structure 22 (3) (2014) 387−396.
[38] M. Sadqi, F. Hernandez, U. Pan, M. Perez, M.D. Schaeberle, J. Avila, et al., Alpha-helix structure in Alzheimer's disease aggregates of tau-protein, Biochemistry 41 (22) (2002) 7150−7155.
[39] M. von Bergen, S. Barghorn, J. Biernat, E.M. Mandelkow, E. Mandelkow, Tau aggregation is driven by a transition from random coil to beta sheet structure, Biochim. Biophys. Acta 1739 (2−3) (2005) 158−166.
[40] W.J. Goux, The conformations of filamentous and soluble tau associated with Alzheimer paired helical filaments, Biochemistry 41 (46) (2002) 13798−13806.
[41] M. von Bergen, P. Friedhoff, J. Biernat, J. Heberle, E.M. Mandelkow, E. Mandelkow, Assembly of tau protein into Alzheimer paired helical filaments depends on a local sequence motif ((306)VQIVYK(311)) forming beta structure, Proc. Natl Acad. Sci. U.S.A. 97 (10) (2000) 5129−5134.
[42] O. Schweers, E. Schonbrunn-Hanebeck, A. Marx, E. Mandelkow, Structural studies of tau protein and Alzheimer paired helical filaments show no evidence for beta-structure, J. Biol. Chem. 269 (39) (1994) 24290−24297.
[43] K. Murakami, K. Irie, A. Morimoto, H. Ohigashi, M. Shindo, M. Nagao, et al., Neurotoxicity and physicochemical properties of Abeta mutant peptides from cerebral amyloid angiopathy: implication for the pathogenesis of cerebral amyloid angiopathy and Alzheimer's disease, J. Biol. Chem. 278 (46) (2003) 46179−46187.
[44] R. Tycko, K.L. Sciarretta, J.P. Orgel, S.C. Meredith, Evidence for novel beta-sheet structures in Iowa mutant beta-amyloid fibrils, Biochemistry 48 (26) (2009) 6072−6084.
[45] J.C. Rochet, K.A. Conway, P.T. Lansbury Jr., Inhibition of fibrillization and accumulation of prefibrillar oligomers in mixtures of human and mouse alpha-synuclein, Biochemistry 39 (35) (2000) 10619−10626.
[46] J.M. Froula, M. Castellana-Cruz, N.M. Anabtawi, J.D. Camino, S.W. Chen, D.R. Thrasher, et al., Defining alpha-synuclein species responsible for Parkinson's disease phenotypes in mice, J. Biol. Chem. 294 (27) (2019) 10392−10406.
[47] N.P. Ulrih, C.H. Barry, A.L. Fink, Impact of Tyr to Ala mutations on alpha-synuclein fibrillation and structural properties, Biochim. Biophys. Acta 1782 (10) (2008) 581−585.

[48] K. Araki, N. Yagi, Y. Ikemoto, H. Yagi, C.J. Choong, H. Hayakawa, et al., Synchrotron FTIR micro-spectroscopy for structural analysis of Lewy bodies in the brain of Parkinson's disease patients, Sci. Rep. 5 (2015) 17625.
[49] J.S. Valentine, P.A. Doucette, S. Zittin Potter, Copper-zinc superoxide dismutase and amyotrophic lateral sclerosis, Annu. Rev. Biochem. 74 (2005) 563–593.
[50] D. Baumer, D. Hilton, S.M. Paine, M.R. Turner, J. Lowe, K. Talbot, et al., Juvenile ALS with basophilic inclusions is a FUS proteinopathy with FUS mutations, Neurology 75 (7) (2010) 611–618.
[51] Y. Tanaka, N. Yoshikura, N. Harada, M. Yamada, A. Koumura, T. Sakurai, et al., Late-onset patients with sporadic amyotrophic lateral sclerosis in Japan have a higher progression rate of ALSFRS-R at the time of diagnosis, Intern. Med. 51 (6) (2012) 579–584.
[52] B.F. Shaw, H.L. Lelie, A. Durazo, A.M. Nersissian, G. Xu, P.K. Chan, et al., Detergent-insoluble aggregates associated with amyotrophic lateral sclerosis in transgenic mice contain primarily full-length, unmodified superoxide dismutase-1, J. Biol. Chem. 283 (13) (2008) 8340–8350.
[53] V.K. Mulligan, A. Kerman, R.C. Laister, P.R. Sharda, P.E. Arslan, A. Chakrabartty, Early steps in oxidation-induced SOD1 misfolding: implications for non-amyloid protein aggregation in familial ALS, J. Mol. Biol. 421 (4–5) (2012) 631–652.
[54] H.L. Lelie, A. Liba, M.W. Bourassa, M. Chattopadhyay, P.K. Chan, E.B. Gralla, et al., Copper and zinc metallation status of copper-zinc superoxide dismutase from amyotrophic lateral sclerosis transgenic mice, J. Biol. Chem. 286 (4) (2011) 2795–2806.
[55] S.S. Leal, I. Cardoso, J.S. Valentine, C.M. Gomes, Calcium ions promote superoxide dismutase 1 (SOD1) aggregation into non-fibrillar amyloid: a link to toxic effects of calcium overload in amyotrophic lateral sclerosis (ALS)? J. Biol. Chem. 288 (35) (2013) 25219–25228.
[56] N.K. Bhatia, A. Srivastava, N. Katyal, N. Jain, M.A. Khan, B. Kundu, et al., Curcumin binds to the pre-fibrillar aggregates of Cu/Zn superoxide dismutase (SOD1) and alters its amyloidogenic pathway resulting in reduced cytotoxicity, Biochim. Biophys. Acta 1854 (5) (2015) 426–436.
[57] M.Z. Kastyak, M. Szczerbowska-Boruchowska, D. Adamek, B. Tomik, M. Lankosz, K.M. Gough, Pigmented creatine deposits in amyotrophic lateral sclerosis central nervous system tissues identified by synchrotron Fourier Transform Infrared microspectroscopy and X-ray fluorescence spectromicroscopy, Neuroscience 166 (4) (2010) 1119–1128.
[58] P. Andjus, S. Stamenkovic, T. Ducic, Synchrotron radiation-based FTIR spectro-microscopy of the brainstem of the hSOD1 G93A rat model of amyotrophic lateral sclerosis, Eur. Biophys. J. 48 (5) (2019) 475–484.
[59] L. Westergard, H.M. Christensen, D.A. Harris, The cellular prion protein (PrPC): its physiological function and role in disease, Biochim. Biophys. Acta (BBA) Mol. Basis Dis. 1772 (6) (2007) 629–644.
[60] H.M. Schätzl, M.D. Costa, L. Taylor, F.E. Cohen, S.B. Prusiner, Prion protein gene variation among primates, J. Mol. Biol. 245 (4) (1995) 362–374.
[61] K.M. Pan, M. Baldwin, J. Nguyen, M. Gasset, A. Serban, D. Groth, et al., Conversion of alpha-helices into beta-sheets features in the formation of the scrapie prion proteins, Proc. Natl Acad. Sci. U.S.A. 90 (23) (1993) 10962–10966.
[62] J. Nguyen, M.A. Baldwin, F.E. Cohen, S.B. Prusiner, Prion protein peptides induce alpha-helix to beta-sheet conformational transitions, Biochemistry 34 (13) (1995) 4186–4192.
[63] R.A. Silva, W. Barber-Armstrong, S.M. Decatur, The organization and assembly of a beta-sheet formed by a prion peptide in solution: an isotope-edited FTIR study, J. Am. Chem. Soc. 125 (45) (2003) 13674–13675.
[64] S.A. Petty, S.M. Decatur, Intersheet rearrangement of polypeptides during nucleation of {beta}-sheet aggregates, Proc. Natl Acad. Sci. U.S.A. 102 (40) (2005) 14272–14277.
[65] S.A. Petty, T. Adalsteinsson, S.M. Decatur, Correlations among morphology, beta-sheet stability, and molecular structure in prion peptide aggregates, Biochemistry 44 (12) (2005) 4720–4726.
[66] K. Elfrink, J. Ollesch, J. Stohr, D. Willbold, D. Riesner, K. Gerwert, Structural changes of membrane-anchored native PrP(C), Proc. Natl Acad. Sci. U.S.A. 105 (31) (2008) 10815–10819.

[67] F. Eghiaian, T. Daubenfeld, Y. Quenet, M. van Audenhaege, A.P. Bouin, G. van der Rest, et al., Diversity in prion protein oligomerization pathways results from domain expansion as revealed by hydrogen/deuterium exchange and disulfide linkage, Proc. Natl Acad. Sci. U.S.A. 104 (18) (2007) 7414–7419.
[68] S. Combet, F. Cousin, H. Rezaei, S. Noinville, Membrane interaction of off-pathway prion oligomers and lipid-induced on-pathway intermediates during prion conversion: a clue for neurotoxicity, Biochim. Biophys. Acta Biomembr. 1861 (2) (2019) 514–523.
[69] A. Kretlow, Q. Wang, J. Kneipp, P. Lasch, M. Beekes, L. Miller, et al., FTIR-microspectroscopy of prion-infected nervous tissue, Biochim. Biophys. Acta 1758 (7) (2006) 948–959.
[70] A.L. Mitchell, K.B. Gajjar, G. Theophilou, F.L. Martin, P.L. Martin-Hirsch, Vibrational spectroscopy of biofluids for disease screening or diagnosis: translation from the laboratory to a clinical setting, J. Biophotonics 7 (3–4) (2014) 153–165.
[71] C.L.M. Morais, M. Paraskevaidi, L. Cui, N.J. Fullwood, M. Isabelle, K.M.G. Lima, et al., Standardization of complex biologically derived spectrochemical datasets, Nat. Protoc. 14 (5) (2019) 1546–1577.
[72] N. Bergner, B.F. Romeike, R. Reichart, R. Kalff, C. Krafft, J. Popp, Tumor margin identification and prediction of the primary tumor from brain metastases using FTIR imaging and support vector machines, Analyst 138 (14) (2013) 3983–3990.
[73] J. Dudala, M. Bialas, A. Surowka, M. Bereza-Buziak, A. Hubalewska-Dydejczyk, A. Budzynski, et al., Biomolecular characterization of adrenal gland tumors by means of SR-FTIR, Analyst 140 (7) (2015) 2101–2106.
[74] F. Grosserueschkamp, T. Bracht, H.C. Diehl, C. Kuepper, M. Ahrens, A. Kallenbach-Thieltges, et al., Spatial and molecular resolution of diffuse malignant mesothelioma heterogeneity by integrating label-free FTIR imaging, laser capture microdissection and proteomics, Sci. Rep. 7 (2017) 44829.
[75] G. Theophilou, K.M. Lima, P.L. Martin-Hirsch, H.F. Stringfellow, F.L. Martin, ATR-FTIR spectroscopy coupled with chemometric analysis discriminates normal, borderline and malignant ovarian tissue: classifying subtypes of human cancer, Analyst 141 (2) (2016) 585–594.
[76] M. Verdonck, N. Wald, J. Janssis, P. Yan, C. Meyer, A. Legat, et al., Breast cancer and melanoma cell line identification by FTIR imaging after formalin-fixation and paraffin-embedding, Analyst 138 (14) (2013) 4083–4091.
[77] D. Bury, C.L.M. Morais, M. Paraskevaidi, K.M. Ashton, T.P. Dawson, F.L. Martin, Spectral classification for diagnosis involving numerous pathologies in a complex clinical setting: a neuro-oncology example, Spectrochim. Acta A Mol. Biomol. Spectrosc. 206 (2019) 89–96.
[78] C. Colabella, L. Corte, L. Roscini, V. Shapaval, A. Kohler, V. Tafintseva, et al., Merging FT-IR and NGS for simultaneous phenotypic and genotypic identification of pathogenic Candida species, PLoS One 12 (12) (2017) e0188104.
[79] Y. El Khoury, N. Collongues, J. De Seze, V. Gulsari, C. Patte-Mensah, G. Marcou, et al., Serum-based differentiation between multiple sclerosis and amyotrophic lateral sclerosis by Random Forest classification of FTIR spectra, Analyst 144 (15) (2019) 4647–4652.
[80] M.N. Sonuc Karaboga, M.K. Sezginturk, Cerebrospinal fluid levels of alpha-synuclein measured using a poly-glutamic acid-modified gold nanoparticle-doped disposable neuro-biosensor system, Analyst 144 (2) (2019) 611–621.
[81] N. Braidy, A. Poljak, C. Marjo, H. Rutlidge, A. Rich, T. Jayasena, et al., Metal and complementary molecular bioimaging in Alzheimer's disease, Front. Aging Neurosci. 6 (2014) 138.
[82] J. Lopes, M. Correia, I. Martins, A.G. Henriques, I. Delgadillo, E.S.O. da Cruz, et al., FTIR and Raman spectroscopy applied to dementia diagnosis through analysis of biological fluids, J. Alzheimers Dis. 52 (3) (2016) 801–812.
[83] M. Paraskevaidi, C.L.M. Morais, K.M.G. Lima, J.S. Snowden, J.A. Saxon, A.M.T. Richardson, et al., Differential diagnosis of Alzheimer's disease using spectrochemical analysis of blood, Proc. Natl Acad. Sci. U.S.A. 114 (38) (2017) E7929–E7938.

[84] S. Mordechai, E. Shufan, B.S. Porat Katz, A. Salman, Early diagnosis of Alzheimer's disease using infrared spectroscopy of isolated blood samples followed by multivariate analyses, Analyst 142 (8) (2017) 1276–1284.
[85] L.S. Libow, Alzheimer's disease: overview of research and clinical advances, Manag. Care Interface 20 (3) (2007) 20–23.
[86] A. Nabers, J. Ollesch, J. Schartner, C. Kötting, J. Genius, H. Hafermann, et al., Amyloid-β-secondary structure distribution in cerebrospinal fluid and blood measured by an immuno-infrared-sensor: a biomarker candidate for Alzheimer's disease, Anal. Chem. 88 (5) (2016) 2755–2762.
[87] D.H. Burns, S. Rosendahl, D. Bandilla, O.C. Maes, H.M. Chertkow, H.M. Schipper, Near-infrared spectroscopy of blood plasma for diagnosis of sporadic Alzheimer's disease, J. Alzheimers Dis. 17 (2) (2009) 391–397.

CHAPTER 6

Near-infrared spectroscopy and imaging in protein research

Mika Ishigaki[1,2] and Yukihiro Ozaki[3]

[1]Raman Project Center for Medical and Biological Applications, Shimane University, Matsue, Japan, [2]Faculty of Life and Environmental Sciences, Shimane University, Matsue, Japan, [3]School of Science and Technology, Kwansei Gakuin University, Hyogo, Japan

6.1 Introduction

Near-infrared (NIR) spectroscopy deals with absorption, emission, reflection, and diffuse reflectance of NIR light in the region of 800−2500 nm (12,500−4000 cm^{-1}) [1−6]. It consists of electronic NIR spectroscopy and vibrational NIR spectroscopy. The NIR region may be divided into three regions, the 800−1200 nm (12,500−8000 cm^{-1}) region, the 1200−2000 nm (8000−5000 cm^{-1}) region, and the 2000−2500 nm (5000−4000 cm^{-1}) region. The first region is linked with visible (VIS) region, forming a so-called VIS−NIR region. This region shows particularly high permeability, providing a window for biological bodies; it has been used practically for medical applications and agricultural applications [1−6]. In this region, many bands due to electronic transitions, higher-order overtones, and combinations appear. The second region contains a number of bands arising from the first and second overtones and combination modes. This region is very useful for quantitative and qualitative analysis and applications to structural studies such as those of hydrogen bonds. The third region deals mostly with the combination modes. This region is useful for various purposes, for example, studies of secondary structures of proteins, but permeability of this region is relatively low.

Many studies of NIR electronic spectroscopy of proteins, such as oxyhemoglobin, deoxyhemoglobin, and cytochrome *c* oxidase, have been carried out in relation to medical applications using the windows of biological bodies [4,6,7]. However, in this chapter we do not describe NIR electronic spectroscopy of proteins. In the NIR vibrational spectra, bands are observed due to overtones and combinations of stretching and bending modes. Since these overtones and combinations are intrinsically forbidden according to the approximation of harmonic oscillator, the observed band intensities are very weak. Thus the NIR region

Vibrational Spectroscopy in Protein Research.
DOI: https://doi.org/10.1016/B978-0-12-818610-7.00006-2

inversely has a superiority in permeability compared to the VIS and infrared (IR) regions, and the feature of NIR light enables nondestructive and noninvasive analyses of objects [1−5].

The history of NIR spectroscopy is relatively new, and its applications was started only after the 1960s [1−5]. One of the reasons why NIR spectroscopy had not been used until then so much was a difficulty in its spectral analysis. In NIR spectra, a number of bands arising from overtones and combinations heavily overlap each other. Therefore, it is not always easy to extract useful information from the NIR spectra. The idea of introducing a statistical method for the analysis of complicated NIR spectra became a breakthrough to wake up the "Sleeping Giant." Carl Norris who was an agricultural engineer of United States Department of Agriculture (USDA) thought of this idea during his studies of nondestructive analysis of agricultural products. In the 1960s Norris et al. succeeded in monitoring water content in grain by constructing calibration models about NIR data by use of multivariate analysis. The innovative work of electronic NIR spectra of proteins in relation to medical application was the in vivo monitoring of the redox behavior of cytochrome *c* oxidase by Jöbsis in 1977 [7]. In the wake of these pioneering studies, NIR spectroscopy has been spreading to various fields such as biological and biomedical sciences, food and agricultural engineering, chemical engineering and online monitoring, polymer science and engineering, and environmental sciences [1−10]. In the past decades, new NIR techniques such as online spectrometers, handheld spectrometers, and NIR imaging systems have been developed, and they dramatically extended the applications of NIR spectroscopy [1,2,8−15]. Since NIR imaging provides much information about the spatial distribution of sample components, it has been paid special attention. In fact, much research utilizing NIR imaging was recently reported in various fields, such as biomedical, pharmaceutical, foods, polymers, and so on [12,16−38].

NIR spectroscopy has been extensively used to investigate structures and functions of biomolecules such as water, proteins, lipids, and carbohydrates [1−4]. NIR spectroscopy has the advantages of exploring hydrogen bonds, inter- and intramolecular interactions, and hydration in the biomolecules because most of the NIR bands are attributed to functional groups containing a hydrogen atom (e.g., OH, CH, and NH). Also, a band shift due to the formation of a hydrogen bonding is much larger in the NIR region compared with the IR region, which feature can be cited as a strong point to investigate structural changes of biomolecules. Wu et al. explored temperature-induced variations in the secondary structure and hydration of human serum albumin (HSA) using two-dimensional (2D) NIR correlation spectroscopy [39,40]. The linked changes of protein and water bands provided explicit evidence of hydration changes between water and proteins [39]. An investigation of C−H · · · O = C hydrogen bond in poly(3-hydroxybutyrate) (PHB), which is a biopolymer synthesized by bacteria, is another good example of NIR studies on hydrogen bonds in biomolecules [41]. For crystalline and amorphous phases of PHB, their band positions of C−H and C = O stretching modes were significantly

different from each other. The results were interpreted as the C−H and C = O bonds in the crystalline state were involved in the C−H · · · O = C hydrogen bonding [41]. In this way, NIR spectroscopy demonstrates its unique advantages in the studies of molecular structural changes associated with hydrogen bonds and hydration.

The purpose of this review is to describe the advantages and usefulness of NIR spectroscopy in protein research. To do that, characteristic features of NIR spectroscopy in protein research are described first, and then band assignments of NIR spectra due to proteins are discussed. After that, we introduce several examples of NIR studies of proteins; the concrete examples of structural changes caused by thermal denaturation are discussed for three kinds of proteins. As an example of application of NIR imaging to biology, recent studies about fish egg development are exhibited. In the end, we summarize the future prospects of NIR spectroscopy and imaging for protein research.

6.2 Applications of near-infrared spectroscopy to protein science

6.2.1 How to apply near-infrared spectroscopy to protein science

Proteins are the main components for all life and they play central roles in biological responses [42]. Proteins which work to express genetic information and metabolites produced by genetic information are also proteins. Therefore it has been proposed for decades to investigate the relationship between protein functions and its conformation. There exist several powerful methods to investigate protein structure such as nuclear magnetic resonance [43,44], X-ray crystallography [45−47], Raman spectroscopy [48−50], and so on. IR spectroscopy is also one of the valuable tools to get information about higher order structure of proteins. An IR spectrum of a protein exhibits an absorbance pattern of fundamental modes of molecular vibrations, and bands relating to secondary structures of proteins have been investigated in detail [49,51,52]. However, absorbance due to water in the IR region is extremely strong, and sample forms containing water are strictly restricted for proper measurements.

In the NIR region, bands due to overtones and combinations of molecular vibration are observed, and their weak absorbance guarantees high transparency [1−4]. Therefore NIR spectroscopy is a valuable method to measure objects such as aqueous solutions and live tissues in situ without being invasive.

The advantages of NIR spectroscopy in protein research can be summarized as follows.

1. NIR spectroscopy is very useful to investigate the hydration of proteins.
2. It enables nondestructive analysis of proteins in tissues in vivo because of its relatively small disturbance due to water absorption. For example, we succeeded in measuring NIR spectra of proteins in a fish egg in situ and even those of membrane proteins in an egg membrane taken from the fish egg.

3. One can explore a change in hydration as well as that in the protein secondary structure at the same time. IR and Raman spectroscopies can hardly investigate them simultaneously.
4. As in the case of an IR spectrum, a band shift for a particular band caused by a hydrogen bond or an interaction between molecules can be observed. The shift of the NIR band is far larger than that of the IR band.
5. Electronic NIR spectra of proteins, such as hemoglobin and cytochrome *c* oxidase, show absorption bands in the short wavelength NIR region, and optical imaging of brain function and metabolism in vivo can be performed [7,53,54].

The relationship between NIR spectra and protein structure has been investigated for decades, and it is also applied in vitro and in vivo to the studies of protein hydration and structure [39,55,56]. However, complicated spectra frequently make it difficult to assign the origin of vibrational modes. Vibrational modes due to side chains of amino acids sometimes couple with those arising from the main chain determining the peptide secondary structure, and spectral patterns show variations depending on the composition of amino acids, even those with the same higher order structure of proteins [57,58]. In order to identify NIR bands reflecting protein structures, spectral variations caused by structural changes should be explored using proteins that can take clearly different forms. Many reports have been published for NIR band assignments relating to higher order structure of proteins [8,57–60]. In the present chapter, we discuss the band assignments of NIR spectra of proteins first, and then introduce three concrete examples of NIR spectral variations [ovalbumin (OVA), β-lactoglobulin, and bovine serum albumin (BSA)] due to protein secondary structural changes by thermal denaturation while referring to other studies.

6.2.2 Near-infrared spectral analysis

Methods for NIR band assignments have been conventionally based on group frequencies, spectra–structure relationships, and comparison of the NIR spectrum of a compound with those of related compounds. Calculations of second derivatives of NIR spectra and their difference spectra are very useful In order to decompose them into some spectral components overlapping with each other. However, these conventional methods are not always sufficient to unravel complicated NIR spectra. Thus chemometrics has often been applied to the NIR dataset to extract useful information [1,2,8,9,61–68]. A main part of chemometrics is multivariate data analysis, which is essential for qualitative and quantitative assays based on NIR spectra. Besides well-known methods of multivariate data analysis, such as principal component analysis/regression (PCA/PCR) and partial least squares regression (PLSR), self-modeling curve resolution (SMCR) is also important. SMCR can predict pure component spectra and pure component concentration profiles from a set of NIR spectra [1,2,8,9,61–64]. Furthermore, 2D correlation spectroscopy is unique in

the analysis of NIR spectra in addition to the methods mentioned above [40,69]. It can capture complex spectra as 2D synchronous and asynchronous maps which disentangle overlapped NIR bands by taking into account spectral correlation. Moreover, quantum chemical calculations, such as density function theory calculations, are becoming popular in the analysis of NIR spectra [70–72]. One can calculate the intensities and frequencies of overtones and combination bands using quantum chemical calculations.

6.2.3 Near-infrared bands due to amide groups

NIR spectra of proteins and peptides show several bands due to overtones and combinations of an amide group [1–4,8]. The amide group has mainly three fundamental vibrational modes of amide I, II, and III. In the NIR region, bands due to the second overtone of amide I, and combination modes of N-H stretching and amide II or amide III modes can be observed [8,59,73,74]. Protein bands with different secondary structures have been reported to be detected at different wave numbers [8,57–60]. In order to investigate variations in the band positions depending on secondary structure of proteins, Fig. 6.1 exhibits NIR absorbance spectra of OVA, β-lactoglobulin, and BSA with different proportions of α-helix

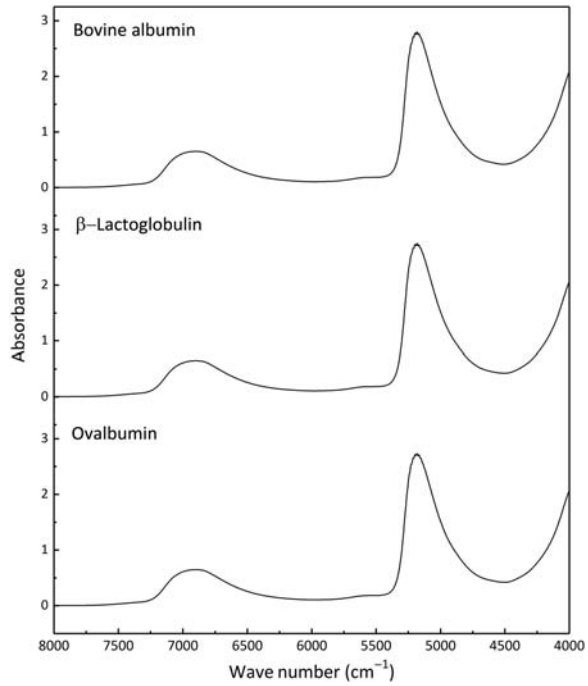


Figure 6.1
Near-infrared (NIR) spectra of protein aqueous solutions; ovalbumin (OVA), β-lactoglobulin, and bovine serum albumin (BSA).

Table 6.1: The properties of three proteins: ovalbumin (OVA), β-lactoglobulin, and bovine serum albumin (BSA).

	OVA	β-Lactoglobulin	BSA
α-Helix (%)	41	15	66
β-Sheet (%)	34	54	3
Molecular weight	45,000	18,400	66,500
Thermal denaturation temperature (°C)	78	70	62–64

and β-sheet components; β-lactoglobulin and BSA have β-sheet- and α-helix-rich structures, respectively, and OVA has a comparable ratio of them [75–83]. Table 6.1 summarizes the properties of these three proteins [75–83]. In Fig. 6.1, broad features at around 6900 and 5200 cm^{-1} in the three spectra are due to the combination mode of the antisymmetric and symmetric O–H stretching modes of water and those of the antisymmetric O–H stretching and O–H bending modes of water, respectively [8]. In order to investigate the weak absorbance bands due to proteins, second derivative spectra were calculated.

Fig. 6.2 depicts the second derivative spectra in the 4900–4200 cm^{-1} region. Peaks around 4864 cm^{-1} are assigned to the combination modes of N–H stretching vibration and amide II (amide A + II) [73,74], and those at approximately 4610 cm^{-1} are due to the combination modes of amide A and amide III (amide A + III) [59]. Their peak positions are varied for each protein. The comparison of protein NIR spectra with different ratios of secondary structure revealed that the band at around 4615 cm^{-1} may be related with an α-helix structure, and the band at around 4540 cm^{-1} with a β-sheet [57,59]. Fig. 6.2 surely demonstrates the spectral features relating to the secondary structures, and the order of the peak intensity corresponds to the proportion of β-sheet structure in Table 6.1; the peak intensity at 4540 cm^{-1} was most intense for β-lactoglobulin followed by OVA. Furthermore, the peak at 4368 cm^{-1} observed in the BSA spectrum (Fig. 6.2) was assigned to the peak relating with α-helix [57,59]. The BSA has the highest ratio of α-helix component compared to the other two samples, and thus the peak was supposed to be clearly observed. For the other two proteins, the peaks may be broadened or shifted by overlapping with the combination modes of C–H stretching and C–H bending in the 4400–4200 cm^{-1} region [8]. The peaks in the vicinity of 4256–4262 cm^{-1} did not arise from the amide modes and were due to the combination of C–H stretching and C–H bending modes [8].

6.2.4 Thermal denaturation

In order to identify NIR bands reflecting protein secondary structures, these three protein samples were heated to change the secondary structures, and temperature-dependent spectral variations were investigated. Fig. 6.3A–D compare the temperature-dependent

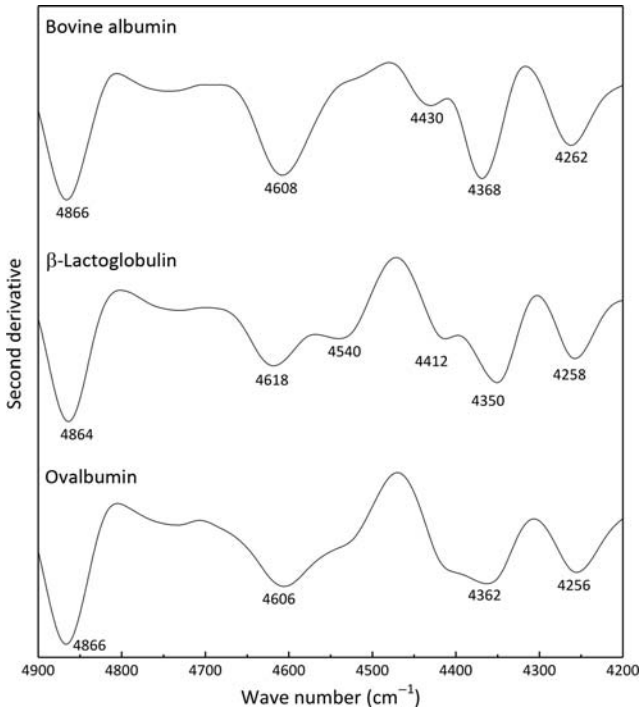


Figure 6.2
Second derivative spectra in the 4900−4200 cm^{-1} region of the spectra shown in Fig. 6.1.

changes in the second derivative spectra in the 4900−4800, 4700−4550, 4600−4450, and 4400−4300 cm^{-1} regions, respectively. These protein secondary structures are known to increase the ratio of β-sheet at the expense of α-helix by thermal denaturation [74,75], and several spectral variations with the denaturation were observed in each marker band region.

The peak at around 4860 cm^{-1} shifted to a lower wave number (Fig. 6.3A). Liu et al. reported that denatured pepsin which had originally a β-sheet-rich structure exhibited a higher frequency shift of the amide A/II mode in the NIR region in conjunction with the upward shift of amide A in the corresponding IR spectra [60]. In the structural transition of silk fibroins, the corresponding band was shifted downward from 4875 to 4860 cm^{-1} as reported by Miyazawa et al. [58]. Therefore amide A/II can be a marker band to monitor the variation of hydrogen bonds in amide groups caused by structural changes of proteins.

The peaks at around 4610 cm^{-1} have been reported to be also related to α-helix structure [57] (Fig. 6.3B), and the peak intensity of silk fibroin was decreased with thermal denaturation to be replaced by the increment of the ones at around 4520 cm^{-1} assigned to the β-sheet structure mentioned below [58]. In the present results shown in Fig. 6.3B, the peak intensities were consistently decreased in the β-lactoglobulin and BSA spectra with the

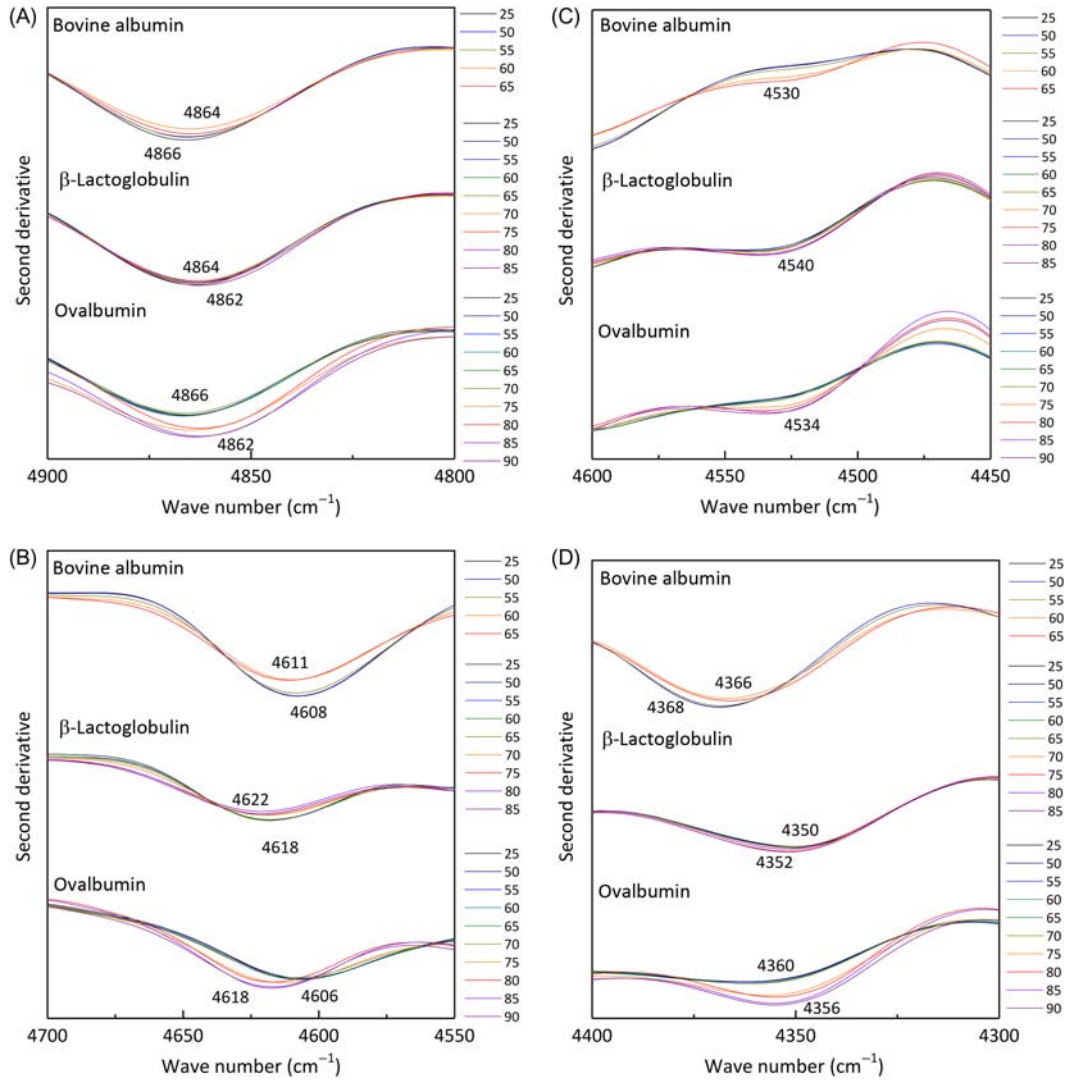


Figure 6.3
Temperature-dependent second derivative spectra variations of ovalbumin (OVA), β-lactoglobulin, and bovine serum albumin (BSA) in (A) 4900−4800 cm^{-1}, (B) 4700−4550 cm^{-1}, (C) 4600−4450 cm^{-1}, and (D) 4400−4300 cm^{-1} regions.

thermal denaturation, but in the OVA spectrum, on the other hand, the large amount of peak shifts to higher frequency were observed from 4606 to 4618 cm^{-1}. The other two samples (β-lactoglobulin and BSA) also yielded a higher frequency shift of the peaks, and it can be a possible indicator of structural changes of proteins.

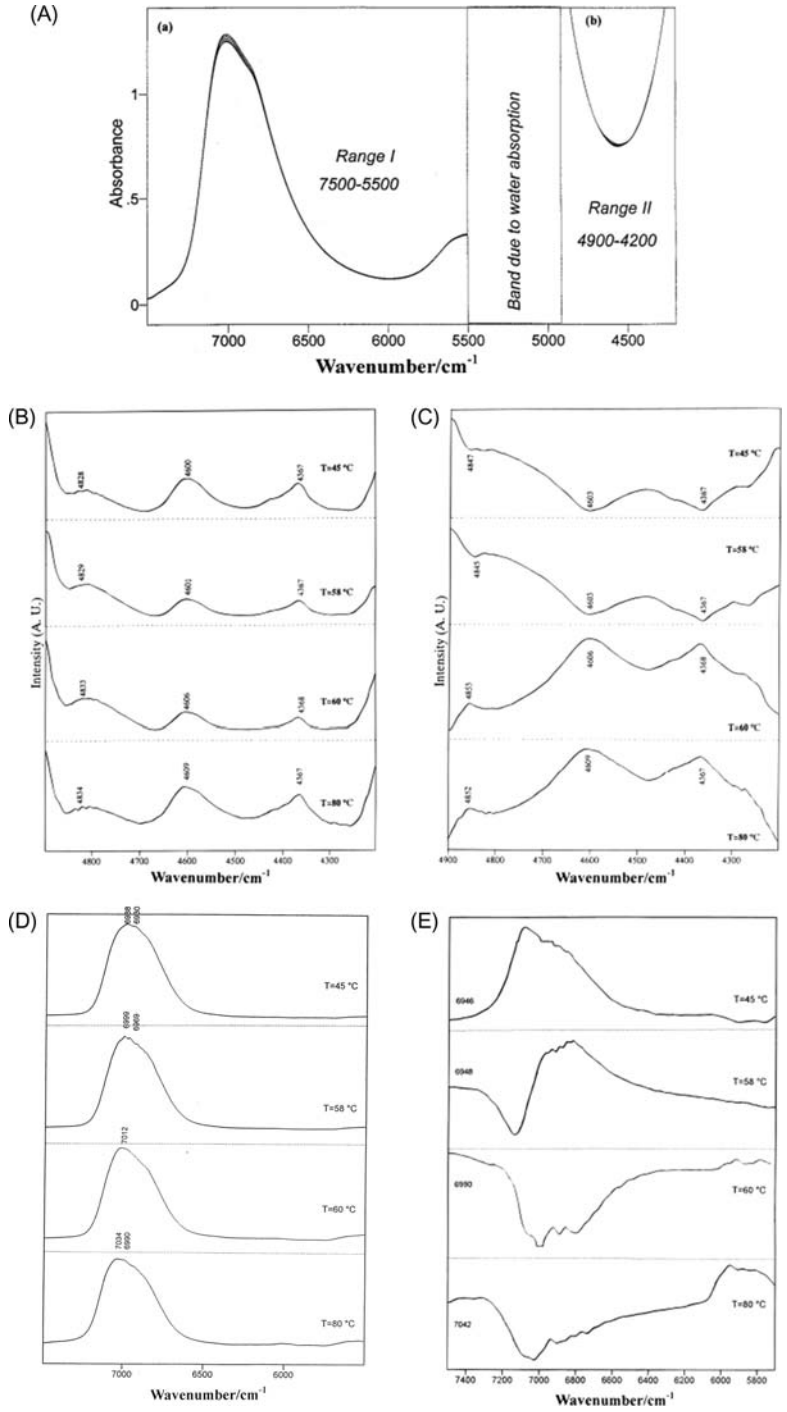
In Fig. 6.3C, the peaks at 4534 and 4530 cm^{-1} appeared in the spectra of OVA and BSA, respectively, and the peak intensity at 4540 cm^{-1} increased for β-lactoglobulin with the

increment of β-sheet components by thermal denaturation. Proteins that are rich in β-sheet structure such as concanavalin and β-lactoglobulin showed clear bands at around 4525 cm^{-1} [58]. Robert et al. also assigned the peak at around 4535 cm^{-1} to the β-sheet structure based on using 12 model proteins [59]. These results indicate that the peak at around 4525 cm^{-1} is surely related to the β-sheet structure.

The enlarged second derivative spectra in the 4400−4300 cm^{-1} region are shown in Fig. 6.3D. The peak intensity at 4368 cm^{-1} in the BSA spectrum which was supposed to be from α-helix structure [8,9,57] became weaker, and the peak position shifted from 4368 to 4366 cm^{-1}. The peak also gave a lower shift in the spectra of OVA, both of which included significant components of α-helix. These results were consistent with other reports concluding that the peak due to α-helix shifted to a lower wave number in transition to a β-sheet structure [58]. Therefore it can be confirmed that the origin of the peak is due to the α-helix.

6.2.5 Protein hydration study of human serum albumin by near-infrared spectroscopy

In the introduction, the advantage of NIR spectroscopy in the studies of hydration of biomolecules was emphasized. Here, the study of hydration changes in HSA by temperature-induced variation in the secondary structure is outlined [39]. Wu et al. measured Fourier-transform near-infrared (FT-NIR) spectra of HSA in aqueous solutions with the concentrations of 1.0, 2.0, 3.0, 4.0, and 5.0 wt.% over a temperature range of 45°C−80°C in the regions of 7500−5500 and 4900−4200 cm^{-1} (Fig. 6.4A) [39]. To investigate the correlation between the temperature-induced secondary structural changes and hydration of protein molecules, power and slice spectra were calculated, respectively, from 2D synchronous and asynchronous correlation spectra generated by the temperature-dependent spectra of HSA [40]. 2D-correlation spectroscopy (2D-COS) can decompose NIR spectra showing complex and heavily overlapping features into spectral components by spreading peaks along the second dimension [84−86]. It provides information about the order of spectral intensity changes and the value of the controlling variable affecting the spectra. The correlated variations of some bands assist to identify various inter- and intramolecular interactions. Moreover, the correlation between NIR spectra and other spectra such as Raman or IR spectra can also be analyzed. Fig. 6.4B and C depict the power and slice spectra in the 4900−4200 cm^{-1} region, respectively [39]. In the power spectra, a band at around 4600 cm^{-1} due to the combination mode of amide B and amide II (amide B/II) of HSA shows a shift by 5 cm^{-1} between 58°C and 60°C with the secondary structural changes of HSA (Fig. 6.4B). Both the power and slice spectra in the 7500−5500 cm^{-1} region also show remarkable changes near 60°C (Fig. 6.4D and E). Comparison of these temperature-dependent frequency shifts between the bands due to amide B/II mode of HSA around 4600 cm^{-1} and the combination band due to water near 7000 cm^{-1} indicates that the protein unfolding occurs almost in parallel with the change in


(A)
(a)
(b)
Absorbance
Range I
7500-5500
Band due to water absorption
Range II
4900-4200
Wavenumber/cm-1
(B)
T=45 °C
T=58 °C
T=60 °C
T=80 °C
Intensity (A. U.)
Wavenumber/cm-1
(C)
T=45 °C
T=58 °C
T=60 °C
T=80 °C
Intensity (A. U.)
Wavenumber/cm-1
(D)
T=45 °C
T=58 °C
T=60 °C
T=80 °C
Wavenumber/cm-1
(E)
T=45 °C
T=58 °C
T=60 °C
T=80 °C
Wavenumber/cm-1

(*Continued*)

the protein hydration. This study has demonstrated the usefulness of NIR spectroscopy in monitoring subtle changes in protein dynamics, especially for hydration [39].

6.2.6 Near-infrared studies of protein secondary structure

There are several other interesting NIR studies of protein secondary structure. Izutsu et al. investigated various proteins, such as BSA, lysozyme, OVA, myoglobin, and cytochrome-*c*, in aqueous solutions and those in freeze-dried solids to explore the characteristic peaks of protein secondary structure [57]. They identified NIR bands due to α-helix (4090, 4365−4370, 4615, and 5755 cm^{-1}) and β-sheet (4060, 4405, 4525−4540, 4865, and 5915−5925 cm^{-1}) structures in the combination (5000−4000 cm^{-1}) and the first overtone (6000−5600 cm^{-1}) regions [57]. Miyazawa et al. measured several globular proteins with a variety of relative amounts of protein secondary structures [58]. In conjunction with the NIR spectral variations in silk fibroin proteins with different secondary structures, the characteristic peaks of protein secondary structures in their NIR second derivative spectra were summarized [58]. There are many more interesting NIR studies on proteins. Wang et al. reported 2D-COS NIR study of heat denaturation of OVA in aqueous solutions [86]. They studied also the relation between changes in hydration and those in the secondary structure [86]. Murayama et al. investigated a hydration study of HSA in acidic pH region using NIR 2D-COS [87]. The applicability of NIR reflectance spectroscopy for probing structural changes in protein during the process of denaturation was investigated by Cho et al. [88]. Protein solubility and digestibility were indicated to be predicted using a correlated variation of absorbance changes at around 2144, 2168, and 2208 nm.

6.3 Near-infrared imaging

6.3.1 Advantages of near-infrared imaging

NIR imaging collects data in a spatially resolved manner; the data of chemical imaging, in general, can be represented as a three-dimensional (3D) cube spanning one wavelength and two

◀

Figure 6.4

(A) Fourier-transform near-infrared (FT-NIR) spectra of human serum albumin (HSA) in aqueous solutions with the concentrations of 1.0, 2.0, 3.0, 4.0, and 5.0 wt.% over a temperature range of 45−80°C. (B) Power and (C) slice spectra of the synchronous and asynchronous two-dimensional (2D) correlation spectra in the 4900−4200 cm^{-1} region calculated from the spectra shown in (A). (D) The corresponding power and (E) slice spectra in the 7500−5500 cm^{-1} region. Source: *Reproduced from Y. Wu, B. Czarnik-Matusewicz, K. Murayama, Y. Ozaki, Two-dimensional near-infrared spectroscopy study of human serum albumin in aqueous solutions: using overtones and combination modes to monitor temperature-dependent changes in the secondary structure, J. Phys. Chem. 104 (2002) 5840. with permission. Copyright (2002) American Chemical Society.*

spatial dimensions [11,12]. Therefore, novel information about the spatial distribution of chemical components can be obtained. To handle these hypercubes with a great number of data, chemometrics methods are often used. NIR imaging has the following advantages [11–15].

1. NIR imaging enables in situ, noncontact, and nondestructive analysis.
2. Weaker absorbance of water band makes it easier to analyze aqueous samples than IR imaging.
3. Since NIR light can penetrate much deeper into a sample than IR light, NIR imaging can be easily employed for investigating thick samples or bulk materials without any or much sample preparations.
4. Variations of hydration state can also be visualized.

6.3.2 Instruments for near-infrared imaging

In order to record NIR imaging data with high quality, high sensitivity, and high spectral resolution, spatial resolution, imaging speed, and imaged area are important factors [89]. Three Japanese companies have recently developed very unique NIR imaging devices [13,20,27,90,91], and the performance and technology of the imaging instrument have been much improved.

The first device is D-NIRs (Yokogawa Electric Corporation, Japan) [20]. The photodiode (PD) array detector with 640 elements enables high wavelength resolution in the 1000–1700 nm region with a spectral resolution of 1.25 nm. Adoption of this detector together with a new charge amplifier array integrated circuit contributes to high sensitivity and high speed of the instrument. The second one is a hyperspectral camera with a newly developed InGaAs photodiode array (T2SLSWIR focal plane array, SSW230A, Sumitomo Electric Industries Ltd., Yokohama, Japan) [27]. It accomplished the wide area and high-speed monitoring so as to obtain NIR data in the 1000–2350 nm region of a 150×200 mm^2 within 2–5 seconds [27]. The third is an imaging-type two-dimensional Fourier spectroscopic (ITFS) system (NT00-T011, NT00-T012, AOI ELECTRONICS Co., Ltd, Japan). The device is characterized by its spectroscopic mechanism with a partial movable mirror. An optical path difference to the object light, not the reference light, is given using the mirror [90,91]. The system is more immune to mechanical vibrations, and it enables a simpler optical configuration without adding any antivibration mechanism [90,91]. Furthermore, the light interference signals are observed only for rays that come from the same point in the focal plane, and the ones from outside the focal plane, on the other hand, are detected as the direct current component on the interferogram. That is, the system makes it possible to obtain confocal imaging data, and 3D NIR imaging data can be obtained by changing the focal plane in the z-direction. It is not too much to say that these device developments have opened up a new era of expanding the newest application of NIR imaging.

6.4 Application of near-infrared imaging to embryogenesis of fish eggs

In this section, our recent studies investigating embryogenesis of fish eggs are described as an application of NIR imaging to biology [30–34,92]. In the visualization of the fish egg development, the concentration of proteins and lipids changed with the development, and distributions of membrane proteins and lipoproteins were also explored [30–32]. Ultrahigh speed imaging using a newly developed device was also demonstrated [33]. Furthermore, a NIR imaging research expanded from a new perspective taking into account optical interference was also shown, which enabled nonstaining blood flow imaging in addition to biomolecular distribution such as water, proteins, and lipids [34].

6.4.1 Nonstaining visualization of embryogenesis in Japanese medaka (Oryzias latipes) fish egg by near-infrared imaging

Our recent research about NIR imaging of embryogenesis in Japanese medaka fish egg is introduced here as an example of the concrete application of NIR imaging to biology [30–34]. The fish species has been used in various biological research, such as embryonic development, genome analysis, carcinogenic testing, and genetic diseases [93–96]. The eggs produced are approximately 1.5 mm in diameter, and they are suitable for spectroscopic analysis because they are almost transparent. The egg consists of three major parts: oil droplets, egg yolk, and embryonic body after fertilization, and the detailed structure of embryonic body can be clearly seen on the seventh day, as shown in Fig. 6.5. The eggs hatched at around 2 weeks after fertilization, and the eggs on the day before

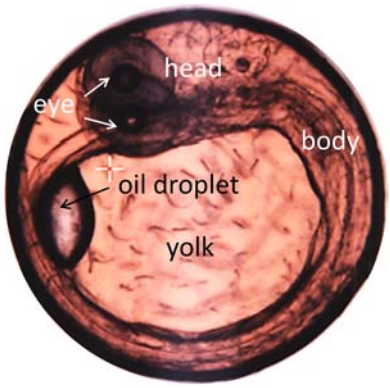


Figure 6.5
An optical image of a medaka fish egg on the seventh day after fertilization. Source: *Reproduced from P. Puangchit, M. Ishigaki, Y. Yasui, M. Kajita, P. Ritthiruangdej, Y. Ozaki, Non-staining visualization of embryogenesis and energy metabolism in medaka fish eggs using near-infrared spectroscopy and imaging, Analyst 142 (24) (2017) 4765–4772 with permission from The Royal Society of Chemistry.*

hatching were defined in our research as "just before hatching" (JBH) [30–33]. To regulate the optical path length, NIR measurements were performed by sandwiching an egg between two glass slides with pinchcocks and spacers.

Fig. 6.6A and B show NIR spectra in the 7500–4000 cm^{-1} region and their second derivatives in the 4900–4000 cm^{-1} region of five parts of medaka fish eggs (body, eye, head, oil droplet, and yolk) on the fifth day after fertilization. Broad bands at around 6950 and 5200 cm^{-1} arise from water in the egg [8]. Prominent peaks at 4258 and 4340 cm^{-1} in oil droplet spectra are due to the combination mode of C–H stretching and bending modes, and a residual band at 4412 cm^{-1} also comes from lipid components [44,97,98]. Fig. 6.6C

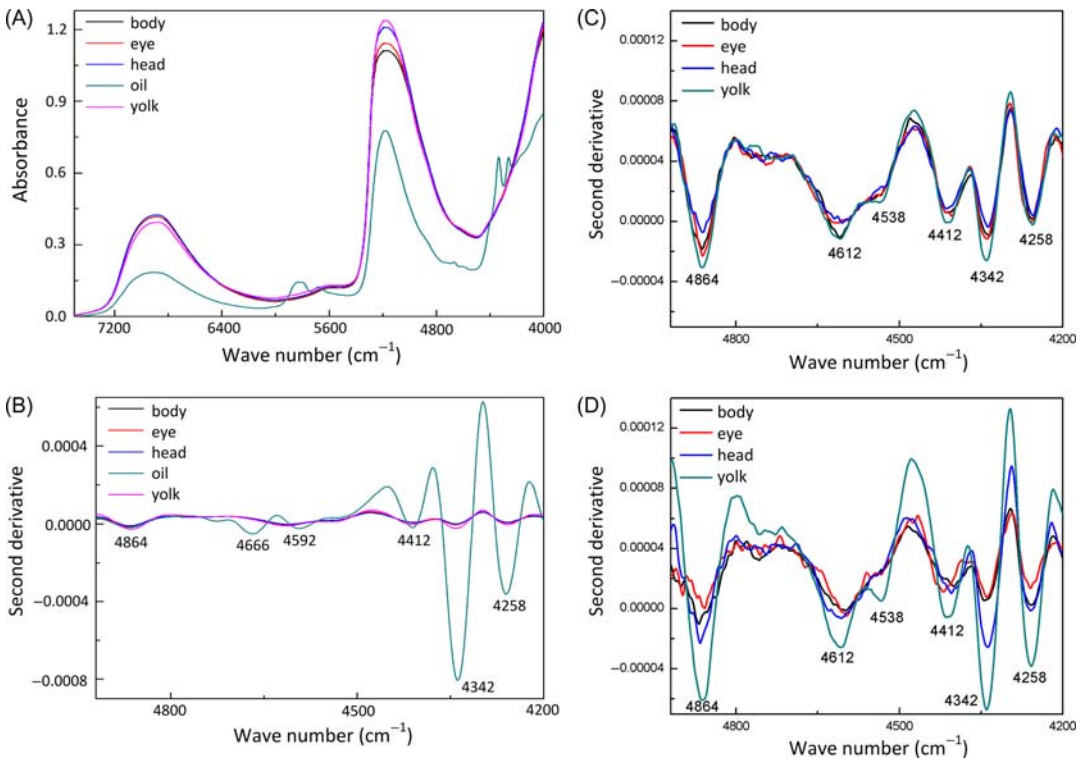


Figure 6.6

(A) Near-infrared (NIR) spectra (7500–4000 cm^{-1}) and (B) their second derivatives (4900–4000 cm^{-1}) of the body, eye, head, oil droplets, and yolk of a medaka egg on the fifth day after fertilization. The second derivative spectra in the 4900–4200 cm^{-1} region of the body, eye, head, and yolk parts on the (C) third day and (D) just before hatching (JBH). The spectra were obtained by a Perkin–Elmer instrument (Spectrum Spotlight 300). Source*: Reproduced from Reproduced from P. Puangchit, M. Ishigaki, Y. Yasui, M. Kajita, P. Ritthiruangdej, Y. Ozaki, Non-staining visualization of embryogenesis and energy metabolism in medaka fish eggs using near-infrared spectroscopy and imaging, Analyst 142 (24) (2017) 4765–4772 with permission from The Royal Society of Chemistry.*

and D present second derivative spectra of the body, eye, head, and yolk parts on the third day and JBH, respectively. In addition to the three bands due to lipid components (4412, 4342, and 4258 cm^{-1}), three peaks were observed at 4864, 4612, and 4538 cm^{-1}; they were assigned to amide combination modes of proteins [8,57–60]. The oil droplets are rich in lipids, while the egg yolk, eye, and head contain both protein and lipid components. Within the embryo, the body, eye, and head showed similar spectral patterns to those of yolk on the third day because embryonic body and yolk parts overlapped each other. However, on the day JBH, egg developments were completed enough to differentiate these spectra each other.

The distribution of biomolecules such as proteins and lipids, and their variations depending on egg development can be visualized by plotting a second derivative intensity of a characteristic peak of each component [32]. Fig. 6.7 shows VIS images and day-dependent

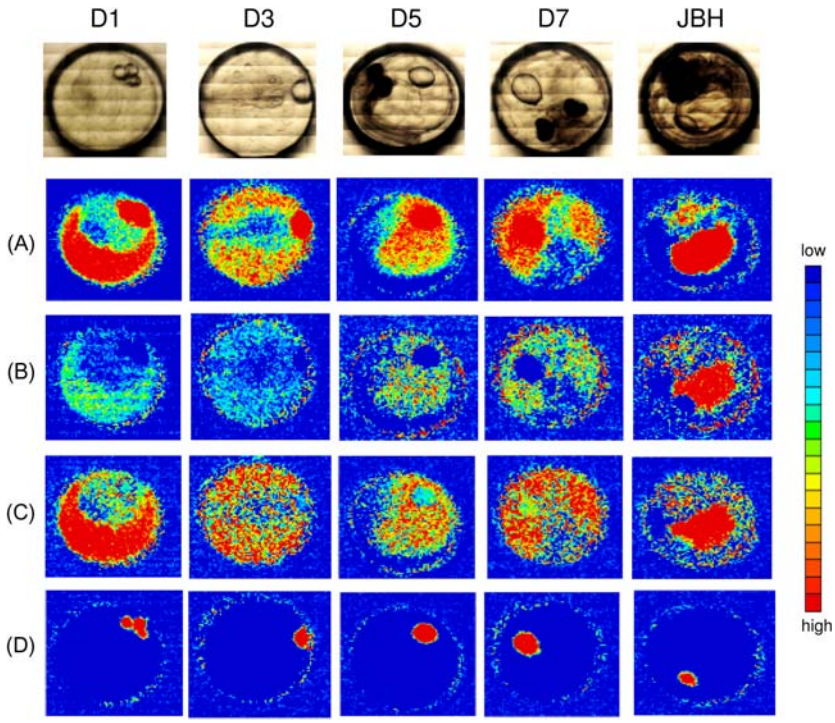


Figure 6.7

Day-dependent variations in near-infrared (NIR) images of fertilized medaka eggs [first, third, fifth, seventh, and just before hatching (JBH)] constructed by the second derivative intensities at (A) 4864, (B) 4616, (C) 4340, and (D) 4666 cm^{-1}. Source*: Reproduced from Reproduced from P. Puangchit, M. Ishigaki, Y. Yasui, M. Kajita, P. Ritthiruangdej, Y. Ozaki, Non-staining visualization of embryogenesis and energy metabolism in medaka fish eggs using near-infrared spectroscopy and imaging, Analyst 142 (24) (2017) 4765–4772 with permission from The Royal Society of Chemistry.*

biomolecular variations in NIR images of fertilized medaka eggs (first, third, fifth, seventh, and JBH) developed by the second derivative intensities at (A) 4864, (B) 4616, (C) 4340, and (D) 4666 cm^{-1}. The distribution of proteins is visualized by images which were constructed by the peak intensities at 4864 and 4616 cm^{-1}, respectively (Fig. 6.7A and B). The peak at 4864 cm^{-1} is assigned to the combination of N−H stretching and amide II (amide A + II) [73,74] modes, and the one at 4616 cm^{-1} comes from the combination of amide B and II modes (amide B + II) [60]; both bands have been reported to reflect the β-sheet and α-helix structures of proteins, respectively [57−60]. Fig. 6.7A and B shows similar patterns of protein distributions in the yolk and membrane. The protein concentration in the yolk part with a higher ratio of α-helix structure increased as the embryonic development progressed as shown in Fig. 6.7B, but the intensity of the amide band at 4616 cm^{-1} ascribed to the β-sheet structure was stronger on day 1 and JBH than during the intermediate period (Fig. 6.7A). One of the functions of the yolk during the early stages of fish embryo development is to provide nutriments for organogenesis, while its role in the later stages is to supply energy via metabolic processes [99−101]. Thus the yolk component seems to have particularly important roles in the first and last stages of egg development, and we may have identified biomaterial changes that are required for these specific roles.

Fig. 6.8A exhibits the averaged NIR second derivative spectra of egg membrane. Ten NIR spectra of egg membrane were recorded in point mode per an egg at five different stages (first, third, fifth, seventh, and JBH) and mean spectrum was calculated using 50 extracted spectra. The bands at 4342 and 4258 cm^{-1} result from a combination mode of C−H stretching and bending of hydrocarbons and aliphatic compounds [97,98], and the ones at 4864, 4538, and 4412 cm^{-1} and at 4612 cm^{-1} are sensitive to β-sheet and α-helix structures in proteins, respectively [8,56−58]. Generally, membrane proteins assume both α-helix and β-barrel secondary structures [102,103]. The NIR image developed by 4538 cm^{-1} peak intensity due to β-sheet structure made the membrane structure clear in Fig. 6.8B. In this way, NIR images successfully captured the membrane protein features constructing these secondary structures.

The peak at 4340 cm^{-1} resulted from a combination of C−H stretching and bending modes of hydrocarbons and aliphatic compounds [8,97,98], and their images demonstrate prominent features from oil droplets and yolk parts (Fig. 6.7C). A band at 4666 cm^{-1} arises from unsaturated fatty acids [97], and the images developed from its band intensity showed clear features of oil droplets. Generally, fish eggs contain polyunsaturated fatty acids such as docosahexaenoic acid and eicosapentaenoic acid, and therefore these NIR images are considered to correctly visualize the distribution of unsaturated fatty acids. Furthermore, the highlighted contour of egg membranes indicated the presence of unsaturated fatty acids also within the membrane. Unsaturated fatty acids in cell membranes play important roles to protect cells from outside stimulus [104], and their presence in the membranes is a well-known fact.

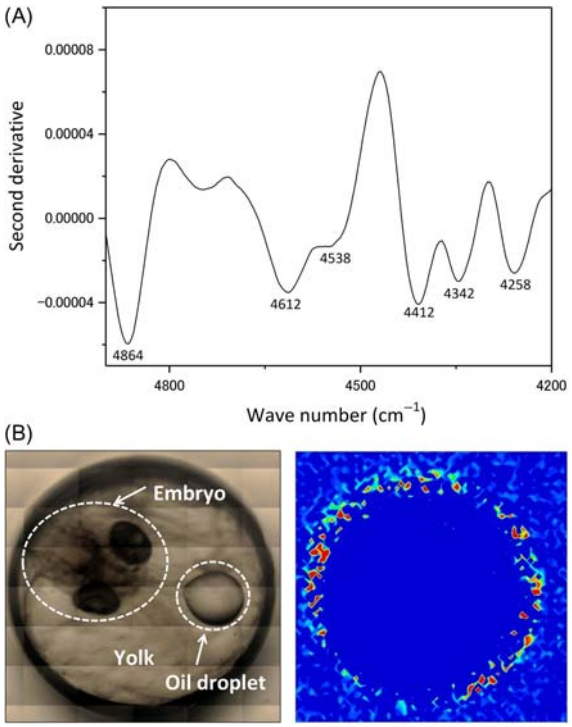


Figure 6.8

(A) The averaged near-infrared (NIR) second derivative spectrum obtained from egg membrane. (B) Comparison of an optical image of a medaka egg on the fifth day and a NIR image developed bythe second derivative intensity at 4530 cm^{-1} due to β-sheet structure of proteins. Source*: (A) Reproduced from Reproduced from P. Puangchit, M. Ishigaki, Y. Yasui, M. Kajita, P. Ritthiruangdej, Y. Ozaki, Non-staining visualization of embryogenesis and energy metabolism in medaka fish eggs using near-infrared spectroscopy and imaging, Analyst 142 (24) (2017) 4765–4772 with permission from The Royal Society of Chemistry. (B) Reproduced from M. Ishigaki, Y. Yasui, P. Puangchit, S. Kawasaki, Y. Ozaki, In vivo monitoring of the growth of fertilized eggs of medaka fish (Oryzias latipes) by near-infrared spectroscopy and near-Infrared imaging—a marked change in the relative content of weakly hydrogen-bonded water in egg yolk just before hatching, Molecules 21 (8) (2016) 1003 with permission from Multidisciplinary Digital Publishing Institute.*

6.4.2 Near-infrared images of the influence of bioactivity on water molecular structure

NIR spectroscopy has an advantage that water in a biological tissue can be easily analyzed because absorbance due to water in a NIR spectrum is relatively weak compared to that in an IR spectrum. A main constituent of living bodies is water and biomolecular interactions also occur within water. Therefore, the concentration gradient of biomolecules and their molecular structures should affect the molecular structure of water by hydrogen bonds. If

fact, water absorbance was detected in the NIR spectra from the samples of medaka fish eggs, and the influence of bioactivity on water structure was investigated.

Egg development is stimulated by fertilization, and protein synthesis is activated by gene expressions. In our study, the variations of yolk components such as proteins, lipids, and water were investigated and compared for four kinds of egg samples: (1) the eggs activated by fertilization, and three nonactivated groups where their embryogenesis was stopped or not started by (2) culturing under cold temperature, (3) instant freezing, or (4) not being fertilized. In the NIR absorbance and second derivative spectra obtained from egg yolk of four species, water protein, and lipid bands were observed (Fig. 6.9A). The broad features at around 6950 and 5200 cm^{-1} were due to water, the peaks at 4258 and 4340 cm^{-1} from lipid components, and the one at 4864 cm^{-1} of was from proteins. In order to investigate differences in the yolk components between the activated and nonactivated eggs, PCA was performed to the dataset of NIR second derivatives including all four groups to get a set of orthogonal principal components (PCs) to account for the maximum variance in the spectral dataset [105].

The dataset was classified into two groups by PC1 and PC2 as shown in Fig. 6.9B when the data set was restricted in the water absorbance region (5500−5000 cm^{-1}), and the discriminant component between the activated and nonactivated ones was indicated to be related to water structure (Fig. 6.9C). Generally, water consists of two main species at normal temperature for organisms: strongly and weakly hydrogen bonding water species [106−109]. Regarding the water band at approximately 5250 cm^{-1}, Czarnik-Matusewicz et al. assigned two bands at around 5200 and 5000 cm^{-1} to weakly and strongly hydrogen-bonded water, respectively [110]. That is, the present results revealed that the yolk in the activated eggs had a higher ratio of weakly hydrogen bonding water than that in the nonactivated ones. The bioactivity was likely to affect the hydrogen bonds of water and the relationship between protein and water structures can be expected to be simultaneously investigated.

Fig. 6.10 exhibits NIR images constructed by plotting second derivative intensities at 4336 cm^{-1} due to lipids (i), and 4864 cm^{-1} arising from proteins (ii), and an intensity ratio, 5172/5250 cm^{-1} of strongly and weakly hydrogen bonding water (iii). Lipid and protein components were again confirmed to be much included in the yolk part and less in the blastocyst which would be an embryonic body. The images of strongly hydrogen bonding water were more highlighted in the yolk part, and its water component was less distributed in the blastodisc constructed by cytoplasm after fertilization. In an unfertilized egg, oil droplets were scattered at random within the egg, and cytoplasm and yolk parts were not distinguished (Fig. 6.10D). Eggs of which development was stopped under low temperature showed more distribution of strongly hydrogen bonding water even though they were kept in the material bias similar to the activated eggs (Fig. 6.10B). In a once frozen egg, the

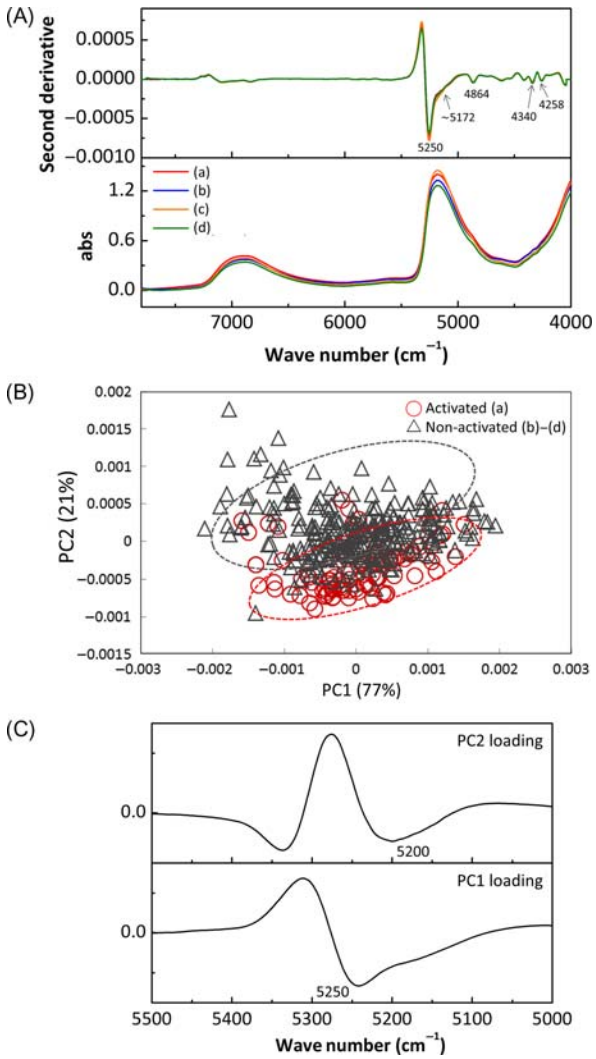


Figure 6.9

(A) Near-infrared (NIR) absorbance and their second derivative spectra obtained from egg yolk of four species. (B) Score plots and (C) loading plots of principal component 1 (PC1) and principal component 2 (PC2) of the data set of NIR second derivative spectra including four kinds of fish eggs (A)−(D).

internal structures that had been constructed by fertilization (Figs. 6.10A and B) seemed to be completely destroyed by freezing. With the deactivated of egg developments, they showed water structural changes yielding the higher ratio of strongly hydrogen bonding water. These NIR imaging patterns of showing the water structural differences between the activated and nonactivated eggs are expected to relate to the phenomena of protein

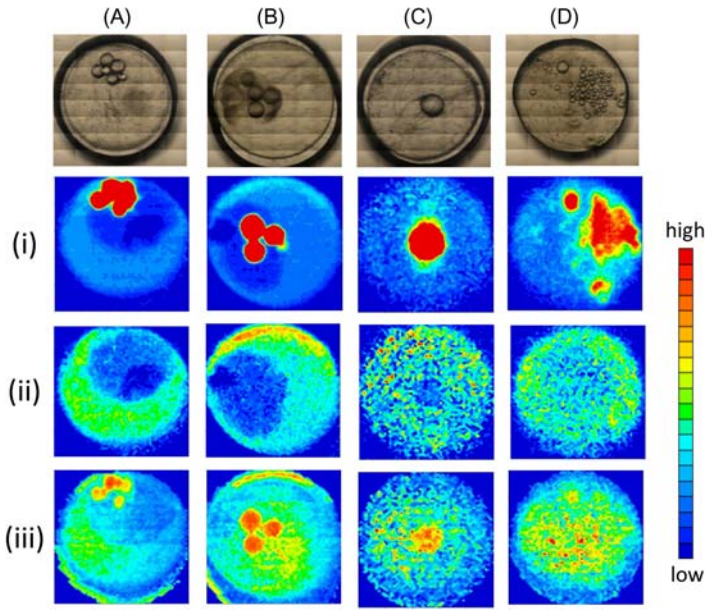


Figure 6.10
Visible (VIS) images of the four kinds of fish egg groups: (A) the eggs activated by fertilization, and three nonactivated groups where their embryogenesis were stopped or not started by (B) culturing under cold temperature, (C) instant freezing, or (D) not be fertilized. Near-infrared (NIR) images constructed by second derivative intensities at (i) 4336, (ii) 4864 cm^{-1}, and (iii) the ratio defied as 5172/5250 cm^{-1}.

concentration variations, protein structural differences, protein phosphorylation, and so on. In this way, NIR imaging can capture the protein−water interaction in vivo and in situ, and bioactivity can be assessed by its interaction.

6.4.3 High-speed near-infrared imaging of the embryonic development in fertilized fish eggs

NIR images using the higher wave number region (9000−5500 cm^{-1} were also constructed. The imaging data were obtained by a different microscope NIR imaging system with a hyperspectral camera [89,111]. The novelty of the system was characterized by the hyperspectral camera, which allowed us to build NIR images of fish eggs in a few seconds. This is approximately 1000 times faster than the measurement time of conventional NIR imaging systems.

Fig. 6.11A shows averaged NIR spectra in the 1000−2200 nm (10,000−4550 cm^{-1}) region of the three major parts (yolk, oil droplets, and embryo) of fertilized medaka eggs measured on the first day after fertilization and the day JBH. Broad absorbance bands due to water

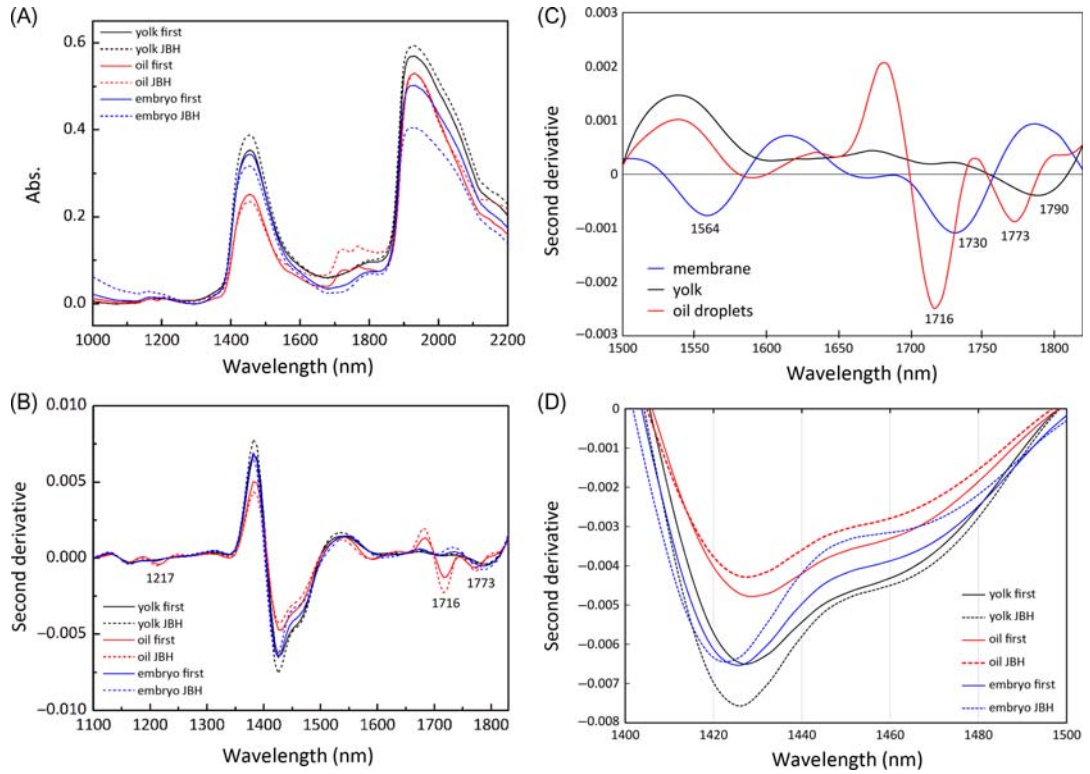


Figure 6.11

(A) NIR spectra in the 1000–2200 nm (10,000–4550 cm^{-1}) region of the yolk, oil droplets, and embryo of medaka eggs on the first day after fertilization (*solid line*) and the day just before hatching (*dashed line*) with baseline corrections. (B) Second derivative spectra in the 1100–1820 nm (9100–5500 cm^{-1}) region. (C) The comparison of second-derivative spectra in the 1500–1820 nm (6660–5500 cm^{-1}) region between egg membrane, yolk, and oil droplets. (D) Second-derivative spectra in the 1400–1500 nm region. Source*: Reproduced from M. Ishigaki, T. Nishii, P. Puangchit, Y. Yasui, C.W. Huck, Y. Ozaki, Noninvasive, high-speed, near-infrared imaging of the biomolecular distribution and molecular mechanism of embryonic development in fertilized fish eggs, J. Biophoton. 11 (4) (2018) e201700115 with permission. Copyright (2002) John Wiley & Sons, Ltd.*

were also detected at around 1430 (7000) and 1910 (5200 cm^{-1}) nm. Fig. 6.11B depicts their corresponding second derivative spectra in the 1100–1820 nm (9100–5500 cm^{-1}) region. The spectra from oil droplets had several peaks due to lipids. A band at 1217 nm (8216 cm^{-1}) resulted from the second overtone of the C–H stretching mode of the CH_2 groups, and two prominent bands at 1716 (5827) and 1773 (5640 cm^{-1}) nm were assigned to the first overtone of the C–H stretching of the CH_2 groups in hydrocarbons and aliphatic compounds [8,98,112]. A broad feature centered at approximately 2050 nm (4878 cm^{-1}) in the yolk and embryo spectra is ascribed to a combination of N–H stretching and amide II modes [1,2]. Fig. 6.11C shows the comparison of second derivative spectra in the

1500−1820 nm (6660−5500 cm^{-1}) region among the egg membrane, yolk, and oil droplets. The membrane spectrum exhibited two characteristic peaks at around 1564 (6394) and 1730 (5780 cm^{-1}) nm, which were attributed to the first overtone of N−H stretching mode of amide groups and that of the C−H stretching modes of CH_2 groups, respectively [1,2,57,59,78]. These significant peaks proved that the egg membrane includes rich aliphatic compounds and proteins, which are well described features of a cell membrane. Furthermore, the first overtone of C−H stretching in unsaturated fatty acids has been reported to have significant peaks in 1600−1900 nm region (6250−5263 cm^{-1}), and in saturated fatty acids, on the other hand, the corresponding bands have relatively low frequency suggested by quantum chemical calculation [113]. The peak position of C−H stretching band in the unsaturated fatty acids was shifted from approximately 1725 nm (5800 cm^{-1}) to 1709−1717 nm (5850−5824 cm^{-1}) [97,114]. Therefore the peak due to the first overtone of C−H stretching band at 1716 nm of oil droplets is very likely to be from the unsaturated fatty acids.

The spectral shape in the 1880−1980 and 1400−1500 nm regions due to water absorption differs significantly between oil droplets and other parts of the eggs, and between on first day and JBH as shown in Fig. 6.11A. This indicates that the water environment surrounding oil droplets was significantly different from that surrounding the other parts of the egg, and water structure changes with the egg development. Fig. 6.11D shows the second derivatives of NIR spectra in the 1400−1500 nm region. The relative contributions of two water components (bands at 1420 and 1460 nm were due to weaker and stronger hydrogen bonds, respectively [106]) were evaluated by calculating the ratio of the second derivative intensities defined as 1460/1426 nm to discuss the trend of the water behavior characteristic for these parts. The embryonic parts have been suggested to contain lower concentrations of strongly hydrogen-bonded water species relative to those in the yolk, and oil droplets have the highest concentration of strongly hydrogen-bonded water species.

Fig. 6.12A exhibits VIS images of a medaka fish egg from the first day to the day JBH, and Fig. 6.12B−F show NIR images developed by plotting the second derivative intensities of some notable bands due to proteins, lipids, and water, and the distribution of weakly hydrogen-bonding water. Fig. 6.12B, which is constructed by plotting the second derivative intensity at 1767 nm due to the first overtone of the C−H stretching modes in the middle of the bands between egg membrane (1730 nm) and oil droplets (1773 nm), demonstrates the distribution of aliphatic compounds, such as fatty acids and hydrocarbons [8,98,112]. The shape of the egg membrane, the contour of oil droplets, and the outline of the embryo were all highlighted. The phospholipid bilayer structure of a cell membrane including a large number of C−H groups was visualized. Furthermore, eye structures were made clear after the third day. Eyes are known to include sphingolipids and glycerolipids with long carbon chains, which made their structure clear without staining. The images of C−H stretching band at 1716 nm of CH_2 groups in Fig. 6.12C showed the distribution of particularly

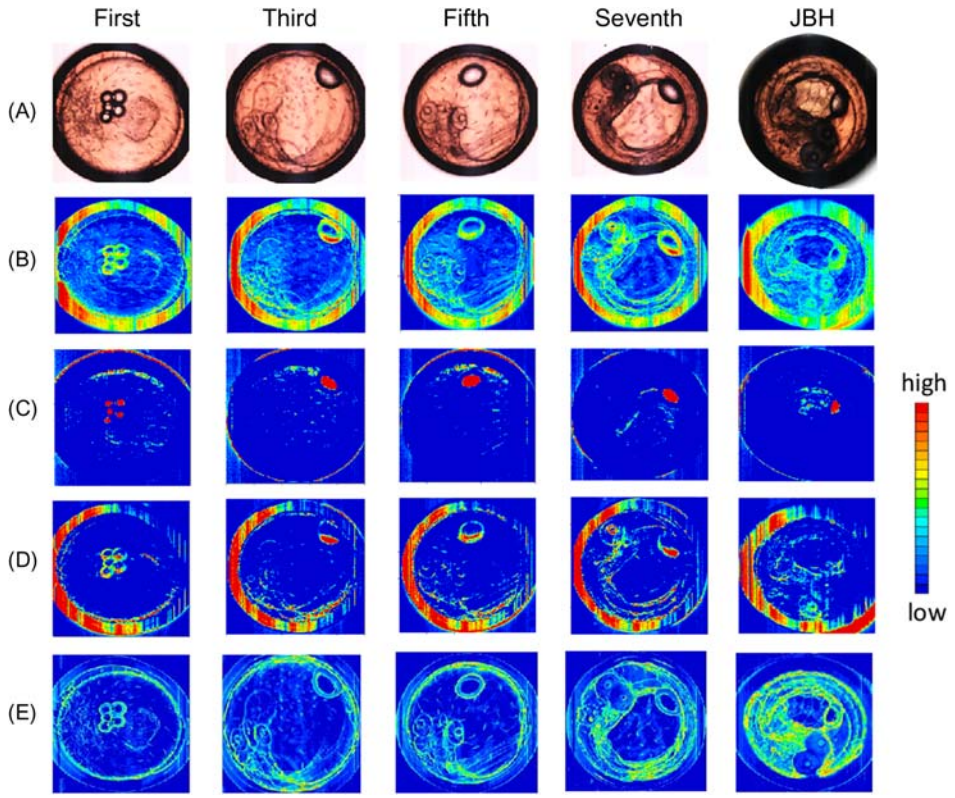


Figure 6.12

(A) Visible (VIS) images of medaka eggs from the first day after fertilization to the day just before hatching. Near-infrared (NIR) images of eggs based on band intensities in second-derivative spectra at (B) 1767 nm (aliphatic compounds), (C) 1716 nm (unsaturated fatty acids), (D) 1564 nm (proteins), and (E) scores calculated by projecting principal component 1 (PC1) loadings onto imaging data to identify the contribution of weakly hydrogen-bonded water. Source*: Reproduced from M. Ishigaki, T. Nishii, P. Puangchit, Y. Yasui, C.W. Huck, Y. Ozaki, Noninvasive, high-speed, near-infrared imaging of the biomolecular distribution and molecular mechanism of embryonic development in fertilized fish eggs, J. Biophoton. 11 (4) (2018) e201700115 with permission. Copyright (2002) John Wiely & Sons, Ltd.*

unsaturated fatty acids in various kinds of fatty acids and showed their presence in oil droplets and egg membrane. Fig. 6.12D was constructed based on the intensity of the band at 1564 nm ($\sim$6400 cm^{-1}), attributed to the first overtone of N−H stretching mode of amide groups [104,115,116]. First, the structure of the egg membrane is clearly depicted because it contains many membrane proteins, also shown in the second derivative spectra of egg membrane (Fig. 6.11). In addition, heterogeneous structures within oil droplets are apparent. The positions of these heterogeneous components overlap between Fig. 6.12B and D, and they are likely to show the distribution of lipoproteins. Furthermore, the structures

of eyes are supposed to be made clear because they contain a lot of collagen in the vitreous body. In the distribution map of weakly hydrogen bonding water (Fig. 6.12E), detailed structures such as eyes in the embryonic body, egg membranes, and the inner and the outer contours of oil droplets were apparent. It is very interesting to have made clear the different water structures at the interfaces between different substances. The high-speed NIR system achieved great advances in visualizing the distribution of biomaterials and its molecular mechanism of embryonic development nondestructively, noninvasively, and almost instantaneously at the molecular level without staining. Especially, the results make clear the distribution of protein components, such as membrane proteins, lipoproteins, and collagen individually.

6.4.4 Near-infrared in vivo imaging of blood flow and molecular distribution in a developing fish egg using an imaging-type two-dimensional Fourier spectrometer

This section introduces our recent research of NIR imaging by another type of imaging system: (AOI ELECTRONICS Co., Ltd., NT00-T011), to demonstrate the future possibility of NIR imaging [34]. The system has a significantly different spectroscopic mechanism compared with the conventional so-called FT-NIR spectrometer [88,89]. An optical path difference to the object light is given using a partially movable mirror, and an interferogram can be obtained by continuously changing the spatial phase difference. The strongest advantage of the system is being able to obtain confocal imaging data, and even 3D NIR imaging data may be obtained by changing the focal plane in the z-direction.

Fig. 6.13A shows an optical image of a medaka fish egg on the fifth day after fertilization, and Fig. 6.13B and C displays the interferogram recorded in the reflectance mode from the yolk part (A) and the heart part (B) (Fig. 6.13A). The interferogram detected at the heart part (B) yielded a noise-like background (Fig. 6.13C) in addition to the normal center burst (Fig. 6.13B). The NIR spectra were calculated by Fourier transformation of the interferogram. Fig. 6.14A depicts an averaged NIR absorbance spectrum (1000−2500 nm) of the yolk part calculated by Fourier transformation of the interferogram, and it shows broad features due to water absorption. Fig. 6.14B and C exhibited observed light intensity in the 1000−2500 nm and 2000−15,000 nm regions recorded from the beat part (B) in Fig. 6.13A. It indicates that the light with the specific frequency associated with the beat cycle emerged from the sample. The strongest peak can be seen at 3768 nm, and two pronounced peaks are also observed at 1884 and 1256 nm. The peak at 3768 nm was revealed to be due to the fundamental mode of the heartbeat, and the two peaks (1884 and 1256 nm) in Fig. 6.14B correspond to the first and second overtones of the fundamental mode, respectively. By plotting the intensities of (A) detected light and (B) absorbances at 1256 and 1884 nm, nonstaining blood flow images were successfully observed (Fig. 6.15). Furthermore, the simultaneous visualization of (A) heart, (B) blood vessels, and the

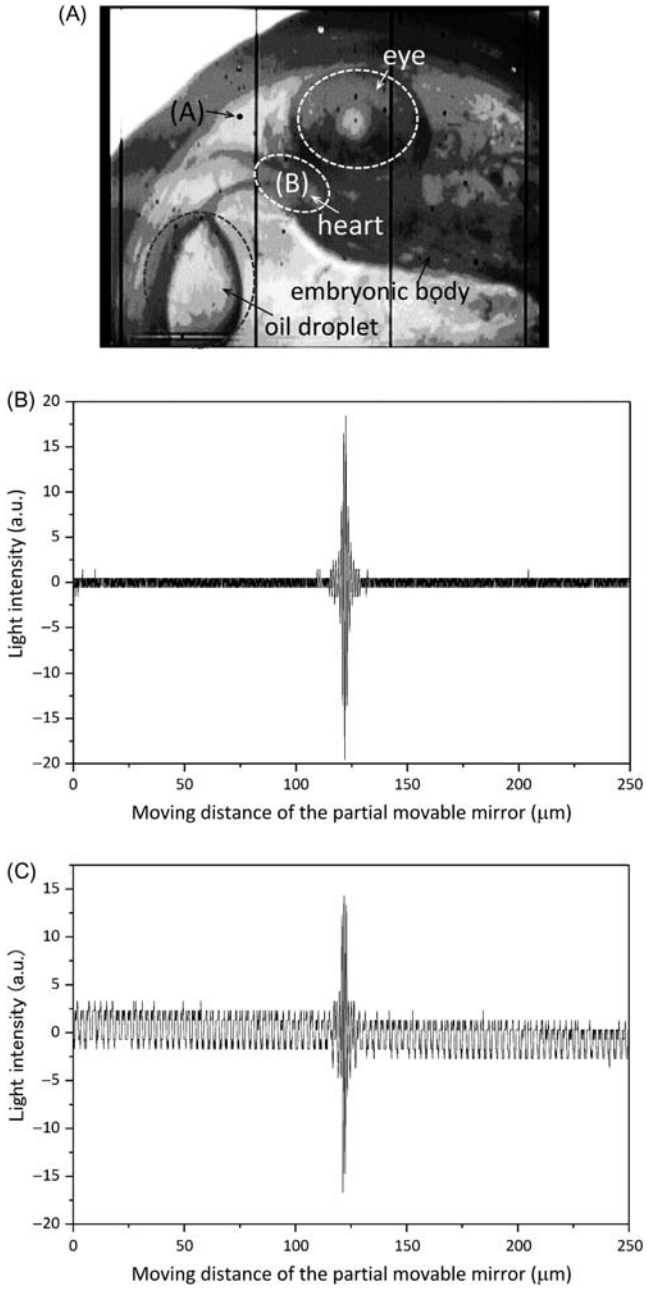


Figure 6.13

(A) An optical image of an embryonic body of a medaka fish egg on the fifth day after fertilization. (B) An interferogram obtained from the yolk part (A), and (C) from the heart part (B) of (A). Source*: Reproduced from M. Ishigaki, P. Puangchit, Y. Yasui, A. Ishida, H. Hayashi, et al., Nonstaining blood flow imaging using optical interference due to Doppler shift and near-infrared imaging of molecular distribution in developing fish egg embryos, Anal. Chem. 90 (8) (2018) 5217–5223 with permission from The Royal Society of Chemistry.*

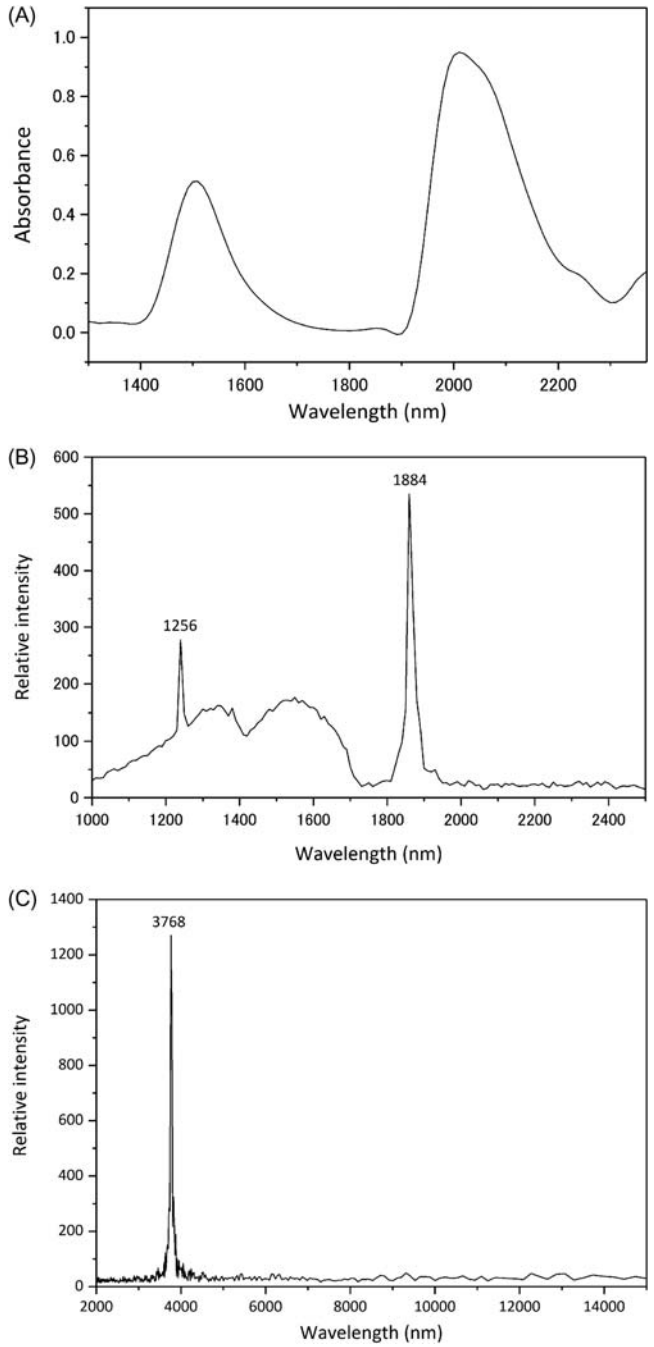


Figure 6.14

(A) An averaged near-infrared (NIR) absorbance spectrum obtained in the vicinity of Point (A). The observed light intensity recorded at Point (B) in Fig. 6.13 in the (B) 1000–2500 nm and (C) 2000–15,000 nm regions. Source*: Reproduced from M. Ishigaki, P. Puangchit, Y. Yasui, A. Ishida, H. Hayashi, et al., Nonstaining blood flow imaging using optical interference due to Doppler shift and near-infrared imaging of molecular distribution in developing fish egg embryos, Anal. Chem. 90 (8) (2018) 5217–5223 with permission from The Royal Society of Chemistry.*

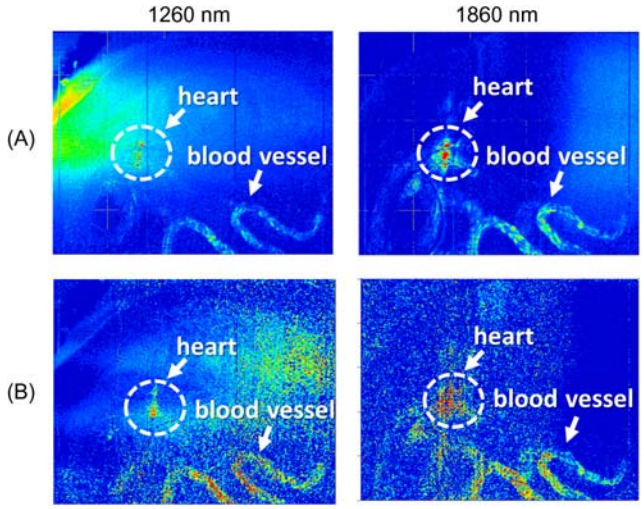


Figure 6.15
Blood flow images of a medaka fish egg on the fifth day after fertilization obtained by plotting intensities of (A) detected light from the sample and (B) absorbance at 1260 and 1860 nm.
Source: *Reproduced from M. Ishigaki, P. Puangchit, Y. Yasui, A. Ishida, H. Hayashi, et al., Nonstaining blood flow imaging using optical interference due to Doppler shift and near-infrared imaging of molecular distribution in developing fish egg embryos, Anal. Chem. 90 (8) (2018) 5217−5223 with permission from The Royal Society of Chemistry.*

biomolecular distributions of (C) water (O−H) and (D) hydrocarbons or aliphatic compounds (C−H) was also successfully accomplished as in Fig. 6.16. Nonstaining 3D images of molecular distribution and blood flow may be obtained in principle using ITFS. In the near future, 3D NIR mapping of biological body, cardiogenesis, and induced pluripotent stem cell differentiation into cardiomyocytes using optical interference is expected to be achieved.

6.5 Future prospects

As described in this review, NIR spectroscopy is a powerful tool to investigate protein structure and hydration in vivo and in vitro both in basic science and applications. The bioactivity of a living body might be assessed from the aspect of interaction variations between proteins and water caused by protein structural changes and protein phosphorylation. Recent developments in NIR devices are remarkable, and NIR imaging systems have achieved very rapid and attractive performances. Although the miniaturization of NIR spectrometers is still in progress, portable NIR spectrometers and imaging systems and hand-held spectrometers have already been used in various places, including factories and fields. Moreover, a new measurement method named "ultrasonic-assisted mid-infrared

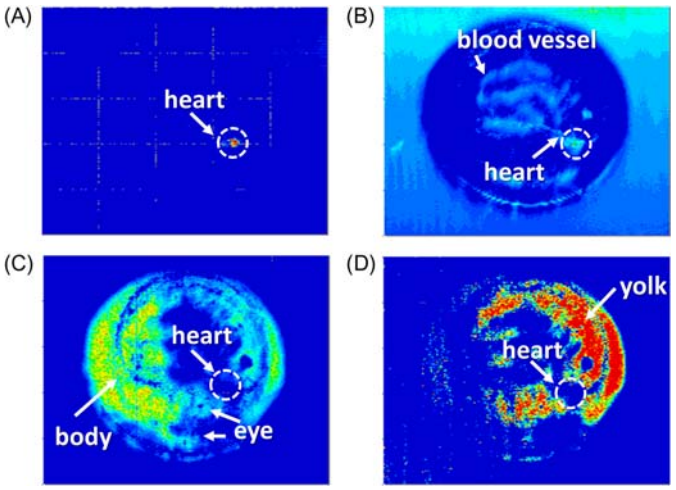


Figure 6.16

Near-infrared (NIR) images obtained by plotting intensities of (A) detected light at 2260 nm and absorbance at (B) 1880 nm, (C) 1940 nm, and (D) 2360 nm. Source: *Reproduced from M. Ishigaki, P. Puangchit, Y. Yasui, A. Ishida, H. Hayashi, et al., Nonstaining blood flow imaging using optical interference due to Doppler shift and near-infrared imaging of molecular distribution in developing fish egg embryos, Anal. Chem. 90 (8) (2018) 5217−5223 with permission from The Royal Society of Chemistry.*

spectroscopic imaging method" has emerged [117]. Ishimaru et al. has proposed to generate a reflection plane at a depth of 100 μm from the surface of the sample by creating an ultrasonic wave [117]. The application of the method to NIR spectroscopy and imaging can enable 3D NIR imaging of biomolecular distribution in living cells within a short time. Another important recent progress is that of quantum chemical calculations of NIR spectra of various molecules [113,118−122]. Even NIR spectra of long-chain fatty acids have been successfully reproduced. In the near future, NIR spectra of peptides may be calculated by quantum chemical calculation, providing a solid base for band assignments of proteins.

Further application possibilities of NIR spectroscopy and imaging are expected to expand in the near future. Moreover, surface plasmon resonance-NIR, which provides about 100 times absorption enhancement compared with the absorption in the corresponding attenuated total reflection NIR spectra [123,124], will extend markedly the detection limit of NIR spectroscopy. It also allows NIR spectral measurement with a much smaller quantity of samples. This technology may expand dramatically the application of NIR spectroscopy. Medical diagnosis based on water structure to assess bioactivity might also become feasible. NIR spectroscopy and imaging are still young spectroscopy and are developing strongly.

References

[1] H.W. Siesler, Y. Ozaki, S. Kawata, H.M. Heise (Eds.), Near-Infrared Spectroscopy, Wiley-VCH, Weinheim, 2002.

[2] Y. Ozaki, W.F. McClure, A.A. Christy (Eds.), Near-Infrared Spectroscopy in Food Science and Technology, John Wiley & Sons, 2006.
[3] Y. Ozaki, C.W. Huck, K.B. Beć, Near infrared spectroscopy and its applications, in: V.P. Gupta (Ed.), Molecular and Laser Spectroscopy, Elsevier, 2017, pp. 11–38.
[4] H.M. Heise, Applications of near-infrared spectroscopy in medical sciences, in: H.W. Siesler, Y. Ozaki, S. Kawata, H.M. Heise (Eds.), Near-Infrared Spectroscopy, Wiley-VCH, Weinheim, 2002, p. 289.
[5] Y. Ozaki, Near-infrared spectroscopy—its versatility in analytical chemistry, Anal. Sci. 28 (2012) 545–563.
[6] Application of near infrared spectroscopy, in: T. Jue, K. Masuda (Eds.), Biomedicine (Handbook of Modern Biophysics), Springer, 2016.
[7] F.F. Jobsis, Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters, Science 198 (4323) (1977) 1264–1267.
[8] J. Workman, L. Weyer Jr, Practical Guide to Interpretive Near-Infrared Spectroscopy, CRC Press, 2007.
[9] E.W. Ciurczak, J.K. Drennen III (Eds.), Pharmaceutical and Medical Applications of Near-Infrared Spectroscopy, CRC Press, Baco Raton, FL, 2002.
[10] E.W. Ciurczak, B. Igne (Eds.), Pharmaceutical and Medical Applications of Near-Infrared Spectroscopy, second ed., CRC Press, Baco Raton, FL, 2014.
[11] R. Salzer, H.W. Siesler (Eds.), Infrared and Raman Spectroscopic Imaging, second version, Wiley-VCH, Weinheim, 2014.
[12] S. Šašíc, Y. Ozaki (Eds.), Raman, Infrared, and Near-Infrared Chemical Imaging, John Wiley & Sons, Inc, New York, 2009.
[13] D. Ishikawa, H. Shinzawa, T. Genkawa, S.G. Kazarian, Y. Ozaki, Recent progress of near infrared (NIR) imaging—development of novel instruments and their applicability for practical situations, Anal. Sci. 30 (2013) 143–150.
[14] E.N. Lewis, L.H. Kidder, Technologies and practical considerations for implementing near infraed chemical imaging, in: S. Šašić, Y. Ozaki (Eds.), Raman, Infrared, and Near-Infrared Chemical Imaging, JohnWiley & Sons, Inc, New York, 2009, p. 75.
[15] H. Sato, Y. Ozaki, J.H. Jiang, R.Q. Yu, H. Shinzawa, Vibrational spectroscopy imaging in polymers, in: S. Šašić, Y. Ozaki (Eds.), Raman, infrared, and near-infrared chemical imaging, John Wiley & Sons, Inc, New York, 2009, p. 263.
[16] S.G. Kazarian, P.S. Wray, Applications of FTIR spectroscopic imaging in pharmaceutical science, in: S. Šašić, Y. Ozaki (Eds.), Raman, Infrared, and Near-Infrared Chemical Imaging, JohnWiley & Sons, Inc, New York, 2009, p. 185.
[17] H. Shinzawa, M. Nishida, A. Tsuge, D. Ishikawa, Y. Ozaki, S. Morita, et al., Thermal behavior of poly (lactic acid)-nanocomposite studied by near-infrared imaging based on roundtrip temperature scan, Appl. Spectrosc. 68 (2014) 371–378.
[18] H. Shinzawa, T.N. Murakami, M. Nishida, W. Kanematsu, I. Noda, Near-infrared (NIR) imaging analysis of polylactic acid (PLA) nanocomposite by multiple-perturbation two-dimensional (2D) correlation spectroscopy, J. Mol. Struct. 1069 (2014) 171–175.
[19] H. Shinzawa, M. Nishida, T. Tanaka, W. Kanematsu, Accelerated weathering-induced degradation of poly (lactic acid) fiber studied by near-infrared (NIR) hyperspectral imaging, Appl. Spectrosc. 66 (4) (2012) 470–474.
[20] D. Ishikawa, K. Murayama, T. Genkawa, K. Awa, M. Komiyama, Y. Ozaki, Development of a compact near infrared imaging device with high-speed and portability for pharmaceutical process monitoring, NIR N. 23 (8) (2012) 19–23.
[21] D. Ishikawa, K. Murayama, T. Genkawa, K. Awa, M. Komiyama, S.G. Kazarian, et al., Application of a Newly Developed Portable NIR Imaging Device to Monitor the Dissolution Process of Tablets, Anal. Bioanal. Chem. 405 (29) (2013) 9401–9409.
[22] F.W. Koehler, E. Lee, L.H. Kidder, E.N. Lewis, Near infrared spectroscopy: the practical chemical imaging solution, Spectrosc. Eur. 14 (3) (2002) 12–19.

[23] T. Furukawa, H. Sato, Y. Kita, K. Matsukawa, H. Yamaguchi, S. Ochiai, et al., Molecular structure, crystallinity and morphology of polyethylene/polypropylene blends studied by Raman mapping, scanning electron microscopy, wide angle X-ray diffraction, and differential scanning calorimetry, Polym. J. 38 (11) (2006) 1127–1136.
[24] H. Shinzawa, K. Awa, Y. Ozaki, Compression-induced morphological and molecular structural changes in cellulose tablets probed by near-infrared imaging, J. Infrared Spectrosc. 19 (2011) 15–22.
[25] M. Unger, Y. Ozaki, H.W. Siesler, Variable-temperature fourier transform near-infrared (FT-NIR) imaging spectroscopy of the diffusion process of Butanol (OD) into polyamide 11, Appl. Spectrosc. 65 (9) (2011) 1051–1055.
[26] N. Suttiwijitpukdee, H. Sato, M. Unger, Y. Ozaki, Effects of hydrogen bond intermolecular interactions on the vrystal spherulite of poly (3-hydroxybutyrate) and cullulose acetate butyrate blends: studied by FT-IR and FT-NIR imaging spectroscopy, Macromolecules 45 (2012) 2738–3748.
[27] D. Ishikawa, T. Nishii, F. Mizuno, S.G. Kazarian, Y. Ozaki, Development of a high-speed monitoring near infrared hyperspectral camera (Compovision) for wide area imaging and its applications, NIR News 24 (5) (2013) 6–11.
[28] D. Ishikawa, T. Nishii, F. Mizuno, H. Sato, S.G. Kazarian, Y. Ozaki, Potential of a newly developed high-speed near infrared (NIR) camera (Compovision) in polymer industrial analyses: monitoring crystallinity and crystal evolution of polylactic acid (PLA) and concentration of PLA in PLA/poly-(R)-3hydroxybutyrate (PHB) blends, Appl. Spectrosc. 67 (12) (2013) 1441–1446.
[29] D. Ishikawa, D. Furukawa, T.T. Wei, K.R. Reddy, A. Motomura, Y. Igarashi, et al., High-speed monitoring of the crystallinity change in poly (lactic acid) during photodegradation by using a newly developed wide area NIR imaging system (Compovision), Anal. Bioanal. Chem. 407 (2015) 397–403.
[30] M. Ishigaki, S. Kawasaki, D. Ishikawa, Y. Ozaki, Near-infrared spectroscopy and imaging studies of fertilized fish eggs: in vivo monitoring of egg growth at the molecular level, Sci. Rep. 6 (2016) 20066.
[31] M. Ishigaki, Y. Yasui, P. Puangchit, S. Kawasaki, Y. Ozaki, In vivo monitoring of the growth of fertilized eggs of medaka fish (*Oryzias latipes*) by near-infrared spectroscopy and near-Infrared imaging—a marked change in the relative content of weakly hydrogen-bonded water in egg yolk just before hatching, Molecules 21 (8) (2016) 1003.
[32] P. Puangchit, M. Ishigaki, Y. Yasui, M. Kajita, P. Ritthiruangdej, Y. Ozaki, Non-staining visualization of embryogenesis and energy metabolism in medaka fish eggs using near-infrared spectroscopy and imaging, Analyst 142 (24) (2017) 4765–4772.
[33] M. Ishigaki, T. Nishii, P. Puangchit, Y. Yasui, C.W. Huck, Y. Ozaki, Noninvasive, high-speed, near-infrared imaging of the biomolecular distribution and molecular mechanism of embryonic development in fertilized fish eggs, J. Biophoton. 11 (4) (2018) e201700115.
[34] M. Ishigaki, P. Puangchit, Y. Yasui, A. Ishida, H. Hayashi, et al., Nonstaining blood flow imaging using optical interference due to Doppler shift and near-infrared imaging of molecular distribution in developing fish egg embryos, Anal. Chem. 90 (8) (2018) 5217–5223.
[35] M. Ferrari, V. Quaresima, A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application, Neuroimage 63 (2012) 921–935.
[36] A. Gerich, J. Dubois, L.H. Kidder, NIR imaging applications in the pharmaceutical industry, in: S. Šašić, Y. Ozaki (Eds.), Raman, Infrared, and Near-Infrared Chemical Imaging, JohnWiley & Sons, Inc, New York, 2009, p. 205.
[37] P. Geladi, M. Manley, Near-infrared hyperspectral imaging in food research, in: S. Šašić, Y. Ozaki (Eds.), Raman, Infrared, and Near-Infrared Chemical Imaging, JohnWiley & Sons, Inc, New York, 2009, p. 243.
[38] R.A. Shaw, V.V. Kupriyanov, O. Jilkima, M.G. Sowa, Near-infrared in vivo spectroscopic imaging: biomedical research and clinical applications, in: S. Šašić, Y. Ozaki (Eds.), Raman, Infrared, and Near-Infrared Chemical Imaging, John Wiley & Sons, Inc, New York, 2009, p. 149.
[39] Y. Wu, B. Czarnik-Matusewicz, K. Murayama, Y. Ozaki, Two-dimensional near-infrared spectroscopy study of human serum albumin in aqueous solutions: using overtones and combination modes to monitor temperature-dependent changes in the secondary structure, J. Phys. Chem. 104 (2002) 5840.

[40] I. Noda, Y. Ozaki (Eds.), Two-Dimensional Correlation Spectroscopy: Applications In Vibrational And Optical Spectroscopy, John Wiley & Sons, Chichester, 2004.
[41] Y. Hu, J. Zhang, H. Sato, Y. Futami, I. Noda, Y. Ozaki, C−H · · · O = C hydrogen bonding and isothermal crystallization kinetics of poly (3-hydroxybutyrate) investigated by near-infrared spectroscopy, Macromolecules 39 (2006) 3841.
[42] D. Voet, J.G. Voet, C.W. Pratt (Eds.), Fundamentals of Biochemistry: Life at the Molecular Level, John Wiley & Sons, 2016.
[43] K. Wüthrich, NMR with proteins and nucleic acids, Europhys. N. 17 (1) (1986) 11−13.
[44] D.S. Wishart, B.D. Sykes, F.M. Richards, The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy, Biochemistry 31 (6) (1992) 1647−1651.
[45] H. Frauenfelder, G.A. Petsko, D. Tsernoglou, Temperature-dependent X-ray diffraction as a probe of protein structural dynamics, Nature 280 (5723) (1979) 558.
[46] D. Picot, P.J. Loll, R.M. Garavito, The X-ray crystal structure of the membrane protein prostaglandin H2 synthase-1, Nature 367 (6460) (1994) 243.
[47] D.I. Svergun, M.V. Petoukhov, M.H. Koch, Determination of domain structure of proteins from X-ray solution scattering, Biophys. J. 80 (6) (2001) 2946−2953.
[48] R.W. Williams, Estimation of protein secondary structure from the laser Raman amide I spectrum, J. Mol. Biol. 166 (4) (1983) 581−603.
[49] J.T. Pelton, L.R. McLean, Spectroscopic methods for analysis of protein secondary structure, Anal. Biochem. 277 (2) (2000) 167−176.
[50] D. Kurouski, R.P. Van Duyne, I.K. Lednev, Exploring the structure and formation mechanism of amyloid fibrils by Raman spectroscopy: a review, Analyst 140 (15) (2015) 4967−4980.
[51] J. Kong, S. Yu, Fourier transform infrared spectroscopic analysis of protein secondary structures, Acta Biochem. Bbiophys. Sin. 39 (8) (2007) 549−559.
[52] D.M. Byler, H. Susi, Examination of the secondary structure of proteins by deconvolved FTIR spectra, Biopolymers 25 (3) (1986) 469−487.
[53] A. Villringer, U. Dirangl (Eds.), Optical Imaging of Brain Function and Metabolism 2, Springer, Berlin, 1997.
[54] A. Villringer, J. Planck, C. Hock, L. Schleinkofer, U. Dirnagl, Near infrared spectroscopy (NIRS): a new tool to study hemodynamic changes during activation of brain function in human adults, Neurosci. Lett. 154 (1-2) (1993) 101−104.
[55] Y. Ozaki, K. Murayama, Y. Wang, Application of two-dimensional near-infrared correlation spectroscopy to protein research, Vib. Spectrosc. 20 (2) (1999) 127−132.
[56] S.W. Bruun, I. Søndergaard, S. Jacobsen, Analysis of protein structures and interactions in complex food by near-infrared spectroscopy. 1. Gluten powder, J. Agric. Food Chem. 55 (18) (2007) 7234−7243.
[57] K.I. Izutsu, Y. Fujimaki, A. Kuwabara, Y. Hiyama, C. Yomota, N. Aoyagi, Near-infrared analysis of protein secondary structure in aqueous solutions and freeze-dried solids, J. Pharm. Sci. 95 (4) (2006) 781−789.
[58] M. Miyazawa, M. Sonoyama, Second derivative near infrared studies on the structural characterisation of proteins, J. Infrared Spectrosc 6 (1998) A253−A257.
[59] P. Robert, M.F. Devaux, N. Mouhous, E. Dufour, Monitoring the secondary structure of proteins by near-infrared spectroscopy, Appl. Spectrosc. 53 (2) (1999) 226−232.
[60] Y. Liu, R.K. Cho, K. Sakurai, T. Miura, Y. Ozaki, Studies on spectra/structure correlations in near-infrared spectra of proteins and polypeptides. Part I: a marker band for hydrogen bonds, Appl. Spectrosc. 48 (10) (1994) 1249−1254.
[61] C.A. Roberts, J. Workman, J.B. Reeves (Eds.), Near-Infrared Spectroscopy in Agriculture, Vol. 44, American Society of Agronomy, Madison, 2004.
[62] D.A. Burns, E.W. Ciurczak (Eds.), Handbook of Near-Infrared Spectroscopy, 3rd ed., CRC Press, Boca Raton, 2008.

[63] B.G. Osborne, Near-infrared spectroscopy in food analysis, in: R.A. Meyers, R.J. McGorrin (Eds.), Encyclopedia of Analytical Chemistry: Applications, Theory and Instrumentation, 2006.
[64] H. Martens, T. Næs (Eds.), Multivariate Calibration, Wiley, New York, 1989.
[65] B.G.M. Vandegiste, D.L. Massart, L.M.C. Buydens, S. de Jong, P.L. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics: Part B, Elsevier, Amsterdam, 1998.
[66] H. Mark, J. Workman Jr., Chemometrics in Spectroscopy, Elsevier, Amsterdam, 2007.
[67] R. Kramer, Chemometrics Techniques for Quantitative Analysis, CRC Press, Boca Raton, 1998.
[68] K. Varmuza, P. Filzmoser, Introduction to Multivariate Statistical Analysis in Chemometrics, CRC Press, Boca Raton, 2009.
[69] Y. Ozaki, I. Noda, Two Dimensional Correlation Spectroscopy, Melville, New York, AIP Conference Proceedings, AIP, Melville, NY, 2000, p. 503.
[70] J.R. Lane, H.G. Kjaergaard, XH-stretching overtone transitions calculated using explicitly correlated coupled cluster methods, J. Phys. Chem. 132 (2010) 174304.
[71] Y. Futami, Y. Ozaki, Y. Hamada, M.J. Wojcik, Y. Ozaki, Frequencies and absorption intensities of fundamentals and overtones of NH stretching vibrations of pyrrole and pyrrole-pyridine complex studied by near-infrared/infrared spectroscopy and density-functional-theory calculations, Chem. Phys. Lett. 482 (2009) 320–324.
[72] Y. Futami, Y. Ozaki, Y. Hamada, M.J. Wojcik, Y. Ozaki, Solvent dependence of absorption intensities and wavenumbers of the fundamental and first overtone of NH stretching vibration of pyrrole studied by near-infrared/infrared spectroscopy and DFT calculations, J. Phys. Chem. A 115 (2011) 1194–1198.
[73] S. Holly, O. Egyed, G. Jalsovszky, Assignment problems of amino acids, di-and tripeptides and proteins in the near infrared region, Spectrochim. Acta A Mol. Spectrosc. 48 (1) (1992) 101–109.
[74] W.Y. Yang, E. Larios, M. Gruebele, On the extended β-conformation propensity of polypeptides at high temperature, J. Am. Chem. Soc. 125 (52) (2003) 16220–16227.
[75] S. Ngarize, H. Herman, A. Adams, N. Howell, Comparison of changes in the secondary structure of unheated, heated, and high-pressure-treated β-lactoglobulin and ovalbumin proteins using Fourier transform Raman spectroscopy and self-deconvolution, J. Agric. Food Chem. 52 (21) (2004) 6470–6477.
[76] B. Yuan, K. Murayama, Y. Wu, R. Tsenkova, X. Dou, S. Era, et al., Temperature-dependent near-infrared spectra of bovine serum albumin in aqueous solutions: spectral analysis by principal component analysis and evolving factor analysis, Appl. Spectrosc. 57 (10) (2003) 1223–1229.
[77] P.Y. Furlan, S.A. Scott, M.H. Peaslee, FTIR-ATR study of pH effects on egg albumin secondary structure, Spectrosc. Lett. 40 (3) (2007) 475–482.
[78] T. Peters Jr (Ed.), All About Albumin: Biochemistry, Genetics, and Medical Applications, Academic Press, 1995.
[79] T. Peters Jr., Serum albumin, in: C.B. Anfinsen, J.T. Edsall, F.M. Richards (Eds.), Advances in Protein Chemistry, 37, Academic Press, 1985, pp. 161–245.
[80] J.N. De Wit, G.A.M. Swinkels, A differential scanning calorimetric study of the thermal denaturation of bovine β-lactoglobulin thermal behaviour at temperatures up to 100 °C, Biochim. Biophys. Acta Protein Struct. 624 (1) (1980) 40–50.
[81] M.Z. Papiz, L. Sawyer, E.E. Eliopoulos, A.C.T. North, J.B.C. Findlay, R. Sivaprasadarao, et al., The structure of β-lactoglobulin and its similarity to plasma retinol-binding protein, Nature 324 (1986) 383–385.
[82] M. Rüegg, U. Moor, B. Blanc, A calorimetric study of the thermal denaturation of whey proteins in simulated milk ultrafiltrate, J. Dairy Res. 44 (3) (1977) 509–520.
[83] B.J.M. Harmsen, W.J.M. Breem, On the conformation of bovine serum albumin after alkaline or thermal denaturation, Int. J. Protein Res. 1 (1969) 225–233.
[84] N.L. Sefara, N.P. Magtoto, H.H. Richardson, Structural characterization of β-lactoglobulin in solution using two-dimensional FT mid-infrared and FT near-infrared correlation spectroscopy, Appl. Spectrosc. 51 (1997) 536.

[85] C.P. Schultz, H. Fabian, H.H. Mantsch, Two-dimensional mid-IR and near-IR correlation spectra of ribonuclease A: using overtones and combination modes to monitor changes in secondary structure, Biospectroscopy 4 (1998) S19–S29.
[86] Y. Wang, K. Murayama, Y. Myojo, R. Tsenkova, N. Hayashi, Y. Ozaki, Two-dimensional Fourier transform near-infrared spectroscopy study of heat denaturation of ovalbumin in aqueous solutions, Phys. Chem. B 102 (34) (1998) 6655–6662.
[87] K. Murayama, T. Negawa, T. Hayashi, S. Era, Y. Wu, Y. Ozaki. Hydration study of human serum albumin in acidic pH region using two-dimensional near-infrared correlation spectroscopy, in: Analytical Sciences/Supplements Proceedings of IUPAC International Congress on Analytical Sciences 2001 (ICAS 2001), The Japan Society for Analytical Chemistry, 2002, pp. 667–670.
[88] R.K. Cho, J.H. Lee, J.J. Ahn, Y. Ozaki, M. Iwamoto, The applicability of near infrared reflectance spectroscopy for determining solubility and digestibility of heated protein under high pressure, J. Infrared Spectrosc. 3 (2) (1995) 73–79.
[89] H. Shinzawa, D. Ishikawa, M. Ishigaki, Y. Ozaki. Near-infrared imaging of polymers; from basic science to industrial applications, in: R.A. Meyers (Ed.), Encyclopedia of Analytical Chemistry, Wiley, 2016.
[90] W. Qi, Y. Suzuki, S. Sato, M. Fujiwara, N. Kawashima, S. Suzuki, et al., Enhanced interference-pattern visibility using multislit optical superposition method for imaging-type two-dimensional Fourier spectroscopy, Appl. Opt. 54 (20) (2015) 6254–6259.
[91] T. Kawajiri, K. Yanogawa, K. Yamamoto, M. Kondo, T. Harada, I. Ishimaru. 3-dimensional spectroscopic-tomography of biological membrane with high spatial resolution by the imaging-type 2-D Fourier spectroscopy, in: J.T. Wen, D. Hodko, Y. Otani, J. Kofman, O. Kaynak (Eds.), Optomechatronic Technologies International Society for Optics and Photonics, vol. 7266, 2008, p. 72660G.
[92] M. Ishigaki, NIR in vivo imaging of blood flow and molecular distribution in a developing fish egg are simultaneously obtained using imaging-type two-dimensional Fourier spectroscopy, NIR News 29 (6) (2018) 8–12.
[93] T. Iwamatsu, Stages of normal development in the medaka *Oryzias latipes*, Mech. Dev. 121 (2004) 605–618.
[94] I. Nanda, M. Kondo, U. Hornung, S. Asakawa, C. Winkler, A. Shimizu, et al., A duplicated copy of *DMRT1* in the sex-determining region of the Y chromosome of the medaka, *Oryzias latipes*, Proc. Natl Acad. Sci. U.S.A. 99 (18) (2002) 11778–11783.
[95] W.E. Hawkins, W.W. Walker, J.W. Fournie, C.S. Manning, R.M. Krol, Use of the Japanese medaka (*Oryzias latipes*) and guppy (*Poecilia reticulata*) in carcinogenesis testing under national toxicology program protocols, Toxicol. Pathol. 31 (2003) 88–91.
[96] E. Mochizuki, K. Fukuta, T. Tada, T. Harada, N. Watanabe, S. Matsuo, et al., Fish mesonephric model of polycystic kidney disease in medaka (*Oryzias latipes*) pc mutant, Kidney Int. 68 (1) (2005) 23–34.
[97] T. Sato, S. Kawano, M. Iwamoto, Near infrared spectral patterns of fatty acid analysis from fats and oils, J. Am. Oil Chem. Soc. 68 (11) (1991) 827–833.
[98] W. Hug, J.M. Chalmers, P.R. Griffith, Handbook of Vibrational Spectroscopy, John Wiley and Son Ltd, Chichester, 2002.
[99] T. Iwamatsu, T. Muramatsu, H. Kobayashi, Oil droplets and yolk spheres during development of Medaka embryos, Ichthyol. Res. 55 (2008) 344–348.
[100] C. Huynh-Delerme, M. Edery, H. Huet, S. Puiseux-Dao, C. Bernard, J. Fontaine, et al., Microcystin-LR and embryo-–larval development of medaka fish, *Oryzias latipes*. I. Effects on the digestive tract and associated systems, Toxicon 46 (1) (2005) 16–23.
[101] A. Monroy, M. Ishida, E. Nakano, The pattern of transfer of the yolk material to the embryo during the development of the teleostean fish, *Oryzias Latipes*, Dev. Growth Differ 6 (1961) 151–158.
[102] N. Grigorieff, T.A. Ceska, K.H. Downing, J.M. Baldwin, R. Henderson, Electron-crystallographic refinement of the structure of bacteriorhodopsin, J. Mol. Biol. 2593 (1996) 393–421.
[103] M.S. Weiss, G.E. Schulz, Structure of porin refined at 1.8 Å resolution, J. Mol. Biol. 227 (1992) 493–509.

[104] W.L. Smith, The eicosanoids and their biochemical mechanisms of action, Biochem. J. 259 (2) (1989) 315.
[105] S. Wold, K. Esbensen, P. Geladi, Principal component analysis, Chemom. Intell. Lab. 2 (1987) 37–52.
[106] S. Šašic, V.H. Segtnan, Y. Ozaki, Self-modeling curve resolution study of temperature-dependent near-infrared spectra of water and the investigation of water structure, J. Phys. Chem. A 106 (2002) 760.
[107] H. Okajima, M. Ando, H.O. Hamaguchi, Formation of "nano-ice" and density maximum anomaly of water, Bull. Chem. Soc. Jpn. (BCSJ) 91 (6) (2018) 991.
[108] V.H. Segtnan, Š. Šašić, T. Isaksson, Y. Ozaki, Studies on the structure of water using two-dimensional near-infrared correlation spectroscopy and principal component analysis, Anal. Chem. 73 (13) (2001) 3153–3161.
[109] H. Maeda, Y. Ozaki, M. Tanaka, N. Hayashi, T. Kojima, Near infrared spectroscopy and chemometrics studies of temperature-dependent spectral variations of water: relationship between spectral changes and hydrogen bonds, J. Infrared Spec. 3 (4) (1995) 191–201.
[110] B. Czarnik-Matusewicz, S. Pilorz, J.P. Hawranek, Temperature-dependent water structural transitions examined by near-IR and mid-IR spectra analyzed by multivariate curve resolution and two-dimensional correlation spectroscopy, Anal. Chim. Acta 544 (2005) 15.
[111] M. Ishigaki, A. Nakanishi, T. Hasunuma, A. Kondo, T. Morishima, T. Okuno, et al., High-speed scanning for the quantitative evaluation of glycogen concentration in bioethanol feedstock *Synechocystis* sp. PCC6803 using a near-infrared hyperspectral imaging system with a new near-infrared spectral camera, Appl. Spectrosc. 71 (3) (2017) 463–471.
[112] F. Westad, A. Schmidt, M. Kermit, Incorporating chemical band-assignment in near infrared spectroscopy regression models, J. Infrared Spectrosc. 16 (3) (2008) 265–273.
[113] J. Grabska, M. Ishigaki, K.B. Beć, M.J. Wójcik, Y. Ozaki, Correlations between structure and near-infrared spectra of saturated and unsaturated carboxylic acids. Insight from anharmonic density functional theory calculations, J. Phys. Chem. A 121 (18) (2017) 3437–3451.
[114] T. Sato, Application of principal-component analysis on near-infrared spectroscopic data of vegetable oils for their classification, J. Am. Oil Chem. Soc. 71 (3) (1994) 293–298.
[115] M.P. Sheetz, S.J. Singer, Biological membranes as bilayer couples. A molecular mechanism of drug-erythrocyte interactions, Proc. Natl Acad. Sci. U.S.A. 71 (11) (1974) 4457–4461.
[116] J.E. Johnson, R.B. Cornell, Amphitropic proteins: regulation by reversible membrane interactions, Mol. Membr. Biol. 16 (3) (1999) 217.
[117] T. Kitazaki, N. Kawashima, N. Yamamoto, H. Nomura, H. Kang, A. Nishiyama, et al., Parametric standing wave generation of a shallow reflection plane in a nonrigid sample for use in a noninvasive blood glucose monitor, J. Biomed. Opt. 24 (3) (2019) 036003.
[118] K.B. Beć, Y. Futami, M.J. Wójcik, Y. Ozaki, A spectroscopic and theoretical study in the near-infrared region of low concentration aliphatic alcohols, Phys. Chem. Chem. Phys. 18 (19) (2016) 13666–13682.
[119] C.W. Huck, K.B. Beć, J. Grabska, Y. Ozaki, Quantum chemical calculation of NIR spectra of practical materials, NIR News 28 (2) (2017) 13–20.
[120] C.G. Kirchler, C.K. Pezzei, K.B. Beć, S. Mayr, M. Ishigaki, Y. Ozaki, et al., Critical evaluation of spectral information of benchtop vs. portable near-infrared spectrometers: quantum chemistry and two-dimensional correlation spectroscopy for a better understanding of PLS regression models of the rosmarinic acid content in *Rosmarini folium*, Analyst 142 (3) (2017) 455–464.
[121] J. Grabska, K.B. Bec, M. Ishigaki, C.W. Huck, Y. Ozaki, NIR spectra simulations by anharmonic DFT-saturated and unsaturated long-chain fatty acids, J. Phys. Chem. B 122 (27) (2018) 6931–6944.
[122] J. Grabska, K.B. Beć, M. Ishigaki, M.J. Wójcik, Y. Ozaki, Spectra-structure correlations of saturated and unsaturated medium-chain fatty acids. Near-infrared and anharmonic DFT study of hexanoic acid and sorbic acid, Spectrochim. Acta A 185 (2017) 35–44.
[123] A. Ikehata, T. Itoh, Y. Ozaki, Surface plasmon resonance near-infrared spectroscopy, Anal. Chem. 76 (21) (2004) 6461–6469.
[124] A. Ikehata, T. Itoh, Y. Ozaki, Quantitative analyses of absorption-sensitive surface plasmon resonance near-infrared spectra, Appl. Spectrosc. 60 (7) (2006) 747–751.

CHAPTER 7

Vibrational imaging of proteins: changes in the tissues and cells in the lifestyle disease studies

Katarzyna Majzner[1,2], Krzysztof Czamara[2], Marta Z. Pacia[2], Jakub Dybas[1,2], Ewelina Bik[1,2], Karolina Chrabaszcz[1,2], Ewelina Wiercigroch[1,2], Aleksandra Dorosz[1,2], Adrianna Wislocka[1,2], Katarzyna M. Marzec[2], Kamilla Malek[1,2] and Malgorzata Baranska[1,2]

[1]Faculty of Chemistry, Jagiellonian University, Krakow, Poland, [2]Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

7.1 Introduction

Proteins are undoubtedly one of the most essential groups of macromolecules with a high variability of structures and a plethora of functions, establishing them as dominant players in the cells of every living organism. According to IUPAC, the term "proteins" relates to all naturally occurring and also synthetic polypeptides having molecular weights greater than about 10,000 u, however, the limit is not precise [1]. The tremendous heterogeneity of proteins is additionally emphasized by their numerous occurrence, for example, it is estimated that one Henrietta Lacks (HeLa) cell has 1−3 billion protein molecules [2]. Structurally, they are formed in four levels of self-organization. Most proteins consist of linear polymers built from a series of up to 20 different L-α-amino acids. The order of the amino acid residues in the polypeptide chain, that is, their sequence, is called the primary structure of the peptide or protein, which is the first step to determine the structure of a given protein. The complete understanding of the biological properties and functions of proteins requires three-dimensional structural information. Polypeptide chains have the ability to form complex spatial structures as a result of, for example, the interactions of hydrogen bonds between amino acid residues in different parts of the same chain or between different chains. This spatial arrangement of the amino acids chains is called the protein secondary structure. The two basic models of the protein secondary structure named as α-helices and β-sheets. Other extended structures, such as turns, loops, and random coil conformations, link important protein folding intermediates. The presence of specific amino

Vibrational Spectroscopy in Protein Research.
DOI: https://doi.org/10.1016/B978-0-12-818610-7.00007-4

acids in the polypeptide chains, for example, cysteine or arginine, allows the formation of an advanced spatial organization (tertiary structure) through the interaction, inter alia, by disulfide bridges and electrostatic interactions between oppositely charged $-NH_3^+$ and $-COO^-$ groups of side fragments of amino acids. The highest level of protein structure organization, the quaternary structure, refers to the number and arrangement of the protein subunits because many proteins are assemblies of multiple polypeptide chains. Examples of proteins with quaternary structure include hemoglobin, alcohol dehydrogenase, and ion channels.

Biosynthesis of a protein with a native spatial structure, which determines a specific biological function, is a complex process. Moreover, proteins are very fragile macromolecules to any environmental changes including heat or pathology development. In consequence, protein molecules can adapt wrongly or change the conformation during the folding process resulting in the loss or gain of different functions. There are many human diseases that are protein folding-related, for example, cystic fibrosis, or Huntington's disease [3]. Some pathogenesis, called amyloidosis, are associated with the aggregation of incorrectly folded polypeptides. The representatives of these types of pathology are Alzheimer's disease (AD), amyotrophic lateral sclerosis, and Parkinson's disease (PD) [4]. Lifestyle diseases, like hypertension or type II diabetes, are also manifested by the alteration of protein structure and their distribution, and this topic is further discussed below (vide infra).

Although vibrational spectroscopy is not a leading technique to study proteins in pathologies and diseases, it is very informative and can be used in a noninvasive way. With the use of Raman and infrared (IR) spectroscopies, several vibrational modes can be applied for the analysis of a protein's structure and its alterations. Most characteristic protein bands originate from the -CONH- group, referred to amide A and B (NH stretching, at *ca.* 3500 and 3100 cm^{-1}, respectively) and I to VII amide bands (I: 1690–1600 cm^{-1}, II: 1580–1480 cm^{-1}, III: 1300–1230 cm^{-1}, IV: 770–625 cm^{-1}, V: 800–640 cm^{-1}, VI: 600–540 cm^{-1}, VII: *ca.* 200 cm^{-1}) [5]. Some of them are useful for the identification of different protein backbone confirmations including amide I (arising mainly from C = O stretching), amide II, and amide III (both associated with coupled C–N stretching and N–H bending vibrations of the peptide group) [5]. Raman spectra of each protein exhibit characteristic fingerprint that can be used for molecular identification, however, the spectral profile strongly depends on the laser excitation wavelength. In Fig. 7.1 the Raman spectra of glucose oxidase measured with three lasers, 488, 532, and 1064 nm, are presented.

Fig. 7.1 shows that by selecting an appropriate excitation wavelength the different information about the structure of proteins can be obtained, or in some cases Raman signals are completely overlapped by fluorescence background. In comparison to lipids, the Raman spectrum of proteins is relatively weaker. The intensities of Raman signals can be

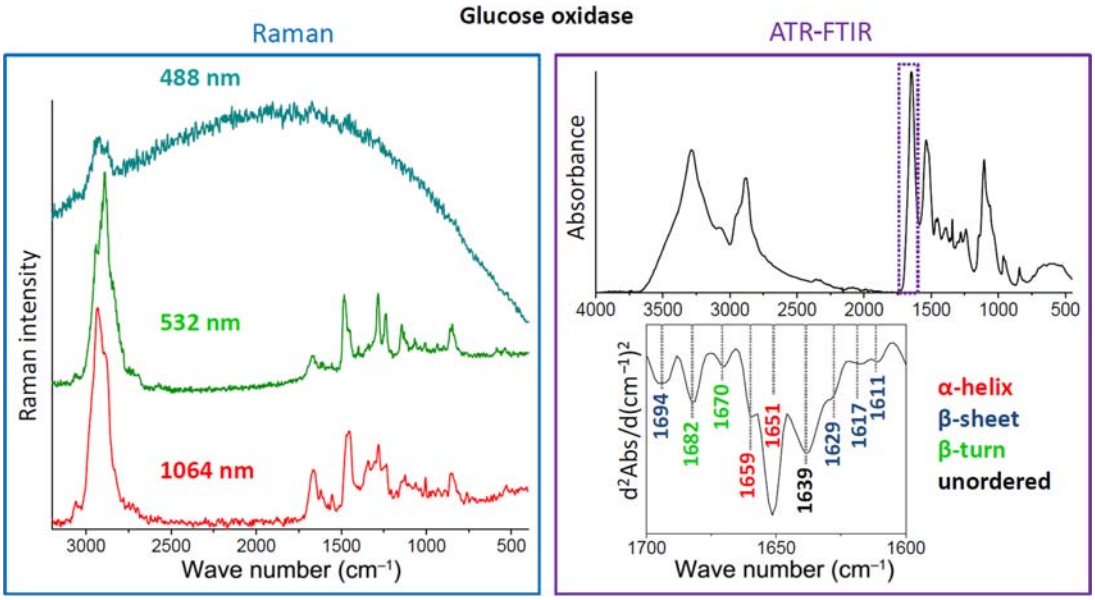


Figure 7.1

Raman and Fourier-transform infrared spectroscopy (ATR-FTIR) spectral profile of glucose oxidase. The spectra were measured by three laser lines: 488, 532, and 1064 nm. The second derivative ATR-IR spectra present the resolved protein secondary structure. Source*: Adapted with permission from A. Rygula, K. Majzner, K.M. Marzec, A. Kaczor, M. Pilarczyk, M. Baranska, Raman spectroscopy of proteins: a review, J. Raman Spectrosc. 44 (8) (2013) 1061—1076. John Wiley and Sons.*

augmented when the exciting laser wavelength is adjusted to the absorption range of the particular chromophore within the sample and thus the resonance enhancement takes place. Resonance Raman spectroscopy (RRS) is widely used to study hemeproteins [6] and systems rich in hemeproteins, that are, red blood cells [7,8] or brain and heart tissue (vide infra). Another way to intensify Raman signal involves metal (plasmon) enhancements. Tip-enhanced Raman spectroscopy (TERS) technique can be applied to resolve the sequence of amino acids of protein primary structure [9,10]. Lipiec et al. show that TERS technique is also valuable in nanoscale chemical mapping of single amyloid fibrils at various stages of aggregation [11]. Vibrational spectroscopy is much more valuable for the investigation of protein secondary structure. For Raman spectroscopy (RS), the spectral region of amide III bands is sensitive to changes in the proteins' secondary structure and the shift of the maximum of this band toward the lower wave numbers is identified with the increased β-sheet conformation and disordered structure in comparison to the α-helix. In many diseases, for example, related to the formation of amyloid (i.e., diabetes type II, AD or PD) and hypertension (vide infra), a similar change in the secondary structure of proteins is observed [4]. On the other hand, the second derivatives of FTIR spectra in the amide I spectral range (Fig. 7.1) enables determination of the secondary structure of the proteins

present in the biological sample, for example, in the vascular wall of diabetic animals [12]. FTIR spectra also provide the qualitative and quantitative information on protein distribution and their alterations arising from pathology development enables identification of early micrometastases in lung tissue [13,14]. Details of model studies on the aorta wall and the spectral characterization of pulmonary metastases in a murine model of breast cancer are presented below.

7.2 Raman in vitro studies of the cell apoptosis

The application of various Raman microscopy techniques in the drug–cell interactions can shed a new light on the complex drug toxicity mechanisms and cell biochemistry, which stands behind them. It is very important to understand the mechanism of cell death, due to the fact that every alteration from their proper functioning may lead to the development of pathology, for example, heart failure, acute lung injury, or cancer formation [15]. Currently, for the investigation of an occurrence of the apoptosis in cells, mostly invasive and expensive biochemical assays are applied. The well-known way to establish the cell apoptosis is to evaluate, whether the caspase 3 and 6 has been activated. Among the morphological changes occurring in cells due to the apoptosis, cell detachment, shrinkage, DNA fragmentation, and phosphatidylserine externalization can be distinguished [16]. Apoptosis in cells can be mediated by two main pathways, that is, mitochondrial (an intrinsic pathway) and receptor (an extrinsic pathway) [17]. When the cell is under a stress conditions, the activation of proapoptotic proteins (Bak and Bax) can lead to the increased permeability of a mitochondrial membrane [18]. As a consequence, the release of proapoptotic factors from mitochondria may be induced, resulting in caspase cascade, DNA fragmentation, and chromatin condensation.

Many cellular events involving changes in protein metabolism may be triggered by an activation of specific membrane receptors, for example, inflammation can be induced by the binding of a ligand, for example, TNF-α (or bacterial lipopolysaccharides), to a respective receptor TNFR-1 (TLR4) and apoptosis by a linkage of FasL to Fas receptor. All these factors trigger the activation of the classical nuclear protein NF-κB pathway, which provides information into the cell nucleus, where the response is processed. Binding of the FasL or TNF-α causes subsequent activation of caspases, protein-degrading enzymes, leading to apoptosis [19]. Alternative pathways of apoptosis induction are also possible, for example, involving the contribution of mitochondria that release cytochrome *c* to the cytoplasm, which binds to the cytosolic protein Apaf-1 to facilitate the formation of the apoptosome, activates caspase 9 and triggers caspase cascade or as a result of disturbance of Ca^{2+} concentrations or redox balance [20]. Cell death may also be initiated by incubation with some chemical compounds, for example, cycloheximide (CHX). CHX belongs to the group of antibiotics and is extracted from the *Streptomyces griseus* bacteria.

Its proapoptotic action is based on inhibiting the translation elongation stage of protein synthesis by joining a large ribosomal subunit. As a result, the cell has a disturbed protein metabolism, which ultimately leads to its death. The action of protein synthesis inhibitors in the process of apoptosis is of a dual nature. On the one hand it inhibits the production of protective proteins, on the other, it limits the expression of caspases. It seems, however, that the first effect is dominant and affects the survival of cells [21].

Apoptosis, or programmed cell death, under normal physiological conditions is responsible for maintaining homeostasis and the process of tissue remodeling. It is characterized by morphological changes in intracellular structures, including condensation and fragmentation of nuclear chromatin, formation of nodules in the cell membrane, disruption of cytoskeleton architecture, cytoplasm, and nucleus shrinkage [22,23]. A lot of spectroscopis studies has been performed to distinguish apoptotic from nonapoptotic cells and to characterize the process of apoptosis. Raman and IR spectroscopy have been successfully used to study the process of apoptosis in various types of cells and under the influence of exposure to various factors, for example, chemical compounds [24–26], including cytotoxic drugs [27–29], or harmful ultraviolet radiation [30,31].

Chemically induced apoptosis can be investigated with the success by means of Raman imaging. An example of studies at the subcellular level involve glyoxal-induced changes observed in control and stressed human embryonic lung fibroblast cells (L132) [32]. The cells were exposed to glyoxal for 24 hours in order to trigger an oxidative stress followed by the apoptosis. Glyoxal cytotoxicity results from the inhibition of protein and DNA synthesis. The spectra of nucleus, cytoplasm, cytoplasmic inclusions with increased RNA contributions, and lipid vesicles were examined. Raman images revealed the decreased size of the nucleus in late apoptotic cells. Additionally, treated cells exhibited more inclusion bodies than control cells. Furthermore, in the Raman spectra of cell nuclei, the decrease in the intensity of a DNA band (786 cm^{-1}) was observed. In the Raman spectrum of cytoplasm of controlled cells the intensity ratios of the most intensive RNA and protein bands near 782 and 1003 cm^{-1} was higher [32].

In our research, the early stage of apoptosis in EA.hy926 endothelial cells induced via two different pathways by FasL and CHX were studied [33]. The chemical changes associated with the early apoptosis within major cellular structures, that is, nuclei, nucleoli, cytoplasm, and endoplasmic reticulum (ER), were determined. For both investigated factors (FasL and CHX), the results showed a significant decrease in the protein content in all above-mentioned organelles of the cell simultaneously with a significant increase in the DNA/RNA signal in nuclei and nucleoli. Spectroscopic observations have been complemented with microscopic images enabling tracking of morphological changes, that is, the involution of capillary-like structures on the cell membrane. As a result, it was possible to characterize alterations associated with an early apoptosis, independently of its pathway.

The knowledge about the mechanism of apoptosis is considerably important for anticancer drug development. Much of the information is focused on the study of the effect of anticancer drugs and their chemotherapeutic response in anticancer therapy [28,34] or dedicated to the classification and differentiation of cells in apoptosis [35,36], as well as a detailed description of biochemical changes in the apoptotic process [37–40]. It is believed that an understanding of this process should enable the introduction of new drugs inducing cancer cells' apoptosis without interfering with healthy cells.

In Ref. [34] the influence of 5-fluorouracil (5-FU, an anticancer, proapoptotic drug) on gastric human carcinoma cells was determined. The authors presented the application of nondestructive Raman spectroscopy to monitor the apoptosis of stomach cancer cells and to test the effectiveness of the treatment of new drugs on the various cancer cells. The spectra of cells incubated with 5-FU with concentration at the level of 25 μg/mL exhibited lower protein (that is, 1576 and 1655 cm^{-1}) and nucleic acids signal (782, 1092, 1320, and 1340 cm^{-1}) in comparison to control cells. The decrease in the intensity of protein Raman bands in comparison to untreated cells can be explained by the cell fragmentation due to the apoptosis [34].

A monitoring of the apoptotic response of individual cells to etoposide over time was investigated on human medulloblastoma cell line (DAOY) [41]. The mode of an action of this chemotherapeutic agent is dependent on DNA and topoisomerase II inhibition (mostly in cancer cells) leading to apoptosis. The Raman spectra were acquired from the same DAOY cell after 12, 24, 36, and 48 hours of incubation with 10 μM etoposide. After 24 hours the decrease in the intensity of a DNA band at ca. 788 cm^{-1} was observed suggesting the disruption of that component. After 36 hours, the decrease in an overall Raman bands' intensity, particularly for phenylalanine (1005 cm^{-1}), and protein and nucleic acids (1342 cm^{-1}) was observed. The authors postulated the decrease in the intensity of those two bands as a biomarker of apoptosis [41].

For nonchemical factors, other than ultraviolet light, an induction of cell apoptosis or necrosis due to an exposure to room temperature (RT) or overheating, respectively can be included. In studies of Brauchle et al [35]. RT-induced apoptosis and heat-triggered necrosis were analyzed in individual Saos-2 and SW-1353 cells by utilizing Raman microscopy. In order to induce both early and late apoptosis, Saos-2 and SW-1253 cells were exposed at RT for 4 and 7 days, respectively, and the Raman measurement was performed subsequently. In turn, the necrosis was induced by heating the cells up to 55°C. In the Raman spectra of the early apoptotic state, the 1375 cm^{-1} Raman band occurred, and in the case of late apoptotic cells this band increased significantly. The intensity of nucleic acids marker band observed in the 780–800 cm^{-1} spectral range decreased during apoptosis, which is in agreement with previously reported results suggesting that the decrease of the 785 cm^{-1} band may correlate with the progress of internucleosomal DNA

cleavage [37]. In late apoptotic cells an increased contribution of lipids (1437 cm^{-1}) was correlated to reduced protein signals (1003 and 1658 cm^{-1}), which in turn can be related to active intracellular apoptotic proteases. On the other hand, the increase in the intensity of band at 1375 cm^{-1} (nuclear condensation) and phenylalanine marker band (1003 cm^{-1}) was observed in the case of necrotic cells along with the decrease in the intensity of an amide I (1658 cm^{-1}) band. Those results suggest that Raman microscopy can be used as a platform to detect and discriminate apoptotic and necrotic cell death since the multivariate classification model successfully identified viable, RT-induced early and late apoptotic cells as well as heat-treated necrotic cells based on Raman spectra [35].

Markers of apoptosis can be identified also in the cytochrome signal. Okada et al. [42] investigated dynamic changes in cytochrome *c* distribution at the Raman band of 750 cm^{-1} after adding an apoptosis inducer to the cells. The Raman band at 750 cm^{-1} is one of the marker bands for hemeproteins in resonance Raman scattering (excitation at 532 nm) and is assigned to pyrrole breathing mode ν_{15}. HeLa cells were incubated with an acinomycin D, which is an apoptosis inducer and inhibits RNA synthesis. Time-lapsed Raman images were acquired after 7 hours of a treatment with subsequent 5 minutes intervals. After 25−35 minutes of the treatment, the change in distribution of the band at 750 cm^{-1} and the decrease in the intensity of cytochrome bands were observed. It was attributed to a release of cytochrome *c* from mitochondria to cytosol. Changes in both distribution and intensity of the signal were observed before any morphological modifications occurred. Additionally, the distributions of proteins have shown no significant differences in either sample, and provided further evidence that the cytochrome *c* distribution changed independently from the cell morphological changes [42].

Alterations of the cytochrome *c* signal were also studied by Salehi et al. [43]. In this case apoptosis was monitored in vitro through a diffusion of cytochrome *c* in the cytoplasm. Changes were observed and the apoptosis condition was induced by a chemotherapeutic agent, paclitaxel. Raman images of cytochrome *c* distribution within cells taken from the cells treated with a lower concentration of paclitaxel, overlapped with the mitochondria position extracted from the cluster analysis. The presence of cytochrome *c* signal outside of mitochondria can be correlated with the start of apoptosis via a mitochondrial apoptosis pathway independent of the caspase 8/t-Bid pathway [43].

7.3 An effect of fixation on endothelial cells

Studies of living cells are always very desirable since allow to reflect their physiological condition. Measurements of live cells are a very helpful and important step for better understanding of functional processes investigated in the cells and very often is crucial in pharmacological studies. However, this approach is not always possible, because of relatively short cell lifetime and dynamic changes in cellular biochemistry due to life

processes occurring in them. The fixation is used to preserve cellular components, their biochemical composition and structure, and prevent from autolysis and bacterial decomposition. Fixation protocols are usually developed to keep biological samples, as close to in vivo conditions as possible. The most required are sample preparations protocols causing no significant biochemical alteration, preserving physiological and mechanical properties of cells and tissues. Changes occurred upon the fixation can be easily detected and investigated by Raman spectroscopy due to its high sensitivity on biochemical changes in biosamples.

Chemical fixation methods can be classified by their chemical nature into inter alia aldehydes, alcohols, and oxidizing agents, which in turn are related to their actions, such as cross-linking or dehydration. One of the most popular approaches is aldehyde fixation. Aldehyde fixation stabilizes structural details of cells for further examination by microscopic techniques. The major mechanism which explains fixative properties of aldehydes is formation of intermolecular bridges of the aldehyde groups with reactive groups of proteins (i.e., primary amines, amides, hydroxyl groups, sulfhydro groups) in cells and tissue. The literature has evaluated the effect of fixation methods for particular cells and tissues. Glutaraldehyde is one of the aldehyde-based fixatives that provoke cross-linking in proteins and nucleic acids inside cells. One of the advantages of glutaraldehyde over paraformaldehyde is the effectiveness of a DNA fixation and providing higher stiffness of the cell. Despite the fact that fixation can alter the proteins' composition, it is commonly used when experimental techniques cannot be performed on live cells [44].

In our research glutaraldehyde fixation is a standard approach. To study the influence of glutaraldehyde on the biochemical content of human aortic endothelial cells were selected. Cells were fixed for 4 minutes with 2.5% glutaraldehyde and were washed three times with PBS. Raman imaging was performed with 488 nm excitation wavelength, integration time of 0.75 second, and 2 μm step size. Then the cluster analysis (KMC) of the dataset was performed in order to separate information regarding specific organelles and cell areas. Applied cluster analysis allowed obtaining averaged spectra assigned to particular classes: nucleus (marked in blue), ER (green), and cytoplasm (orange). Fig. 7.2 shows color coded spectra from live and fixed cells.

By comparing the averaged spectra of live and fixed cells a few of the most prominent differences can be noticed. The band that appears in spectrum of every class as a result of fixation can be found at $\sim 1039\ cm^{-1}$. The band that increases its intensity for the fixed cells in comparison to the live ones is located at $1258\ cm^{-1}$ (amide III). The ER and cytoplasm spectra from the live cells exhibits a higher intensity of the band at $3013\ cm^{-1}$ in compare to fixed cells. Furthermore, the $966\ cm^{-1}$ band occurs in the ER and cytoplasm spectra, and together with the Raman band at $1039\ cm^{-1}$ indicate the effect of fixation on spectroscopic signature of cells. In the spectra from the nucleus and cytoplasm the

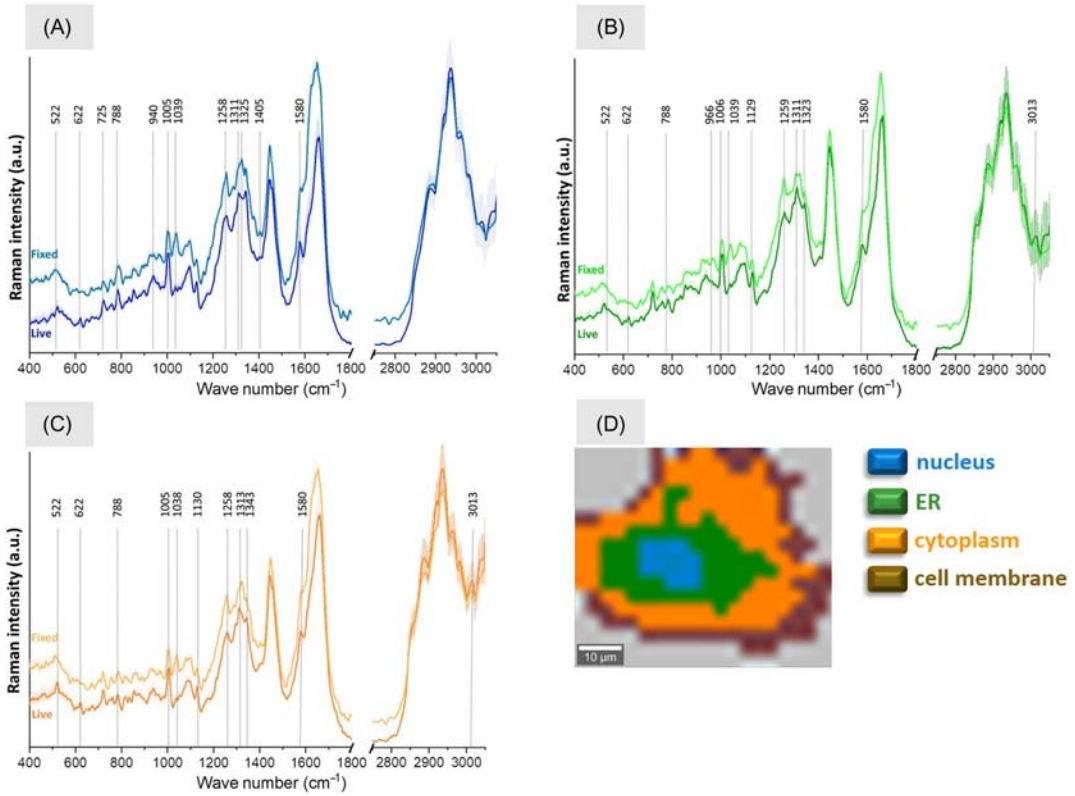


Figure 7.2

The Raman spectra of three main cell clusters corresponding to the cellular region rich in nucleic acids (marked in blue, A), the endoplasmatic reticulum area, rich in phospholipids (green, B), and the cytoplasm, rich in proteins (orange, C) of fixed (brighter shades of colors) and live (darker shades of colors) endothelial cells. All spectra were recorded using the 488 nm laser and were averaged over spectra of 10 cells obtained by means of k-means cluster analysis (KMC) (D). Spectra were presented with the standard deviation on each data point (shadowed area). Spectra were background subtracted, normalized, shifted vertically and maximally extended in the *y*-axis for clarity.

1286 cm^{-1} band is of higher intensity. In general, the spectra of fixed cells exhibit lower intensity in comparison to the Raman spectra measured from live cells. Changes in protein signal can be rather related not to a loss of compounds but to changes in a conformation and structure, which in turn have been attributed to the cross-linking action of the aldehyde.

7.4 Blood plasma proteins and their diagnostic perspectives

Blood circulating throughout the body integrates others tissues and organs and provides a special means of communications. Therefore chemical changes in blood components can reflect many diseases, which is what makes this body fluid useful for clinical diagnosis. The

intracellular matrix of the blood is plasma, which makes up about 55% of the total blood volume and contains mostly water (~90%), dissolved proteins (albumin, globulins, and fibrinogen: ~8%) and organic and inorganic compounds [45]. As Fourier-transform infrared (FTIR) spectroscopy provides information not only about the composition of biological samples but also about the structural changes of macromolecules observed in the complex matrix, this technique has been widely adopted for special applications in biochemistry. Because of the high signal-to-noise ratio in FTIR spectra even in aqueous phase systems blood plasma can be investigated in terms of the identification of disease, its treatment, and induced molecular changes in proteins, nucleic acids, and lipids. For example, Nyden et al. analyzed mixtures of human plasma proteins in aqueous solution to recognize their FTIR signature [46]. An attenuated total reflection-Fourier-transform infrared spectroscopy (ATR-FTIR) study of human plasma components, including the determination of the total content of proteins, was published by Janatsch et al. [47]. High sensitivity of the technique was confirmed on the basis of calibration curves measured for pure reference substances. ATR-FTIR spectroscopy was also employed to establish a model for quantification of albumin, immunoglobulin, total globulin, and albumin/globulin ratio in serum samples. These studies proved the usefulness and applicability of FTIR spectroscopy in the determination of proteins as biochemical parameters in clinical samples like blood plasma [48].

The major advantage of FTIR spectroscopy in studies on bodily fluids is a minimal volume of sample (0.5−1 μL) which is required to gather biochemical composition, including proteins, lipids, carbohydrates, and DNA/RNA [49]. There is practically no limit in the used techniques, that is, ATR, transmission, and transflection modes, but an important factor is a way of depositing the plasma sample on IR windows to provide a constant pathway of sample penetration. In addition, the electric field standing wave effect can lead to significant distortions of transflection FTIR spectra [50,51]. Apart from changes in bands ratios, transflection FTIR spectra of plasma show a significant shift of amide bands correlated with the sample thickness contrary to transmission mode. In consequence, the interpretation of secondary structures of proteins due to disease or stress factors might be false. Staniszewska et al. reported that transmission FTIR spectra of blood plasma exhibit the presence of amide I and II bands at 1658 and 1547 cm^{-1}, respectively, whereas they are up- or downshifted for thick and thin plasma deposits, respectively, in transflection [49].

Alterations in plasma protein composition can be a potential marker of many diseases. Staniszewska-Slezak and coworkers proved by the examination of murine models of civilization diseases that changes in the protein content and or secondary structures are indicators of the appearance of pulmonary hypertension (PAH) and renovascular hypertension, diabetes, and the early phase of systemic hypertension [52−54]. A procedure of plasma samples preparation, data collection, and data processing were optimized and employed in a similar manner in all the above-mentioned reports, see Fig. 7.3. In brief,

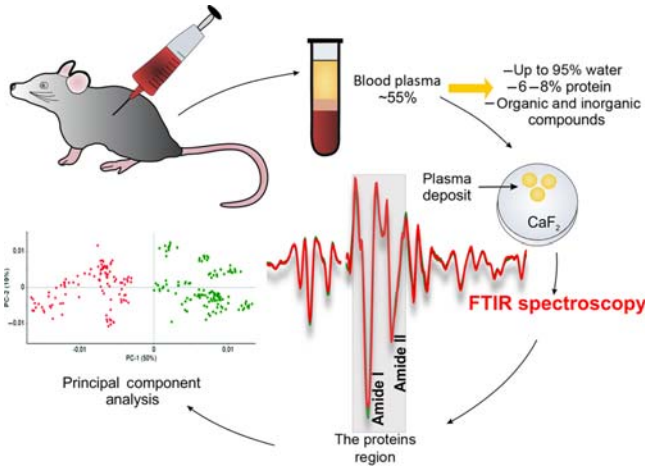


Figure 7.3
Schematic process of obtaining information about plasma protein using Fourier-transform infrared (FTIR) spectroscopy: from blood collection to data analysis.

blood from mice or rats with induced diseases was collected from the right ventricle of the anesthetized animals to the syringe with anticoagulant. Blood samples were centrifuged at 1000g for 10 minutes and platelet-poor plasma (ppp) was immediately separated. Next, three replicates of the ppp sample (0.5 μL) were spotted on CaF_2 windows and left to dry to avoid water influence on amide I and II bands. Then FTIR images were recorded in transmission mode. FTIR spectra collected in such a way are free of distortion and artefacts. An analysis was performed on second derivative IR spectra and included the comparison of bands positions, bands intensities, and their ratios between the experimental groups. The main tool used for discrimination analysis was principal component analysis (PCA) [52–54].

An interesting study was the comparison of FTIR profiles of plasma delivered from rat models of two types of hypertension of a different origin, that is, PAH induced by injection of monocrotaline and renovascular (systemic) hypertension induced by clipping of the left renal artery [52]. It is worth highlighting here that, while systemic hypertension can be diagnosed and monitored by a simple blood pressure measurement, the diagnosis of PAH is more complex and requires exercise tolerance, laboratory tests, magnetic resonance imaging, and echocardiography to assess right heart hypertrophy [55]. So there is a need to establish new analytical methods which recognize the disease or give an insight into molecular changes in biological molecules induced by biochemical processes. FTIR spectroscopy indicated that both diseases exhibit an unique spectral signature resulting from biochemical changes reflected in blood plasma and these changes are mainly associated with the protein content and modifications in protein

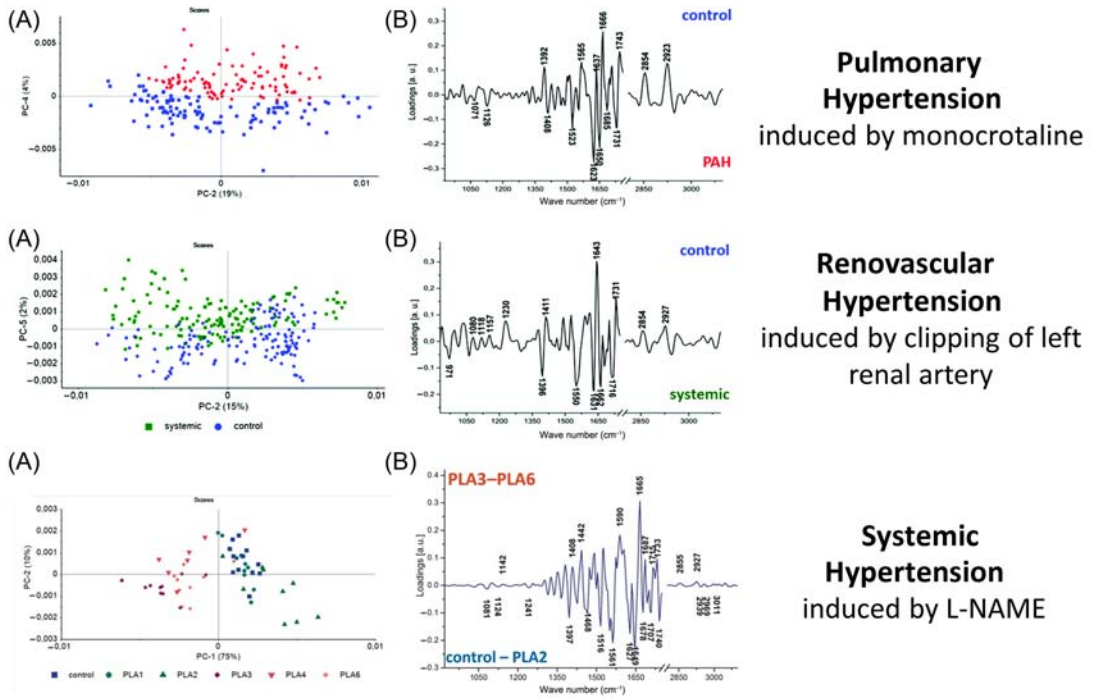


Figure 7.4

Results of principal component analysis (PCA) applied to second derivative FTIR spectra in the regions of 3100−2800 and 1800−900 cm^{-1} for pulmonary (*top*), renovascular (*middle*), and systemic (*down*) hypertensions; (A)—score plots and (B)—loadings plots. Source*: Reproduced from E. Staniszewska-Slezak, et al., Plasma biomarkers of pulmonary hypertension identified by Fourier transform infrared spectroscopy and principal component analysis, Analyst 140 (7) (2015) 2273−2279. with permission from The Royal Society of Chemistry and from E. Staniszewska-Slezak, L. Mateuszuk, S. Chlopicki, M. Baranska, K. Malek, Alterations in plasma biochemical composition in NO deficiency induced by L-NAME in mice analysed by Fourier-transform infrared spectroscopy, J. Biophotonics 9 (10) (2016) 1098−1108, with permission from John Wiley and Sons.*

conformations. It must be stressed here that the FTIR spectrum of a complex biological matrix does not detect a specific protein but rather indicates what groups of proteins are altered due to disease or stress factors. PCA analysis of FTIR spectra of plasma derived from the rat models of pulmonary and renovascular hypertensions discriminated both groups from the control (Fig. 7.4, top and middle panels). The main discriminators were observed in the region of 1500−1750 cm^{-1} assigned mainly to amide I and II bands of proteins and this fact was associated with posttranslational modifications of plasma proteins, for example, tyrosine nitration, protein phosphorylation. For PAH, an increase of the protein content with dominated conformation of α-helices (IR band at 1650 cm^{-1}), extended chains and turns (bands at 1624 and 1685 cm^{-1}, respectively) were proposed as

spectral markers [52]. In addition, the PCA loading plot showed the presence of the discriminator attributed to tyrosine-rich proteins (a band at 1523 cm^{-1}) accompanied by conformational changes of proteins toward β-pleated sheets (Fig. 7.4, top panel). In the case of the FTIR profile of renovascular hypertension, the most intensive PCA loadings suggested the production of proteins and/or polypeptides built of unordered secondary structures and extended chains (amide I bands at 1666 and 1632 cm^{-1}, respectively) which suggest an increase of hypertension in plasma (Fig. 7.4, middle panel) [52]. The authors suggested that these changes may be associated with the phosphorylation process of proteins [52].

A typical systemic hypertension and increased peripheral resistance associates nitric oxide deficiency, induced by prolonged treatment with N^G-nitro-L-arginine methyl ester (L-NAME) administrated to mice [54]. A statistically significant increase of blood pressure was observed after the sixth week of L-NAME administration. FTIR spectra of plasma collected within 1−6 weeks Post L-NAME Administration (PLA 1−6) of the inhibitor application segregated the control, PLA 1 and 2 groups from others, see Fig. 7.4. Spectral differences found in the loading plot of PCA analysis suggested that amide I and amide II bands were again the main discriminators distinguishing consecutive weeks of pathology. In general, it was shown that changes in proteins and amino acids are the most prominent ones occurring during the progression of the pathology, cf. details in Ref. [54]. Despite the fact that the total content of proteins (determined as total absorbance of amide I and amide II bands) was found to be constant within 6 weeks of the disease progression, a detailed analysis of FTIR spectra showed dynamics of changes in protein structures. The highest positively loadings correlated with the late phase of NO deficiency were bands assigned to unordered and antiparallel β-sheet secondary structures of proteins accompanied by a decreased content of tyrosine-rich proteins. In turn, the early stages were characterized by a high content of tyrosine-rich proteins. What is important here is that the spectral profile of plasma of L-NAME-induced hypertension is completely different from profiles of PAH and renovascular hypertensions [52,54].

Diabetes is one of the most common causes of disability and mortality. Dysfunction of the endothelium accompanying this disease increases the risk of atherothrombosis and its cardiovascular complications. The basic diagnostic tool for diabetes is a measure of the glucose level in blood. However, there are no simple clinical methods which could determine cardiovascular complications of diabetes and follow the efficiency of treatment. A FTIR spectroscopic study on plasma delivered from a murine model of type 2 diabetes mellitus (T2DM) proposed additional markers which could be used in clinical diagnostics [53]. Using FTIR spectra it was possible to track biochemical changes specific for T2DM and also to observe reversal endothelial dysfunction due to treatment with an angiotensin-converting enzyme inhibitor (perindopril) showing vasoprotective activity. The authors observed the alternation of the amide II/amide I bands ratio attributed to the β-sheet

formation leading to aggregation of proteins as fibrils in plasma of diabetic mice. Treatment with perindopril at a high dose but not at a low dose reversed this effect. In addition, they proposed the ratio of the tyrosine band (at 1517 cm^{-1}) and the vibration of the phosphate groups (at 1240 cm^{-1}) as a potential marker of the phosphorylation process which occurs in endothelial dysfunction [53].

7.5 Protoporphyrin proteins in leukocytes

Although blood plasma is an acceptable biomaterial, blood cells constitute the important target in clinical diagnosis. But the most common analysis relies only on counting them. Among blood cells, white blood cells (WBCs and leukocytes) play an important role in the body's immune system. They are responsible for fighting bacterial, viral, and fungal infections and accompany an inflammatory state. All leukocytes are produced in the bone marrow and account for $\sim$1% of blood cells [56]. WBCs can be divided into two groups: the first group has granules (granulocytes) and the second does not possess granules in their intracellular environment (agranulocytes) (Fig. 7.5). In this part of the chapter we focus on the proteins of granulocytes, eosinophils and neutrophils, that are easily examined using RS.

Eosinophils are acidophilic granulocytes, which are regular and round cells with a diameter of 10–20 μm and a characteristic morphological feature that are bilobed nuclei, cf. Fig. 7.5 [57,58]. The eosinophil cell contains a pool of specific proteins in its cytoplasmic granules that are eosinophil peroxidase (EPO), eosinophil cationic proteins, and eosinophil-derived neurotoxin [59]. These cells release the content of the granules due to several environmental stimuli. The number of eosinophils in peripheral blood of healthy person is approximately 1%–6% of all WBCs [57]. The production of eosinophils and their accumulation in tissues and blood increases during allergies, parasitic infection, and asthma [58,59]. In allergic inflammation, eosinophils release proteins and lipid mediators contributing to signaling actions. Their other function is the host defense against helminthic

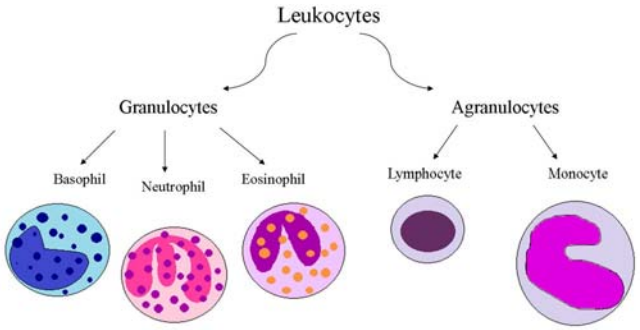


Figure 7.5
Classification of leukocytes and their characteristic morphological features.

parasites. The expression of eosinophilic receptors also affects organ transplant rejection, the development of fibrosis, wound healing, and the evolution of neoplasia [60].

In turn, neutrophils are azurophilic granulocytes and they possess multilobular nuclei [61]. Neutrophils constitute approximately 50%–60% of leukocytes [62]. The neutrophil cell contains specific granules—azurophilic granules and gelatinase granules. Their major component is an azurophilic protein called myeloperoxidase (MPO). The granules are essential for neutrophil functions in innate immunity. They are densely packed with specific proteins which contribute to microbial killing after their extracellular release. The membranes of the granules contain proteins which react with environmental signals and initiate the production of bactericidal oxygen species [63]. The major functions of neutrophils are phagocytosis, digestion, and killing bacteria and microorganisms [64]. Neutrophils themselves do not distinguish between foreign and host antigens, for this purpose other cells of the immune system must be involved [65]. They are also capable of forming reactive oxygen species and other toxic molecules, which play important roles in destroying pathogens. The depletion of the neutrophils causes mucosal homeostasis. Neutrophils directly contribute to the disease pathology in which excessive recruitment and activation leads to the release of toxic products and mass transepithelial migration, causing abscesses and extensive mucosal damage [62].

EPO and MPO heme groups are similar to hemoglobin in erythrocytes [66,67]. The prosthetic group of EPO is protoporphyrin IX with a 6-coordinated spin ferric incorporated in the heme group. In neutrophils the prosthetic group in MPO is heme *a*, which can be found in cytochrome *c* oxidase. Both EPO and MPO are proteins rich in arginine, leucine, and aspartic acid, but a number of basic amino acid residues are higher in EPO than MPO [68]. So they are the excellent object for resonance Raman studies. Electronic absorption spectrum of EPO shows the presence of a strong Soret band at 412 nm, Q band at 500 nm, and charge-transfer (CT) band at 640 nm, whereas the corresponding MPO bands appear at 428, 568, and 675 nm, respectively [64,67,69,70]. This comparison shows that each hemeprotein exhibits a unique electronic absorption profile that can contribute to the differentiation of eosinophils and neutrophils by Raman spectra as shown in several reports [69–72]. Regardless of the employed technique of RS both types of granulocytes can be recognized by the collection of single-point Raman spectra [64] as well as spectroscopic hemograms [70]. For this purpose a laser excitation wavelength above 600 nm is usually used [69–72].

Here, we present results recorded with the use of confocal Raman imaging of eosinophils [73] and neutrophils extracted from the human blood of healthy volunteers. Raman spectra were collected with the excitation line at 532 nm.

Fig. 7.6B and C show the processing of Raman spectra according to the integration of spectral regions specific for organic matter (2830–3030 cm^{-1}) and DNA (770–810 cm^{-1}),

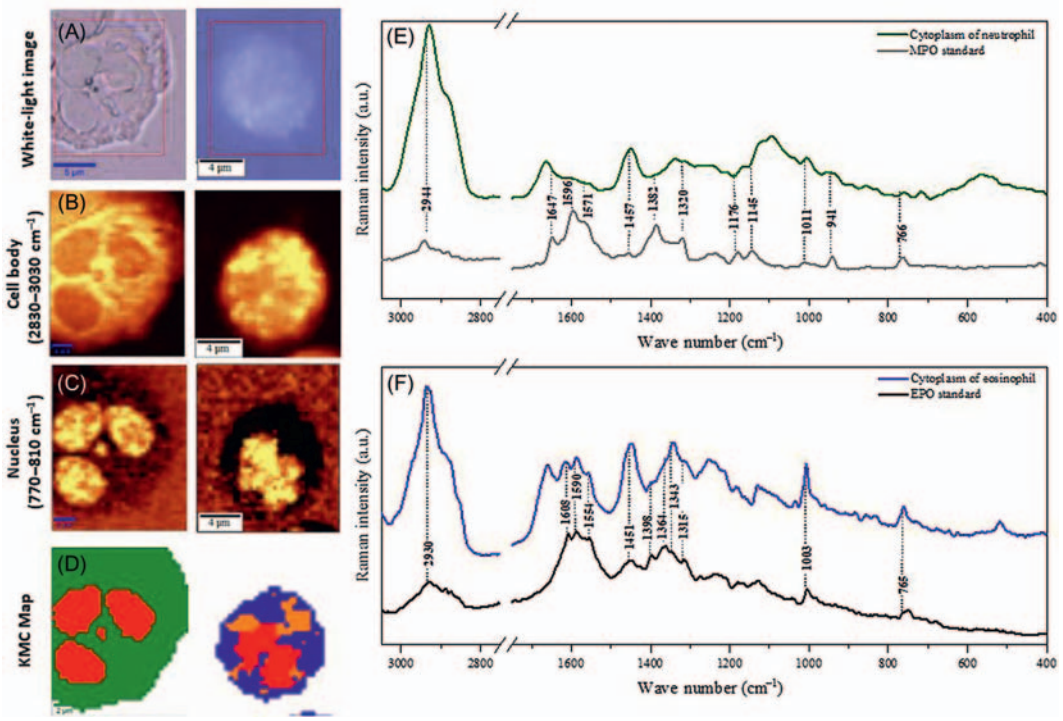


Figure 7.6

White-light microphotographs (A) and Raman distribution images of selected chemical components: cell body (B) and DNA (C) compared to false-color k-means cluster analysis (KMC) maps (D) of human neutrophil (*left*) and eosinophil (*right*). (E, F) The comparison of Raman spectra of neat myeloperoxidase (MPO) (*gray trace* in E) and eosinophil peroxidase (EPO) (*black trace* in F) with mean spectra extracted from the cytoplasmic class of the neutrophil (*green trace* in E) and eosinophil cells (*blue trace* in F). All Raman spectra were collected with a 532 nm laser excitation. Source*: Data for eosinophil were reproduced from A. Rygula et al., Raman imaging highlights biochemical heterogeneity of human eosinophils versus human eosinophilic leukaemia cell line, Br. J. Haematol. 186 (5) (2019) 685–694. Available from: https://doi.org/10.1111/bjh.15971, with permission from John Wiley and Sons.*

respectively, that represent the distribution of Raman images of cell body and nuclei. Nonhierarchical KMC segregated Raman spectra into cytoplasm and nucleus in both cells exhibiting additionally the presence of lipid bodies in the eosinophil (Fig. 7.6D). In turn, the shape of the DNA class in the KMC images clearly indicated bisegmented and multilobular nuclei in the eosinophil and neutrophil, respectively, as expected from their well-known morphology.

Mean Raman spectra of both cell types obtained from KMC analysis for the class assigned to cytoplasm showed the presence of typical Raman bands of proteins, for example, amide I and III bands at 1660 and 1252/1265 cm^{-1}, respectively, a band of the Phe residue at

1006 cm^{-1}, and the S-S bridge band at 520 cm^{-1} (Fig. 7.6D–F). Moreover, these spectra exhibited the presence of bands typical for EPO and MPO as shown in Raman spectra of neat substances (Fig. 7.6E and F). The latter were distinctly different, in particular in the regions of 1670–1530, 1420–1303, 1200–1100 cm^{-1} and differentiated additionally by the presence of Raman signals at 942 and 765 cm^{-1}. Since the 532 nm excitation wavelength was employed to record these spectra, one can expect a contribution of preresonance or resonance effect of the Q bands of the prosthetic heme group. In cytoplasm of the neutrophil (Fig. 7.6E, green trace) Raman bands at 766, 941, 1320, and 1647 cm^{-1} well corresponded to bands of MPO (Fig. 7.6E, gray trace). In the case of eosinophils, the EPO Raman signals centered at 765, 1003, 1343, and 1451 cm^{-1} were clearly observed in the spectra of the reference substance and cytoplasm (Fig. 7.6F). Some of the EPO and MPO bands overlapped with bands specific for cell macromolecules, for example, with amide I (1647 cm^{-1}) and III (1320 cm^{-1}) bands of proteins in the cytoplasm of neutrophils.

The example described above shows that the application of resonance Raman spectroscopy for imaging of neutrophils and eosinophils clearly indicated the characteristic biological features like nuclei and the presence of hemeproteins—MPO and EPO. This spectral signature can be used for differentiation of these cells in blood samples and to follow biochemical alternations due to external stimuli and diseases.

7.6 Resonance Raman spectroscopy in iron-containing proteins in tissues and cells

Iron-containing proteins are a large class of compounds present in each living organism, whose common feature is the presence of the heme—prosthetic group composed of a tetrapyrrole ring and a centrally located six-coordinated iron cation [74]. The most abundant iron-containing protein is hemoglobin (Hb), which is found in the red blood cells (RBCs) and serves as a carrier of respiratory gases (O_2 and CO_2). Other important examples are myoglobin (Mb) occurring in skeletal muscles and acting as an O_2 magazine and cytochrome *c* (C*c*) which is an electron transporter in the respiratory chain in cell mitochondria. Iron-containing proteins include also enzymatic proteins such as catalases, oxidases, and peroxidases.

Due to the presence of the porphyrin ring which absorbs light in the visible range, Raman spectra of iron-containing proteins obtained using a laser line with a wavelength adjusted to the energy of electronic transition show strong resonance enhancement [75,76]. This feature allows the carrying out of measurements of low protein concentrations with the application of low excitation laser power which will not affect the sample [7,76]. It was previously reported that a too high laser power can lead to photo/thermal dissociation of connected exogenous ligands [77]. Moreover, due to the dependence of the resonance enhancement

mechanism on the selected excitation wavelength, it is possible to acquire spectral information from a given group of vibrations. RRS allows to distinguish the oxidation and spin state of the iron cation as well as to identify connected exogenous ligands, both to the central atom and porphyrin ring [8,76,78,79].

In Fig. 7.7 the ultraviolet-visible (UV−Vis) absorption spectra and RR spectra of Hb, Mb, and Cc standard compounds recorded with the application of 488 nm (blue line), 532 nm (green line), and 633 nm (red line) excitation wavelengths are presented. Absorption spectra of iron-containing proteins are characterized mainly by the presence of strong bands in the 380−450 nm range, originating from π-π^* electron transition and named the Soret band after its discoverer. Between 480 and 650 nm, up to four much weaker Q bands are present, whereas for proteins with a high-spin ferric (Fe^{III}) cation there is an additional CT band occurring in the 620−640 nm range [76,80]. The exact number and position of Q bands indicate the molecular state of the iron cation and could be used to differentiate various proteins and their forms [76,81]. The use of a 488 nm laser line which is close enough to the energy of the Soret band allows for the greatest enhancement of completely symmetric vibrations [78]. The intensity of these modes gradually decrease with the application of longer wavelengths (532 and 633 nm) in contrast to asymmetric vibrations [76,82]. It is worth noting that it is possible to obtain resonance (or preresonance) Raman spectra of iron-containing proteins even when 1064 nm excitation is applied [75].

Application of 488 nm excitation wavelength hardly allows for the differentiation of Hb, Mb, and Cc from each other but delivers the most precise information about the spin and oxidation state of the iron cation. All these spectra are dominated by ν_4 mode originated from a half-ring stretching vibration located around 1370 cm^{-1} and often referred to as "breathing" of the porphyrin ring [76,79]. The position of this mode is regarded as the marker of the oxidation state of the iron cation and for ferrous (Fe^{II}) proteins is located at the lower wave numbers (about 1350−1360 cm^{-1}) in contrast to the ferric proteins (about 1370−1380 cm^{-1}). On the other hand, the positions of ν_2 and ν_{10} modes are sensitive to the size of the porphyrin cavity and are regarded as indicators of the spin state of the iron cation [76]. These modes originate from $\nu(C_\alpha C_m)$ vibrations and are shifted to the lower wave numbers in case of high spin (HS) compared to the low spin (LS) iron cation. In the case of Mb and Cc standard, the ν_2 and ν_{10} are located around 1584−1590 cm^{-1} and 1634−1642 cm^{-1}, respectively, indicating their LS character [76]. In contrast ν_{10} in the RR spectrum of Hb standard appears at two positions—1606 and 1638 cm^{-1}—indicating a mixture of both LS and HS spin states which is typical for methemoglobin—a form in which this standard occurs [83]. In the low wave numbers region there are two visible characteristic bands, typical for all iron-containing proteins ascribed to ν_{15} and ν_7 modes and located at around 751−753 cm^{-1} and 677−689 cm^{-1}. The first originates from pyrrole breathing, whereas the latter comes from symmetric pyrrole deformation [78].

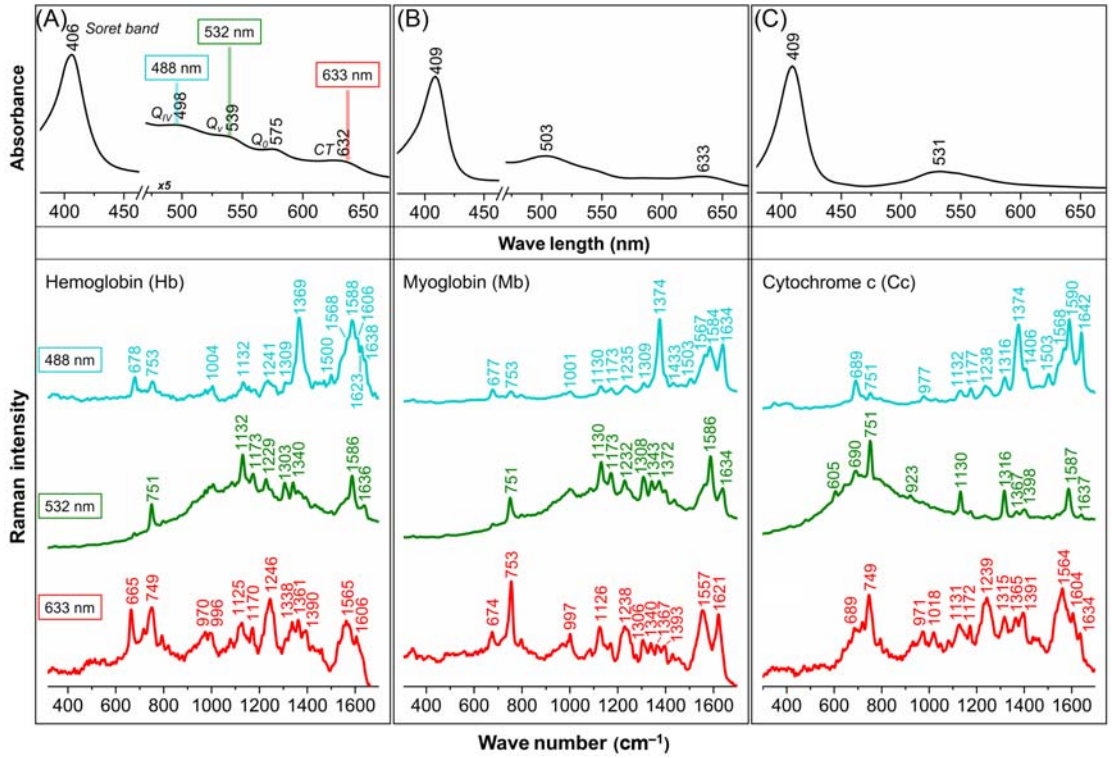


Figure 7.7

The ultraviolet-visible (UV–Vis) (*upper panel*) and resonance Raman spectroscopy (RRS) (*lower panel*) of standards compounds of: (A) hemoglobin—Hb; (B) myoglobin—Mb; and (C) cytochrome *c*—C*c*. The RRS spectra were recorded with the use of 488, 532, and 633 nm excitation wavelengths.

Asymmetric vibrations are enhanced when the 532 nm excitation is applied. The RR spectra of Hb and Mb are dominated by three intense bands ascribed to ν_{15}, ν_{22}, and ν_{37} and located at around 751, 1132, and 1586 cm^{-1}, respectively. The first two are connected with the "breathing" of the porphyrin ring, while the latter is connected to the asymmetrical $\nu(C_\alpha C_m)$ vibration. Based on these modes it is possible to recognize iron-containing proteins with heme *b* but impede their differentiation from each other—spectra of both Hb and Mb are almost identical but with variety in their intensity, especially considering the ν_{37} mode which is greatly increased in the case of Mb. On the other hand, RR spectra of iron-containing proteins with heme *c* are characterized additionally by the presence of the ν_{21} mode located at 1316 cm^{-1} and originates from the bending of methylene groups, which makes it possible to distinguish Cc from Mb or Hb. Besides the proximity of the excitation energy to the 488 nm laser line, RR spectra obtained with 532 nm are characterized by very low intensity of the ν_4 mode.

Excitation with the 633 nm laser line leads to the selective enhancement of both totally and nontotally symmetrical vibrational modes [76,82]. This phenomenon occurs due to the other resonance mechanism compared to the excitations with 488 and 532 nm lines and is connected, among others, with the presence of the CT band in the absorption spectra of iron-containing proteins. This can be clearly seen by the observation of two totally symmetrical modes: ν_7 and ν_4. The first appears as an intense band at 665 cm^{-1}, while the latter, located at 1361 cm^{-1}, is characterized by very low intensity. The ν_7 mode supplies also another interesting feature—the spectrum of Mb is almost invisible, thus allowing it to be used for Mb/Hb differentiation. Between 1238 and 1246 cm^{-1} there occurs another characteristic band for RR spectra of all iron-containing proteins obtained with this excitation originating from in-plane bending vibration of the methine hydrogen [82].

Fig. 7.8 presents the examples of Raman imaging results of tissues fixed with glutaraldehyde (lung and heart) and single functional, nonfixed cells (HAEC—human endothelial cell line) [76,84]. Raman images of the organic matter distribution (Fig. 7.8C) were obtained by the integration of the bands in the spectral region of 2800−3050 cm^{-1}, while iron-containing proteins (Fig. 7.8D) are based on the integration of the ν_{15} mode in the spectral range of 740−760 cm^{-1} [83]. Additionally, for each imaged tissue or cell area the k-means cluster analysis (KMC) was carried out and presented in the image (Fig. 7.8E). The classes obtained from KMC were color coded for a specific type of iron-containing protein occurring in each sample—Hb in lung tissue (blue), Mb in heart tissue (red), and Cc in functional HAEC cell (green). In all cases, the cytoplasm was coded in gray. The average Raman spectra for each KMC class were presented in Fig. 7.8.

The blue spectrum (Fig. 7.9I) extracted from iron-containing protein found in the lung tissue clearly corresponds to the Raman spectrum of the Hb standard, as indicated by the positions of the ν_{15}, ν_{22}, and ν_{37} modes at 756, 1131, and 1591 cm^{-1}, respectively, as well

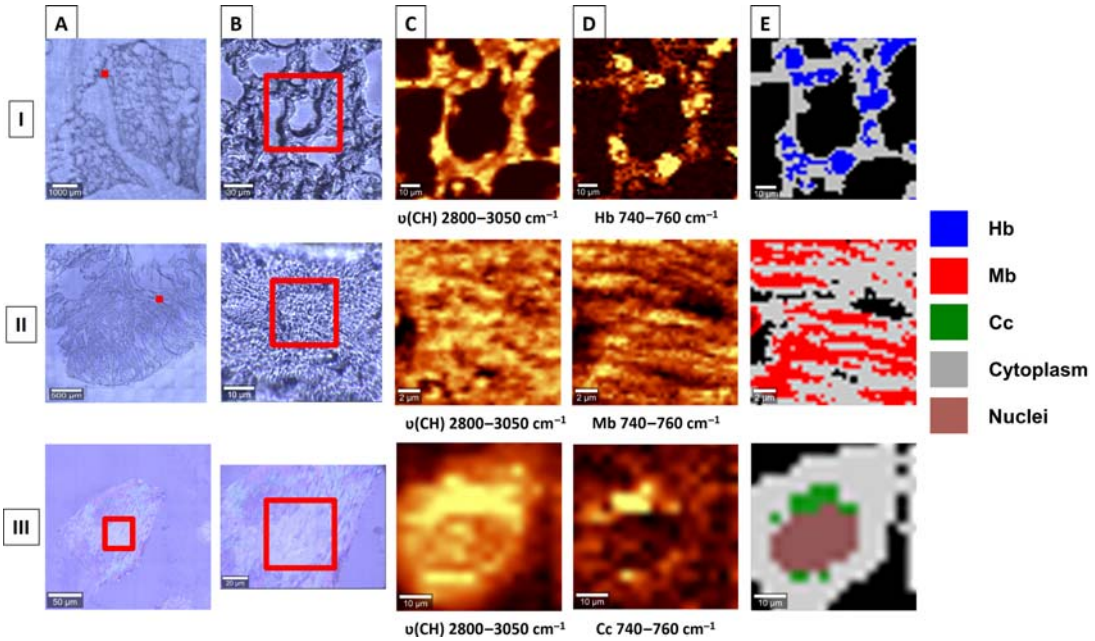


Figure 7.8

The RS imaging of (I) fixed lung tissue; (II) fixed heart tissue; (III) nonfixed HAEC cell. (A) visual image × 20; (B) visual image × 100; (C) Raman integration images (RI) representing organic matter (2800–3050 cm^{-1}); (D) RI representing iron-containing proteins (740–760 cm^{-1}); (E) k-means cluster analysis (KMC) image. The results were performed with the use of a 532 nm excitation wavelength.

as their relative intensity ratios that are approximately equal to 1:1:1. Correlating the above with the information received from the Raman images presented in Fig. 7.9I, it can be concluded that such information is delivered by the presence of RBCs [84]. Definitely the most intense band in the case of the RR spectrum marked in red (Fig. 7.9II) is the ν_{37} mode located at 1588 cm^{-1}, which allows it to be ascribed to Mb present in the heart tissue in large quantities. Mb is a main hemeprotein of longitudinal cardiomyocytes, which are well visible in Fig. 7.8IID–E. It is worth noting, that in the case of Mb, the relative intensities ratio of the of the ν_{15}, ν_{22}, and ν_{37} modes are in approximation equal to 1:1:2, which makes it a valuable observation for Mb/Hb differentiation. The RR spectrum of Cc (Fig. 7.9III) is additionally characterized by the presence of an intense ν_{21} mode located at 1315 cm^{-1} and relative intensities ratio is approximately equal to 1:1:1:1 when compared to the ν_{15}, ν_{22}, and ν_{37} modes. Moreover the RR spectrum of this iron-containing protein is characterized by the absence of any background when compared to the RR spectra of Hb and Mb obtained with 532 nm excitation [76]. As indicted by the Raman images present in Fig. 7.8III all Cc is gathered around the cell nucleus in areas that are characteristic of ER and mitochondria. Based on the previous works we

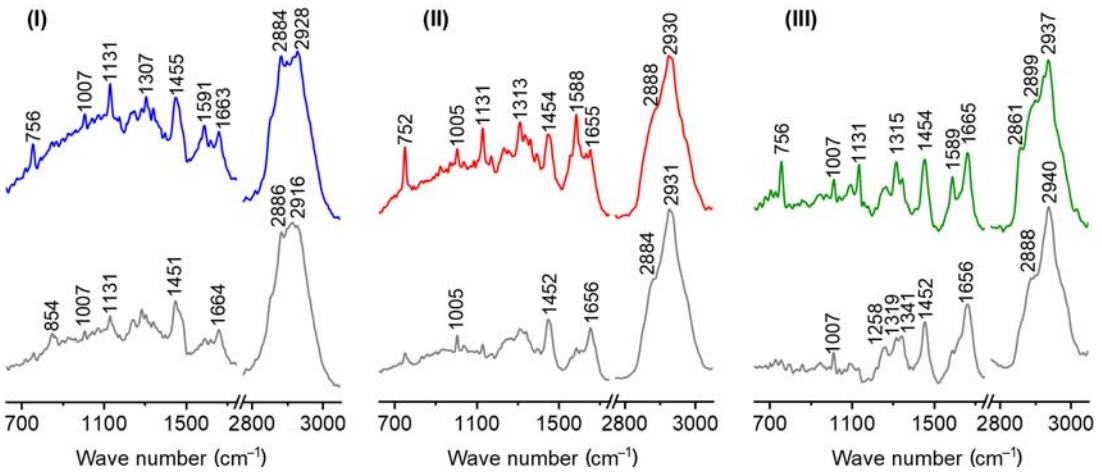


Figure 7.9

The average Raman spectra extracted from k-means cluster analysis (KMC) presented in Fig. 7.2: (I) fixed lung tissue; (II) fixed heart tissue; (III) nonfixed HAEC cell. Colors of spectra correspond to the color coding of KMC in Fig. 7.8: blue—Hb; red—Mb, green—Cc; and gray—cytoplasm. The Raman spectra were obtained with the use of a 532 nm excitation wavelength.

may conclude that as presented in Fig. 7.9III the functional cell has mitochondria filled with the reduced form of Cc as indicated by presence of high-intensity bands on the RR spectrum. Due to fixation such Cc may be converted to the oxidized form, which is characterized by a weak signal, which impedes its observation in fixed cells [42].

7.7 Characterization of lung proteins altered by cancer cell infiltration

The main advantage of infrared spectroscopy is its ability to provide data for the analysis of secondary structures of polypeptides and proteins not only in cells and body fluids but also in tissues [85]. Moreover its nondestructive and label-free character allows for further sample investigations with reference methods such as histological staining, which is "the gold standard" in the classification of tissue compartments [86].

Lung cancer progression is associated with changes appearing in the extracellular matrix (ECM), consisting of basement membranes and connective tissue [87]. They represent networks of macromolecules being the support for surrounding cells [88–90]. In addition, ECM plays an important role in several mechanisms such as homeostasis, developmental organogenesis, injury-repair responses, and regulation of diverse cell functions. It is mainly composed of proteins (collagen, elastin), proteoglycans, glycoproteins, and nonproteoglycan polysaccharides [91]. The deregulation and remodeling of matrix, implicated by cancer, introduce numerous changes in their structure and chemical composition [92].

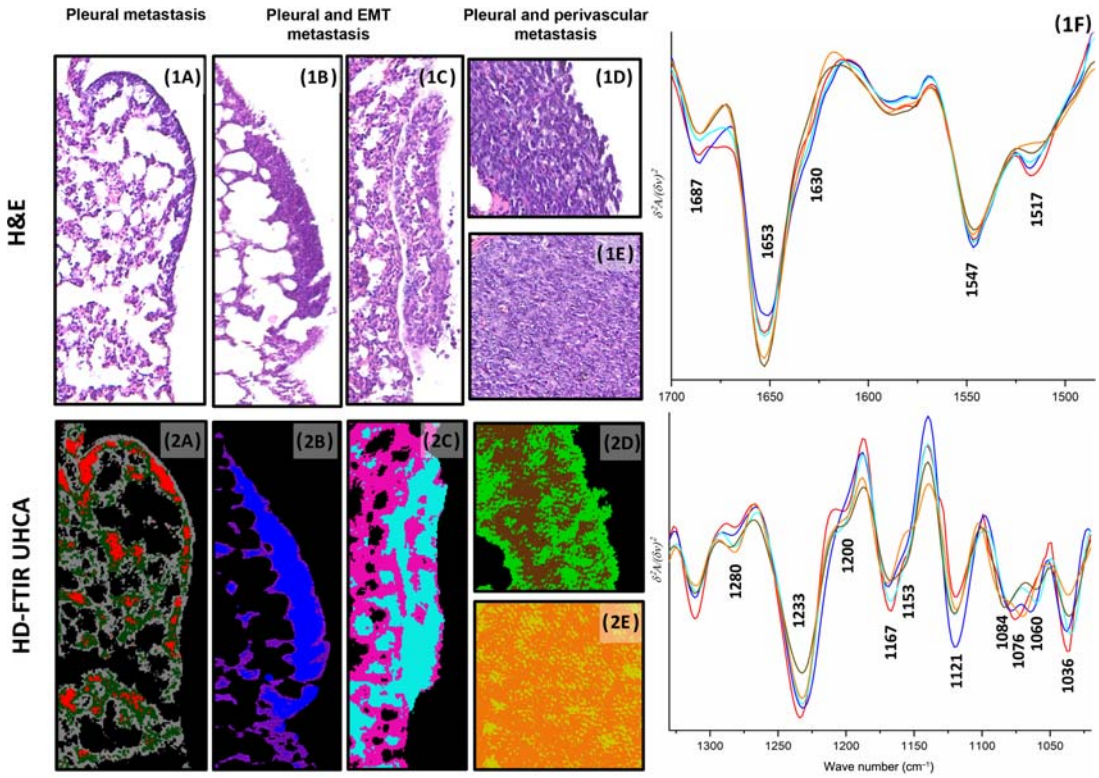


Figure 7.10

Hemotoxylin and eosin (H&E) microphotographs showing the cross section of lung tissue extracted from mice 3 (1A), 4 (1B, 1C), and 5 (1D, 1E) weeks after orthotropic inoculation of variable 4T1 tumor cells imaged using HD Fourier-transform infrared (FTIR) spectroscopic imaging; false-color maps of unsupervised hierarchical cluster analysis (UHCA) for HD FTIR images (2A–2E) and mean second derivative FTIR spectra extracted from UHCA represent metastasis class (2A—red; 2B—blue; 2C—aqua; 2D—green; and 2E—orange); the colors of the spectra correspond to the colors of the classes. Source*: Reproduced from K. Augustyniak et al., High and ultra-high definition of infrared spectral histopathology gives an insight into chemical environment of lung metastases in breast cancer, J. Biophotonics 12 (4) (2019) e201800345, with permission from John Wiley and Sons.*

Cancer cells can spread to distant organs through the bloodstream or lymphatic system, implicating various origins of further developed metastasis sites (Fig. 7.10) [14]. High definition FTIR spectroscopic imaging distinguishes the subcellular structures of cancer metastasis that appear in the lung parenchyma as well as indicates biochemical changes by specific FTIR profiles over the cancer progression. In the micrometastatic phase—3 weeks after orthotropic injection with viable 4T1 tumor cells to mice—FTIR images and their chemometric analysis highlighted the presence of a cluster of cancer cells forming a thin layer under the pleural edge (dark violet color, Fig. 7.10 1A) [14]. Unsupervised hierarchical cluster

analysis (UHCA) analysis distinguished three classes (Fig. 7.10 2A), that is, cancer cells (red class) on the edge of the pleura and inflammatory cells under the pleural cavity (gray and green classes). Within 4 weeks after cancer cells' inoculation the presence of two macrometastases in a lung cross section was observed. The metastatic site depicted in Fig. 7.10 2B illustrates macrometastasis under the pleural cavity adjacent to inflamed cells, where the cluster analysis distinguished two classes (blue and violet) attributed to the solid macrometastasis and surrounding tissue, respectively. The second metastatic site represented epithelial to mesenchymal transition, which revealed neoplastic cells infiltrating cellular matrix within macroparticles secreted by cancer cells (Fig. 7.10 1C). This transition was identified by spindle-shaped cells dissociating from the pleura and migrating to distant organs [14]. Again UHCA analysis grouped FTIR spectra into two classes—a pink class assigned to inflamed parenchyma and an aqua class attributed to the metastatic foci. Spectroscopic investigations of lungs after five weeks from inoculation with viable 4T1 adenocarcinoma breast cancer cells detected pleural and perivascular metastasis. Hemotoxylin and eosin (H&E) staining distinguished differences in morphology: pleural metastasis is a mixture of inflamed and cancer cells with hyaline (Fig. 7.10 1D), while perivascular tumor consisted of only cancer cells with clockface nuclei and high polymorphism (Fig. 7.10 1E). UHCA analysis divided HD IR images into two classes mainly due to differences in tissue density (Fig. 7.10 2D–E) [14].

Mean second derivative FTIR spectra extracted from UHCA analysis showed that the main contributors to classification of IR images are proteins, however, biochemical changes were also found in the spectral regions specific for carbohydrates, phospholipids, and DNA/RNA (Fig. 7.10 1F) [14]. For pleural and perivascular metastasis (5 weeks), we observed changes in the absorbance of the amide I band specific for the α-helical secondary structures of proteins, which increased, along with the disappearance of a signal at 1630 cm^{-1} attributed to β-sheet secondary conformations. An increase in intensity of the 1687 cm^{-1} band (antiparallel β-sheet conformations) was pronounced in both the pleural metastasis found in the third week (red trace) and fourth week (aqua trace) of breast cancer progression and this change was accompanied by a decrease of the content of α-helical secondary structures in proteins (Fig. 7.10 1F). The IR signature of the pleural metastasis (week 3) was similar to the ETM metastasis (week 4) indicating the migrating characteristics of these cells. The intensity of the tyrosine band (1517 cm^{-1}) presents the highest intensity for the third week pleural metastasis and this fact could be associated with a high activity of tyrosine kinase, activating various growth factors (e.g., PDGF, EGF) essential in the multiplication of cancer cells [93]. Spectral alternations in the positions and intensities of collagen bands at approximately 1280, 1233, and 1200 cm^{-1} indicated remodeling of basement membranes due to the attack of cancer cells (Fig. 7.10 1F). In turn, bands at 1153 and 1036 cm^{-1}, assigned to glycosylated proteins and proteoglycans, respectively, were specific for the cancer cells' fraction from the pleural (week 3) and perivascular (week 4) (Fig. 7.10 1F).

Since the detection and localization of the advanced metastasis (AM) formed at the latest metastatic phases can be performed without any difficulty by H&E staining, the early micrometastasis (EM) is often impossible to detect due to their small sizes (up to 2 mm) [13]. Fig. 7.11 illustrates for the first time FTIR imaging of the early metastasis in different environments, that is, three single neoplastic cancer cells embedded in fibrin of blood attached to a vessel wall (Fig. 7.11 1A), a mixture of cancer and inflamed cells in lung

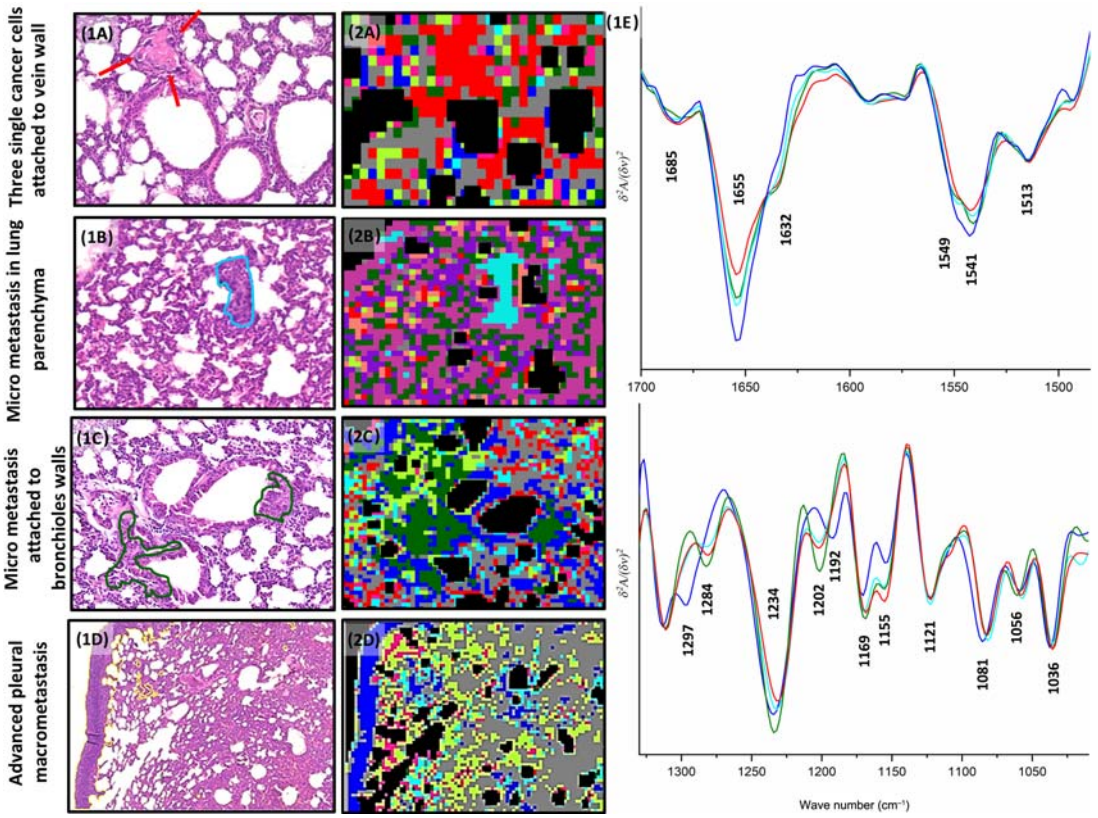


Figure 7.11

(1A–C) Hemotoxylin and eosin (H&E) stains showing cancer cells: c. 123 μm^2 in blood clot (1A); micrometastatic sites surrounded by inflamed cells: c. 4717 μm^2 (1B); and on bronchioles edges left: c. 3047 μm^2 and right: c. 6516 μm^2 (1C). H&E stains of lung cross-sections for advanced metastasis (AM) phase (1D) with marked tumor borders (yellow). (2A–D) Unsupervised hierarchical cluster analysis (UHCA) false-color maps and (1E) their mean second derivative FTIR spectra extracted from classes with micrometastases (*red*, *aqua*, and *green* from 2A, B, and C, respectively) and compared with the spectrum of tumor from the AM phase (blue trace, 2B). Source*: Reproduced from K. Chrabaszcz et al., Label-free FTIR spectroscopy detects and visualizes the early stage of pulmonary micrometastasis seeded from breast carcinoma, Biochim. Biophys. Acta Mol. Basis Dis. 1864 (11) (2018) 3574–3584, with permission from Elsevier.*

parenchyma (Fig. 7.11 1B), and in collagen matrix attached to bronchiole walls (Fig. 7.11 1C). UHCA cluster analysis of IR images displayed in Fig. 7.11 2A–D clearly identified these early and advanced metastatic sites [13]. Spectral features of EM and AM were presented by mean second derivative FTIR spectra (Fig. 7.11 1E). Within groups of neoplastic colonies (red, aqua, and green traces) spectral differences were shown regarding intensities of α-helical conformations of proteins (1655 cm^{-1}) and collagen (1284, 1234, and 1202 cm^{-1}). Cancer cells in the blood clot (red) were characterized by the lowest content of these proteins, whereas a high content of collagen was found in cancer and inflamed cells settled in bronchiole walls (green). The comparison between advanced pleural macrometastasis and the three EM sites revealed the presence of a band at 1632 cm^{-1} related to β-sheet secondary structures in EM. On the contrary, AM expressed high content of α-helical secondary structures and a significant shift of collagen bands [13].

FTIR spectroscopic imaging was also successfully employed in the detection of biochemical changes from 2 weeks after cancer cell infiltration in lung tissue [94]. Even though spectral changes were very subtle it was possible to identify molecular changes for nucleic acids, carbohydrates, and lipids. The amide I band indicated the dominance of α-helical proteins in both groups, but only AM indicated the formation of the unfolded β-sheet (1681 cm^{-1}) and tyrosine-rich proteins (1517 cm^{-1}).

7.8 Proteins of endothelium studied ex vivo

The vascular system is the network of blood vessels and it includes capillaries, veins, and arteries in which blood flows to and from the heart. The structure of a vascular wall, such as the aorta, can be divided into three layers: the tunica adventitia, media, and intima [95]. The fundamental components of tunica adventitia are elastin and collagen fibers as well as connective tissue and nerve structures. The primary objective of the tunica adventitia, as a strong outer covering of arteries and veins, is anchoring the blood vessels in the right place [96]. The tunica media is a three-dimensional network built of smooth muscle cells interspersed with collagen and elastin fibers, while the intima is constructed of a single layer of endothelial cells. The endothelial layer regulates vascular homeostasis through the secretion of a plethora of compounds with prothrombotic (e.g., von Willebrand factor), antithrombotic (e.g., trombomodulin and nitric oxide), and proinflammatory (e.g., cytokines and chemokines) effects [97].

The composition of the layers of the blood vessels is reflected in the profiles of Raman spectra recorded for the vascular wall. The average Raman spectrum of aorta, when irradiated from the intimal side, is dominated by the signal coming from proteins [96,97]. The major bands at 1664 and 1454 cm^{-1} are assigned to the amide I and CH_2 wagging modes, respectively, while the broad and featureless band at 1200–1300 cm^{-1} originates from the amide III [5]. The average Raman spectrum of the aorta exhibits also signals originating from specific amino acids (858 and

1005/1033 cm^{-1} for tyrosine and phenylalanine, respectively), hemeproteins (751, 1131, and 1585 cm^{-1}) and DNA (790 and 1100 cm^{-1}) [79]. The marker bands of elastin can be recognized as the features observed at 1108 and 531 cm^{-1} and attributed to desmosine/isodesmosine, the specific amino acids building the elastin and S−S stretching vibrations, respectively [98,99]. In contrast with the spectrum collected from the intimal side, the Raman spectrum observed for the adventitial side of a normal aorta with several millimeters of adipose tissue, appears to be mainly lipid, and in particular to triglycerides, with almost no contribution from protein [100].

Proper function of endothelium guards the health of the cardiovascular system, while endothelial dysfunction leads to cardiovascular diseases, including diabetes mellitus, hypertension, and atherosclerosis [101]. Previously, RS was successfully used for general characterization of normal functional human artery [96,100,102,103], porcine [104], rabbit [105,106], or murine aorta [107−109]. Furthermore with the use of Raman imaging, the chemical content of endothelium within the blood vessel was measured in murine models of diseases associated with endothelial dysfunction: hypertension [110], metastatic breast cancer [111], T2DM [112], and atherosclerosis [106,113−115]. Spectral differences, mainly connected with alterations in lipid to protein ratio or changes in the secondary structure of proteins enabled successful differentiation of the spectra from healthy mice as compared with mice with endothelial dysfunction with high sensitivity and specificity [116].

Hypertension is a long-term medical condition in which the blood is elevated, which can lead to coronary artery disease, stroke, heart failure, and others. The common feature of hypertension is also endothelial dysfunction, and it is associated with the loss of the balance between the production of endothelium-derived contracting factors (e.g., angiotensin II and endothelin-1) and endothelium-derived relaxing factors (e.g., nitric oxide and NO) [117]. A reduced amount of NO leads to increased production of the reactive oxygen species in the vascular wall, increased vascular stiffness, and alterations in vascular structure [118]. As was described previously in the literature, the major changes in the Raman spectra of the endothelium in hypertension were observed in the 1400−1200 cm^{-1} region and were attributed to alterations in the secondary structure of proteins (Fig. 7.12) [110]. Depicted changes were assigned to the decrease of the relative amount of the α-helix structure (decrease of the intensity of the band at 1281 cm^{-1}) relative to the β-sheet (increase of the intensity of the counterpart at 1254 cm^{-1}) in NO-deficient hypertensive mice compared to the control. Moreover, a shift of the amide I band position (1658 and 1664 cm^{-1} for control and for the endothelium of the hypertensive mice, respectively) was described as a consequence of the changes in the secondary structure of proteins (Fig. 7.12).

Deconvolution of the amide I band allowed for the determination of six counterparts observed at 1620 (aromatic amino acids), 1642 (unordered structure), 1656 (α-helix and lipids), 1668 (β-sheet), 1681 (β-turns), and 1693 cm^{-1} (amino acids: glutamine and arginine). The positions of the counterparts for control and NO-hypertensive endothelium

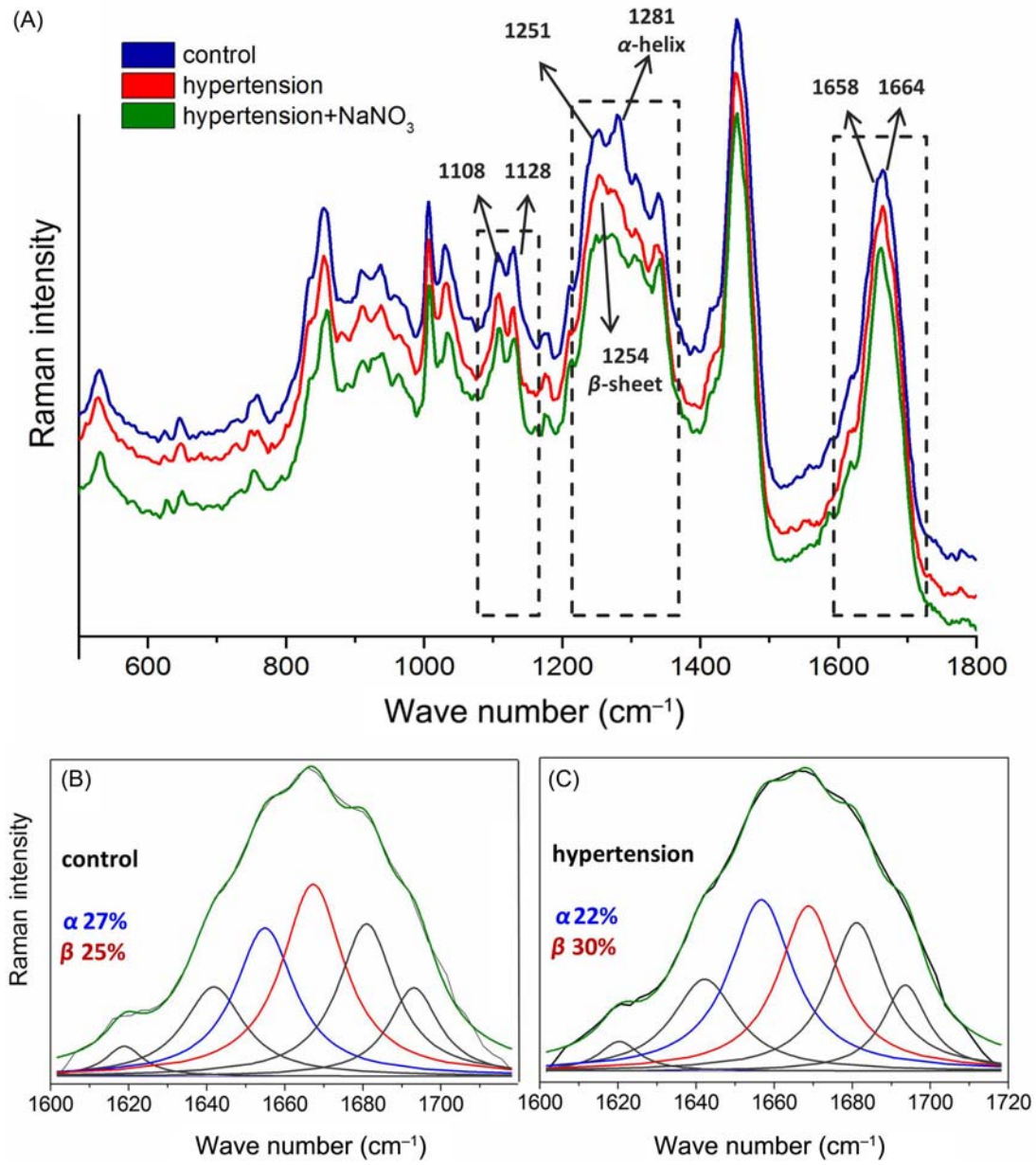


Figure 7.12

The average spectra of the control [*blue*; wild-type (WT) mice], NO-deficient hypertensive [*red*, WT mice supplemented with *N*-nitro-L-arginine methyl ester (L-NAME)] and nitrate-treated (*green*; NO-deficient mice supplemented with nitrate) mice (A) taken from the endothelium region. Deconvolution of the amide I band for control (B) and NO-deficient hypertensive mice (C) showed changes in the estimated content of the α and β secondary structures. The blue line corresponds to the α-helix (1656 cm^{-1}) and the red line to the β-sheet (1668 cm^{-1}), and the green line is the sum of all the bands. Source*: Adapted with permission from M.Z. Pacia, L. Mateuszuk, S. Chlopicki, M. Baranska, A. Kaczor, Biochemical changes of the endothelium in the murine model of NO-deficient hypertension, Analyst 140 (7) (2015) 2178–2184. Copyright (2014) American Chemical Society.*

were the same, while the contributions of two crucial bands located at 1656 (α-helix) and 1668 cm^{-1} (β-sheet) were changed. Although the deconvolution analysis cannot be regarded as a fully quantitative method, it allowed for the determination of the directions of the observed alterations. The estimated content of the α secondary structure decreased from 27% to 22% during the development of hypertension, while the amount of β secondary structure increased from 25% to 30%. It is worth adding that the estimated contents of the α and β secondary structures were determined as the ratio of the integral intensity of the counterpart due to the α or β structure to the integral intensity of the amide I band [110].

Upon development of hypertension, structural abnormalities in the tunica media were also observed in the Raman images as a thickening of elastin fibers in the vascular wall of hypertensive NO-deficient mice in comparison to the control [110]. Remodeling of the vascular wall of hypertensive mice [119,120] or rats [121] is the result of imbalance in the synthesis and degradation of elastin and collagenous components. Under the pathological conditions, the synthesis of elastin is elevated, leading to an increase in the amount of elastin fibers in the tissue [119].

Characterization of chemical changes in the endothelial layer was described in three murine models of disease associated with endothelial dysfunction: T2DM (db/db mice) [112], hypertension (induced by L-NAME) [110], and metastatic breast cancer [111]. Chemical imbalance in the protein and lipid level in the endothelium was a common feature of all three pathologies (Fig. 7.13) [116]. The overall lipid to protein content in endothelium was calculated as the ratio of the integral intensity of the band for lipids and proteins (centered at 2940 cm^{-1} and assigned to the C−H stretching vibrations) to the protein marker band (at 1007 cm^{-1}, assigned to the ring breathing mode of phenylalanine).

Alterations in the endothelial layer in the murine model of diabetes was observed as a statistically significant increase of lipid to protein content of up to 24% relative to control animals ($P = .0009$; Fig. 7.13). On the contrary, a decrease in lipid to protein level was observed in the endothelium under the development of hypertension and cancer metastasis. The ratio of lipid to protein content decreased significantly (6.8%, $P = .041$) for the hypertensive mice but this effect was modest relative to the control. In contrast, the fall in lipid to protein ratio was pronounced in the endothelium in cancer metastasis (19%, $P = .0003$) [116].

All the above-mentioned markers of endothelial dysfunction, like changes in lipid to protein ratio, alterations in the secondary structure of proteins, and reorganization of elastin fibers can be used for assessment of the state of the vascular endothelium.

7.9 Fourier-transform infrared microscopy of proteins

Vibrational spectroscopy based on infrared absorption is a well established method proposed for diagnostic application. FTIR spectroscopy in combination with chemometrics

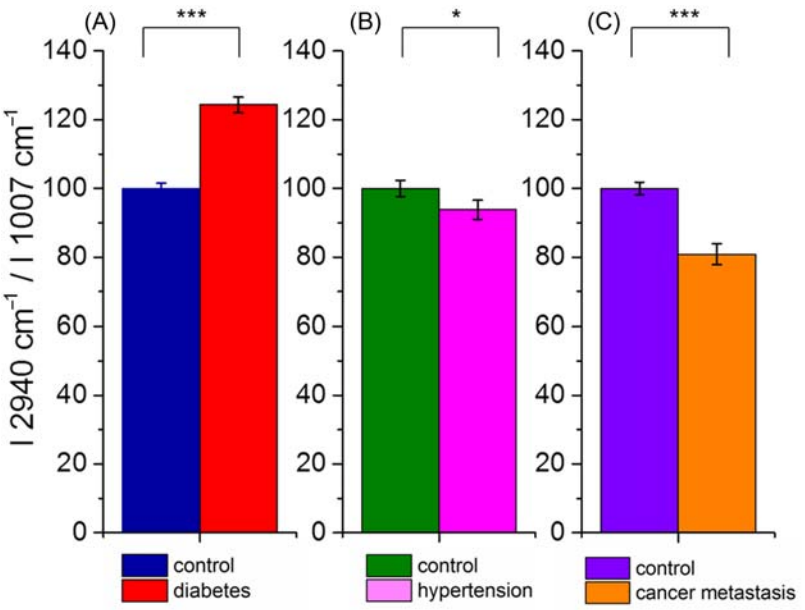


Figure 7.13

Comparison of lipid to protein ratio changes upon development of diseases: diabetes (A), hypertension (B), and cancer metastasis (C). Results are presented as means ± standard error of the mean. A difference between means was considered significant if $P < .05$. For easier and more intuitive comparison, the lipid to protein ratio for each control was set at 100. Source: *Adapted with permission from M.Z. Pacia et al., Rapid biochemical profiling of endothelial dysfunction in diabetes, hypertension and cancer metastasis by hierarchical cluster analysis of Raman spectra, J. Raman Spectrosc. 47 (11) (2016) 1310–1317. Copyright (2014) American Chemical Society.*

is one of the most important contemporary analytical techniques used in the study of biological materials such as tissues or cells, since it allows for identification and analysis of the spatial distribution of the main biochemical components in the bio sample, while ensuring a high level of selectivity and resolution. Chemometrics, which employee statistical and mathematical methods in analyzing chemical data, is a powerful tool and alternative for histopathology. Diagnostic and prognostic tools based on this method have the potential to be a standard in clinical systems leading to improved patient outcome.

The use of FTIR microscopy for the determination of biochemical changes, including secondary structure alterations of proteins, with a view toward diagnostic applications, was proposed for many diseases such as diabetes [12,122–127], cancer [128–136], and hypertension [52]. Goormaghtigh et al. [127] and Majzner et al. [12] have presented an application of a multilinear regression model (MLR) based on mid-infrared absorption spectra (FTIR) to analyze subtle changes in the secondary structure of protein in tissues caused by diabetic disease. Majzner et al. [12] described the application of a protein

secondary structure determination method developed by Goormaghtigh et al. [127] for the spatial analysis of tissues. The study establishes a spectral database of proteins of known structures which covers a wide range of helix and sheet fractional content values and different protein domain folds as described by CATH classification [137,138]. The Goormaghtigh group reported on the importance of a complete reference database [139]. However, it should be mentioned that it is difficult to create a database which includes a wide coverage of extreme values, especially for T and G structures where the asymmetric distribution of values is significant. From the point of view of the obtained tissue results, the most important are α-helix and β-sheet secondary structures, and for them the created database [12,127] covers a very wide range of content, including extreme values (i.e., α-helix: myoglobin—77%, ferritin—71%, superoxide dismutase (Cu and Zn)—2.6%, avidin—0%; β-sheet: xylanase—61,4%, avidin—48%, albumin human—0%, myoglobin—2%).

Each protein contributes to the model unique information about the structure and exact values share of secondary structures for each of them obtained using the dual signal subspace projection (DSSP) algorithm implemented in the Protein Data Bank (PDB) database [137,140]. Protein secondary structure fractional compositions were assigned according to DSSP where H means α-helix structure, B is a residue in an isolated β-bridge, E is an extended strand participating in the β-ladder, G means 3-helix (3/10 helix), I is 5 helix (p-helix), T is a hydrogen-bonded turn, S is a bend, and NA (unassigned) means a blank in the DSSP secondary structure determination that stands for a loop or irregularity. Examples of secondary structure assignment obtained based on the DSSP database for four selected proteins are presented in Table 7.1. FTIR spectra and second derivative spectra are presented in Fig. 7.14.

Table 7.1: Examples of proteins with different secondary structure content: their source, Protein Data Bank (PDB) code, and % of secondary structure.

Protein name	Source	Acronym	PDP ID	Secondary structure (%)[a]							
				H	E	T	G	S	I	B	NA
Avidin	Egg white	AVI	1avd	0	48	4.4	3.6	15.7	0	0	28.3
Ferritin	Horse spleen	FTN	1hrs	71.3	0	8.6	1.7	4.0	0	0	14.4
Myoglobin	Horse skeleton muscle	MBN	1mbn	77.1	2	0	0	9.8	0	0	11.1
Superoxide dismutase (Cu and Zn)	Bovine erythrocyte	SOD	1sxc	2.6	38.9	14.5	3.6	14.2	0	2.6	23.6

[a]Protein secondary structure fractional compositions as assigned by DSSP: H = alpha helix, B = residue in isolated beta-bridge, E = extended strand, participates in beta ladder, G = 3-helix (3/10 helix), I = 5 helix (pi helix), T = hydrogen-bonded turn, S = bend., NA (unassigned) = a blank in the DSSP secondary structure determination stands for loop or irregular [12].

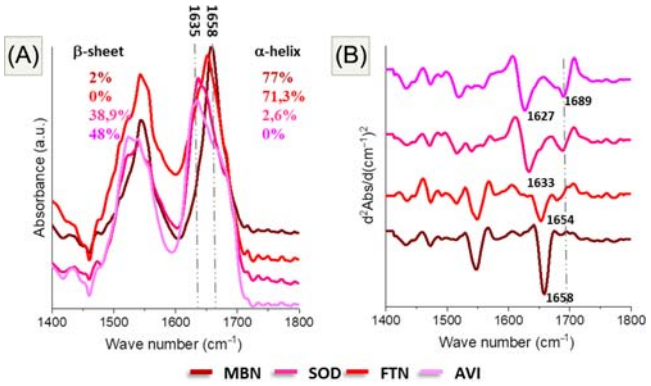


Figure 7.14

Fourier-transform infrared (FTIR) spectra (A) and second derivative FTIR spectra (B) of a series of four selected proteins sorted in order of increasing β-sheet content. The spectra have been vector normalized in the amide I–II region. The lines in panel (A) indicate the wave number of the α-helix (1658 cm^{-1}) and β-sheet contribution (1635 cm^{-1}). The line in panel B (1689 cm^{-1}) indicates the wave number of the antiparallel β-sheet contribution. *AVI*, Avidin; *FTN*, ferritin; *MBN*, myoglobin; *SOD*, superoxide dismutase.

The presented spectra and their second derivatives clearly show the impact of changed secondary structure on the amide I band position since the band shifts from 1658 (α-helix marker band) to 1635 cm^{-1} (β-sheet marker band) when the contribution of the α-helix structure is decreasing (Table 7.1). Moreover, the second derivative of the FTIR spectra presents the marker band of the antiparallel β-sheet (1689 cm^{-1}) [141] when the contribution of the β-sheet is increasing (Fig. 7.14B).

A mathematical model based on FTIR measurements for the prediction of alterations in the content of secondary structures of proteins analyzed by FTIR was prepared and measured by means of two FTIR techniques: transflection and ATR. The ATR spectra, formed from the ATR database, were used without any additional treatment apart from the intensity correction. The main disadvantage of ATR spectra database was the use of proteins in solid state—very often in lyophilized form which is changed from the native structure after the rapid removal of water molecules during the lyophilization process. Changes in secondary structure caused by lyophilization of their aqueous solutions have been investigated by means of Fourier-transform infrared spectroscopy [142,143].

The main assumption of the study was that each wave number can be crucial for the considered structure and have an important role in predicting the secondary structure content. According to this the mathematical approach was chosen as the conceptual study.

The idea of creating a model for predicting the content of secondary protein structures can be presented in the form of a matrix (7.1) [12,127]:

$$\begin{bmatrix} 1 & A_{1,w_i} & \cdots & A_{1,w_k} & A_{1,w_n} \\ 1 & A_{2,w_i} & \cdots & A_{2,w_k} & A_{2,w_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & A_{j-1,w_i} & \cdots & A_{j-1,w_k} & A_{j-1,w_n} \\ 1 & A_{j,w_i} & \cdots & A_{j,w_k} & A_{j,w_n} \end{bmatrix} x \begin{pmatrix} a_1 \\ a_2 \\ \cdots \\ a_n \end{pmatrix} = \begin{pmatrix} c_1 \\ c_c \\ \cdots \\ c_{j-1} \\ c_j \end{pmatrix} \tag{7.1}$$

An absorbance value (A) was used to create a model describing the average content of secondary protein structures at selected values of wave numbers (w). The proposed model is a linear (multilinear) for the absorbance range $A_{j,wn}$, where A is the absorbance of the rescaled spectra and n is the spectral number. The model includes the constant a_1 and the proportionality factor a_2. For each wave number w_n, the content of the secondary structure of proteins c_j was related to the absorbance $A_{j,wi}$ through relationship (7.2):

$$c_j = a_1 + a_2 \times A_{j,wi} + \cdots + a_n \times A_{j,wn}. \tag{7.2}$$

Since the values appearing in the above matrix are unknowns, which in the model building process were determined, it was necessary to know the content of secondary structures for each protein, that is, c_j (examples in Table 7.1). Next, in constructing equations of the model the IR bands' wave numbers (variables) were chosen according to mathematical calculations. After finding the first best variable (wave number), in a next step, the prediction error calculation was performed again using the first wave number and testing the remaining wave numbers (Fig. 7.15). Then two best wave numbers were used and another one was tested. This procedure was repeated until a satisfactory drop of the level of prediction error [the root mean squared deviation (RMSD) value] was achieved. In order to avoid overfitting, the number of the wave numbers was kept as low as possible. As the prediction error the RMSD was used. RMSD is a measure of compliance and is the most commonly used measure of differences between the values predicted by the model (estimator) and real values—in other words RMSD is a measure of the self-prediction error of the model. An example of dependence of the prediction error (RMSD) from the wave number for the beta-turn structure is presented in Fig. 7.15.

The choice of wave numbers was repeated in the amide I and II regions until a relatively low value of RMSD was obtained. For each secondary structure the calculated RMSD values to sequentially added variables were prepared. The evolution of RMSD indicates the optimal number of variables necessary to build a model that has the best correlation with the true secondary structure of proteins with simultaneously limited possible overfitting. On this basis the final MLR has been obtained.

The possible pitfalls of the standard validation approach should be mentioned along with the fact that the validation and wave number selection procedure is not free of such

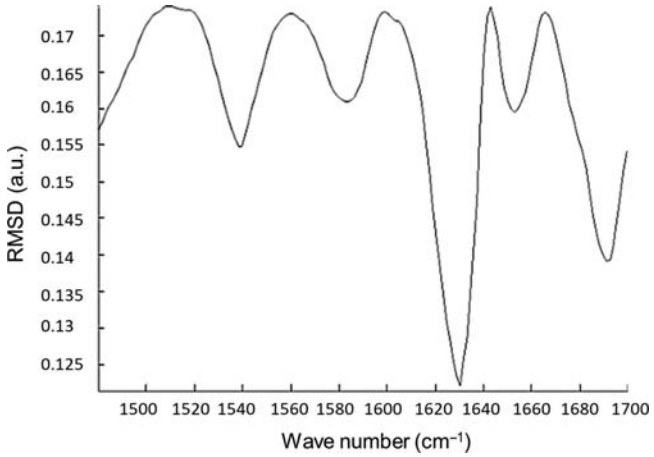


Figure 7.15
The dependence of the root mean squared deviation (RMSD) value on the wave number in predicting the contribution of the β-sheet structure. Based on analysis of the relationship the next variable in the model predicting the participation of the secondary structure was selected. The variable with the lowest prediction error is wave number at 1630 cm^{-1}. The RMSD value for this variable is about 12%.

dangers. It is quite subjective to choose the number of wave numbers based on the RMSD (of self-prediction) dependence. As described in the paper [12], very often a clear minimum is not present and the choice is based on the appearance of a plateau. To check the risk of overfitting the calibration to the training samples is to predict the percentage of the secondary structure for seven samples not used to optimize the calibration. The coefficients of correlation for linear regressions of the predicted and real values of the content of H and E secondary structures of proteins where $R^2 = 0.831$ and 0.674 for H and E, respectively, confirm that the obtained models gave satisfactory results [12]. The obtained MLR models were then used for a total secondary structure contents imaging of tissue sections with the aim of characterizing a difference between aortae of rats and mice challenged and unchallenged with induced diabetes.

The model was tested on a relatively well-defined protein structure tissue (rat tail), and finally was used to analyze in detail a secondary structure of proteins in the aortic walls of diabetic mice (db/db) and rats (streptozotocin-induced diabetic model). Rat tail was used for the test as a less complex tissue built of a more predictable and known composition of proteins. Such an approach allows for verification of the model for a biological system and to be sure that the obtained results are quite consistent with the actual secondary structure of the major protein. However, it is clear that each extrapolation from a calibration model yields a risk of inaccuracy.

The most important factor for this study was a tendency for observed changes between control and diabetic tissue. For each examined tissue we observed the same pattern of changes in the amide I/II area, so the main explanation of this is the change in the secondary structure. However, this change does not have to be caused by changing the secondary structure of the protein, but by changing the biochemical composition, for example, by the appearance of new proteins and biochemical structures. The changes in the chemical composition of the tissue in disease state are highly probable. Differences in an undefined, uncharacterized total protein content and biochemical composition may be manifested by differences in the secondary structure.

The application of the described method might be very useful for the study of different diseases at the tissue level, as it exemplifies the conclusions obtained in the in-depth study of vascular wall that the authors have performed. The obtained model was also used to characterize the protein profile in healthy and atherosclerotic tissue [144]. The calculated secondary structure contents of proteins in atherosclerotic tissue showed an increase in structures related to β-sheet (E and T) and a decrease of helical (H) and unassigned structures.

7.10 Conclusions and perspectives

We have shown the potential of vibrational spectroscopy in the analysis of proteins generally, and in some cases more specifically, in studies on lifestyle diseases. The interpretation of spectra is based mainly on the literature and model studies using reference methods and samples. We are able to easily identify proteins as a group and follow changes in their content, usually in relation to other compounds, for example, lipids. By applying two spectroscopic methods, Raman and FTIR, in a form of advanced modern techniques, we are able to make the study more complementary, comprehensive, and reliable. It is possible to understand the molecular structure of proteins and their changes, as well as protein interactions, and correlate these with specific pathology development. Proteins seem to be crucial for better understanding the processes going on at the cellular level-assisted development of diseases, and vibrational spectroscopy has potential for their in vitro and ex vivo analysis, so the prospects for the clinical applications of these methods are very promising.

Acknowledgments

M. Baranska acknowledges financial support by National Science Center, Poland (DEC-2018/29/B/ST4/00335). K. Majzner acknowledges financial support by National Science Center, Poland (DEC-2016/21/D/ST4/00870). M. Z. Pacia acknowledges financial support by National Science Center, Poland (DEC- 2017/24/C/ST4/00075). J. Dybas acknowledges financial support by National Science Center, Poland (DEC-016/23/N/ST4/00929). K. Chrabaszcz and K. M. Marzec acknowledge financial support by the Polish Ministry of Science and Higher Education (Grant No. IP2015 048474). K. Malek acknowledges financial support by National Science Center, Poland (DEC-2016/23/B/NZ4/01379).

References

[1] G.P. Moss, P.A.S. Smith, D. Tavernier, Glossary of class names of organic compounds and reactivity intermediates based on structure (IUPAC Recommendations 1995), Pure Appl. Chem. 67 (8–9) (1995) 1307–1375.

[2] R. Milo, What is the total number of protein molecules per cell volume? A call to rethink some published values, Bioessays 35 (12) (2013) 1050–1055.

[3] F.U. Hartl, Protein misfolding diseases, Annu. Rev. Biochem. 86 (1) (2017) 21–26.

[4] F. Chiti, C.M. Dobson, Protein Misfolding, functional amyloid, and human disease, Annu. Rev. Biochem. 75 (1) (2006) 333–366.

[5] A. Rygula, K. Majzner, K.M. Marzec, A. Kaczor, M. Pilarczyk, M. Baranska, Raman spectroscopy of proteins: a review, J. Raman Spectrosc. 44 (8) (2013) 1061–1076.

[6] T.G. Spiro, T.C. Strekas, Resonance Raman spectra of heme proteins. Effects of oxidation and spin state, J. Am. Chem. Soc. 96 (2) (1974) 338–345.

[7] J. Dybas, et al., Spectroscopy-based characterization of Hb–NO adducts in human red blood cells exposed to NO-donor and endothelium-derived NO, Analyst 143 (18) (2018) 4335–4346.

[8] K.M. Marzec, J. Dybas, S. Chlopicki, M. Baranska, Resonance Raman in vitro detection and differentiation of the nitrite-induced hemoglobin adducts in functional human red blood cells, J. Phys. Chem. B 120 (48) (2016) 12249–12260.

[9] D. Kurouski, T. Postiglione, T. Deckert-Gaudig, V. Deckert, I.K. Lednev, Amide I vibrational mode suppression in surface (SERS) and tip (TERS) enhanced Raman spectra of protein specimens, Analyst 138 (6) (2013) 1665–1673.

[10] T. Deckert-Gaudig, V. Deckert, Tip-enhanced Raman scattering (TERS) and high-resolution bio nano-analysis—a comparison, Phys. Chem. Chem. Phys. 12 (38) (2010) 12040–12049.

[11] E. Lipiec, D. Perez-Guaita, J. Kaderli, B.R. Wood, R. Zenobi, Direct nanospectroscopic verification of the amyloid aggregation pathway, Angew. Chem. Int. Ed. 57 (28) (2018) 8519–8524.

[12] K. Majzner, T.P. Wrobel, A. Fedorowicz, S. Chlopicki, M. Baranska, Secondary structure of proteins analyzed ex vivo in vascular wall in diabetic animals using FT-IR spectroscopy, Analyst 138 (24) (2013) 7400.

[13] K. Chrabaszcz, et al., Label-free FTIR spectroscopy detects and visualizes the early stage of pulmonary micrometastasis seeded from breast carcinoma, Biochim. Biophys. Acta Mol. Basis Dis. 1864 (11) (2018) 3574–3584.

[14] K. Augustyniak, et al., High and ultra-high definition of infrared spectral histopathology gives an insight into chemical environment of lung metastases in breast cancer, J. Biophotonics 12 (4) (2019) e201800345.

[15] M.E. Peter, P. Legembre, B.C. Barnhart, Does CD95 have tumor promoting activities? Biochim. Biophys. Acta Rev. Cancer 1755 (1) (2005) 25–36.

[16] R.C. Taylor, S.P. Cullen, S.J. Martin, Apoptosis: controlled demolition at the cellular level, Nat. Rev. Mol. Cell Biol. 9 (3) (2008) 231–241.

[17] S. Elmore, Apoptosis: a review of programmed cell death, Toxicol. Pathol. 35 (4) (2007) 495–516.

[18] G. Dewson, R.M. Kluck, Mechanisms by which Bak and Bax permeabilise mitochondria during apoptosis, J. Cell Sci. 122 (16) (2009) 2801–2808.

[19] M.O. Hengartner, The biochemistry of apoptosis, Nature 407 (6805) (2000) 770–776.

[20] R.V. Rao, H.M. Ellerby, D.E. Bredesen, Coupling endoplasmic reticulum stress to the cell death program, Cell Death Differ. 11 (4) (2004) 372–380.

[21] C. Lemaire, K. Andréau, V. Souvannavong, A. Adam, Specific dual effect of cycloheximide on B lymphocyte apoptosis: involvement of CPP32/caspase-3, Biochem. Pharmacol. 58 (1) (1999) 85–93.

[22] G. Häcker, The morphology of apoptosis, Cell Tissue Res. 301 (1) (2000) 5–17.

[23] M. Bhatia, Apoptosis versus necrosis in acute pancreatitis, Am. J. Physiol. Liver Physiol 286 (2) (2004) G189–G196.

[24] D. Zhang, et al., Raman spectrum reveals the cell cycle arrest of triptolide-induced leukemic T-lymphocytes apoptosis, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 141 (2015) 216–222.
[25] S. Hu, et al., Raman spectral changes of Artemisinin-induced Raji cells apoptosis, Vib. Spectrosc. 81 (2015) 83–89.
[26] R. Yamaguchi, et al., In situ real-time monitoring of apoptosis on leukemia cells by surface infrared spectroscopy, J. Appl. Phys. 105 (2) (2009) 24701.
[27] H. Huang, et al., Confocal Raman spectroscopic analysis of the cytotoxic response to cisplatin in nasopharyngeal carcinoma cells, Anal. Methods 5 (1) (2013) 260–266.
[28] T.J. Moritz, D.S. Taylor, D.M. Krol, J. Fritch, J.W. Chan, Detection of doxorubicin-induced apoptosis of leukemic T-lymphocytes by laser tweezers Raman spectroscopy, Biomed. Opt. Express 1 (4) (2010) 1138–1147.
[29] F. Gasparri, M. Muzio, Monitoring of apoptosis of HL60 cells by Fourier-transform infrared spectroscopy, Biochem. J. 369 (2) (2003) 239–248.
[30] E. Lipiec, et al., Monitoring uvr induced damage in single cells and isolated nuclei using sr-ftir microspectroscopy and 3d confocal raman imaging, Analyst 139 (17) (2014) 4200–4209.
[31] L. Di Giambattista, et al., UVB radiation induced effects on cells studied by FTIR spectroscopy, Eur. Biophys. J. 39 (6) (2010) 929–934.
[32] C. Krafft, T. Knetschke, R.H.W. Funk, R. Salzer, Studies on stress-induced changes at the subcellular level by Raman microspectroscopic mapping, Anal. Chem. 78 (13) (2006) 4424–4429.
[33] K. Czamara, F. Petko, M. Baranska, A. Kaczor, Raman microscopy at the subcellular level: a study on early apoptosis in endothelial cells induced by Fas ligand and cycloheximide, Analyst 141 (2) (2016) 1390–1397.
[34] H. Yao, et al., Raman spectroscopic analysis of apoptosis of single human gastric cancer cells, Vib. Spectrosc. 50 (2) (2009) 193–197.
[35] E. Brauchle, S. Thude, S.Y. Brucker, K. Schenke-Layland, Cell death stages in single apoptotic and necrotic cells monitored by Raman microspectroscopy, Sci. Rep. 4 (1) (2015) 4698.
[36] J. Panza, J. Maier, Imaging, manipulation, and analysis of biomolecules, cells, and tissues V, Proc. SPIE 6441 (2007) 6441081–64410812.
[37] S. Verrier, I. Notingher, J.M. Polak, L.L. Hench, In Situ monitoring of cell death using Raman microspectroscopy, Biopolymers 74 (1–2) (2004) 157–162.
[38] A. Zoladek, F.C. Pascut, P. Patel, I. Notingher, Non-invasive time-course imaging of apoptotic cells by confocal Raman micro-spectroscopy, J. Raman Spectrosc. 42 (3) (2011) 251–258.
[39] G.A.R. Ahmed, F.A.R. Khorshid, T.A. Kumosani, FT-IR spectroscopy as a tool for identification of apoptosis-induced structural changes in A549 cells treated with PM 701, Int. J. Nano Biomater. 2 (1/2/3/4/5) (2009) 396.
[40] S. Gaudenzi, D. Pozzi, P. Toro, I. Silvestri, S. Morrone, A.C. Castellano, Cell apoptosis specific marker found by Fourier transform infrared spectroscopy, Spectroscopy 18 (3) (2004) 415–422.
[41] R. Buckmaster, F. Asphahani, M. Thein, J. Xu, M. Zhang, Detection of drug-induced cellular changes using confocal Raman spectroscopy on patterned single-cell biosensors, Analyst 134 (7) (2009) 1440.
[42] M. Okada, et al., Label-free Raman observation of cytochrome c dynamics during apoptosis, Proc. Natl. Acad. Sci. 109 (1) (2012) 28–32.
[43] H. Salehi, E. Middendorp, A.-G. Végh, S.-K. Ramakrishnan, C. Gergely, F.J.G. Cuisinier, Detection of apoptosis caused by anticancer drug paclitaxel in MCF-7 cells by confocal Raman microscopy, Nanoscale Imaging Sensing Actuation Biomed. Appl. X 8594 (2013) 85940F.
[44] A.J. Hobro, N.I. Smith, An evaluation of fixation methods: spatial and compositional cellular changes observed by Raman imaging, Vib. Spectrosc. 91 (2017) 31–45.
[45] Hickman, C. P., Roberts, L. S., Larson, A., I'Arson, H., Integrated Princple of Zoology, Mc-Graw Hill (2008) 922 pp.
[46] M.R. Nyden, G.P. Forney, K. Chittur, Spectroscopic quantitative analysis of strongly interacting systems: human plasma protein mixtures, Appl. Spectrosc. 42 (4) (1988) 588–594.

[47] G. Janatsch, J.D. Kruse-Jarres, R. Marbach, H.M. Heise, Multivariate calibration for assays in clinical chemistry using attenuated total reflection infrared spectra of human blood plasma, Anal. Chem. 61 (18) (1989) 2016–2023.
[48] D. Perez-Guaita, J. Ventura-Gayete, C. Pérez-Rambla, M. Sancho-Andreu, S. Garrigues, M. De La Guardia, Protein determination in serum and whole blood by attenuated total reflectance infrared spectroscopy, Anal. Bioanal. Chem. 404 (3) (2012) 649–656.
[49] E. Staniszewska-Slezak, A. Rygula, K. Malek, M. Baranska, Transmission versus transflection mode in FTIR analysis of blood plasma: is the electric field standing wave effect the only reason for observed spectral distortions? Analyst 140 (7) (2015) 2412–2421.
[50] J. Filik, M.D. Frogley, J.K. Pijanka, K. Wehbe, G. Cinque, Electric field standing wave artefacts in FTIR micro-spectroscopy of biological materials, Analyst 137 (4) (2012) 853–861.
[51] P. Bassan, et al., The inherent problem of transflection-mode infrared spectroscopic microscopy and the ramifications for biomedical single point and imaging applications, Analyst 138 (1) (2013) 144–157.
[52] E. Staniszewska-Slezak, et al., Plasma biomarkers of pulmonary hypertension identified by Fourier transform infrared spectroscopy and principal component analysis, Analyst 140 (7) (2015) 2273–2279.
[53] E. Staniszewska-Slezak, et al., A possible Fourier transform infrared-based plasma fingerprint of angiotensin-converting enzyme inhibitor-induced reversal of endothelial dysfunction in diabetic mice, J. Biophotonics 11 (2) (2018) e201700044.
[54] E. Staniszewska-Slezak, L. Mateuszuk, S. Chlopicki, M. Baranska, K. Malek, Alterations in plasma biochemical composition in NO deficiency induced by L-NAME in mice analysed by Fourier transform infrared spectroscopy, J. Biophotonics 9 (10) (2016) 1098–1108.
[55] Y.-C. Lai, K.C. Potoka, H.C. Champion, A.L. Mora, M.T. Gladwin, Pulmonary arterial hypertension: the clinical syndrome, Circ. Res. 115 (1) (2014) 115–130.
[56] A. Maton, et al., Human Biology and Health, Prentice Hall, Englewood Cliffs, NJ, USA, 1993.
[57] M.A. Giembycz, M.A. Lindsay, Pharmacology of the eosinophil, Pharmacol. Rev. 51 (2) (1999). 213 LP-340.
[58] H.F. Rosenberg, K.D. Dyer, P.S. Foster, Eosinophils: changing perspectives in health and disease, Nat. Rev. Immunol. 13 (2012) 9.
[59] P.F. Weller, The immunobiology of eosinophils, N. Engl. J. Med. 324 (16) (1991) 1110–1118.
[60] P.F. Weller, Human eosinophils, J. Allergy Clin. Immunol. 100 (3) (1997) 283–287.
[61] C. Rosales, Neutrophil: a cell with many roles in inflammation or several cell types? Front. Physiol. 9 (2018) 113.
[62] B.M. Fournier, C.A. Parkos, The role of neutrophils during intestinal inflammation, Mucosal Immunol. 5 (2012) 354.
[63] J.B. Cowland, N. Borregaard, Granulopoiesis and granules of human neutrophils, Immunol. Rev. 273 (1) (2016) 11–28.
[64] G.J. Puppels, H.S. Garritsen, G.M. Segers-Nolten, F.F. de Mul, J. Greve, Raman microspectroscopic approach to the study of human granulocytes, Biophys. J. 60 (5) (1991) 1046–1056.
[65] V. Witko-Sarsat, P. Rieu, B. Descamps-Latscha, P. Lesavre, L. Halbwachs-Mecarelli, Neutrophils: molecules, functions and pathophysiological aspects, Lab. Investig. 80 (2000) 617.
[66] R. Fenna, J. Zeng, C. Davey, Structure of the green heme in myeloperoxidase, Arch. Biochem. Biophys. 316 (1) (1995) 653–656.
[67] B.L. Salmaso, G.J. Puppels, P.J. Caspers, R. Floris, R. Wever, J. Greve, Resonance Raman microspectroscopic characterization of eosinophil peroxidase in human eosinophilic granulocytes, Biophys. J. 67 (1) (1994) 436–446.
[68] B.G.J.M. Bolscher, H. Plat, R. Wever, Some properties of human eosinophil peroxidase, a comparison with other peroxidases, Biochim. Biophys. Acta Protein Struct. Mol. Enzymol. 784 (2) (1984) 177–186.
[69] S.S. Sibbett, S.J. Klebanoff, J.K. Hurst, Resonance Raman characterization of the heme prosthetic group in eosinophil peroxidase, FEBS Lett. 189 (2) (1985) 271–275.

[70] A. Ramoji, et al., Toward a spectroscopic hemogram: Raman spectroscopic differentiation of the two most abundant leukocytes from peripheral blood, Anal. Chem. 84 (12) (2012) 5335−5342.
[71] H. Noothalapati, et al., Towards the development of a non-bioptic diagnostic technique for eosinophilic esophagitis using Raman spectroscopy, Vib. Spectrosc. 85 (2016) 7−10.
[72] H.-J. van Manen, Y.M. Kraan, D. Roos, C. Otto, Single-cell Raman and fluorescence microscopy reveal the association of lipid bodies with phagosomes in leukocytes, Proc. Natl Acad. Sci. U.S.A 102 (29) (2005) 10159−10164.
[73] A. Rygula, et al., Raman imaging highlights biochemical heterogeneity of human eosinophils versus human eosinophilic leukaemia cell line, Br. J. Haematol. 186 (5) (2019) 685−694. Available from: https://doi.org/10.1111/bjh.15971.
[74] K. Kaushansky, et al., Hematology, ninth ed., McGraw-Hill Education, New York, 2016.
[75] B.R. Wood, et al., Resonance Raman spectroscopy reveals new insight into the electronic structure of beta-hematin and malaria pigment, J. Am. Chem. Soc. 126 (30) (2004) 9233−9239.
[76] J. Dybas, Molecular spectroscopy studies on chosen hemoprotein adducts and their changes in biological systems, PhD thesis, Dokt. 2019/002, Jagiellonian University, 2019.
[77] K.M. Marzec, A. Rygula, B.R. Wood, S. Chlopicki, M. Baranska, High-resolution Raman imaging reveals spatial location of heme oxidation sites in single red blood cells of dried smears, J. Raman Spectrosc. 46 (1) (2014) 76−83.
[78] B.R. Wood, D. McNaughton, Raman excitation wavelength investigation of single red blood cells in vivo, J. Raman Spectrosc. 33 (7) (2002) 517−523.
[79] J. Dybas, et al., Raman spectroscopy as a sensitive probe of soft tissue composition—imaging of cross-sections of various organs vs. single spectra of tissue homogenates, TrAC. Trends Anal. Chem. 85 (2016) 117−127.
[80] X. Huang, K. Nakanishi, N. Berova, Porphyrins and metalloporphyrins: versatile circular dichroic reporter groups for structural studies, Chirality 12 (2000) 237−255.
[81] L.R. Milgrom, The Colours of Life, Oxford University Press, Oxford, 1997.
[82] B.R. Wood, B. Tait, D. McNaughton, Micro-Raman characterisation of the R to T state transition of haemoglobin within a single living erythrocyte, Biochim. Biophys. Acta Mol. Cell Res. 1539 (1−2) (2001) 58−70.
[83] J. Dybas, M. Grosicki, M. Baranska, K.M. Marzec, Raman imaging of heme metabolism: in situ in macrophages and Kupffer cells, Analyst 143 (14) (2018) 3489−3498.
[84] K.M. Marzec, et al., Raman microimaging of murine lungs: insight into the vitamin A content, Analyst 140 (7) (2015) 2171−2177.
[85] E. Staniszewska, K. Malek, M. Baranska, Rapid approach to analyze biochemical variation in rat organs by ATR FTIR spectroscopy, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 118 (2014) 981−986.
[86] A.H. Fischer, K.A. Jacobson, J. Rose, R. Zeller, Hematoxylin eosin staining tissue and cell sections, Cold Spring Harb. Protoc. 3 (5) (2017) 4986−4988.
[87] S.E. Dunsmore, D.E. Rannels, Extracellular matrix biology in the lung, Am. J. Physiol. 270 (1 Pt 1) (1996) L3−L27.
[88] A. Cho, V.M. Howell, E.K. Colvin, The extracellular matrix in epithelial ovarian cancer—a piece of a puzzle, Front. Oncol. 5 (February) (2015) 245.
[89] A.D. Theocharis, S.S. Skandalis, C. Gialeli, N.K. Karamanos, Extracellular matrix structure, Adv. Drug. Deliv. Rev. 97 (2016) 4−27.
[90] J.L. Leight, A.P. Drain, V.M. Weaver, Extracellular matrix remodeling and stiffening modulate tumor phenotype and treatment response, Annu. Rev. Cancer Biol. 1 (1) (2017) 313−334.
[91] C. Bonnans, J. Chou, Z. Werb, Remodelling the extracellular matrix in development and disease, Nat. Rev. Mol. Cell Biol. 15 (12) (2014) 786−801.
[92] P. Lu, K. Takai, V.M. Weaver, Z. Werb, Extracellular matrix degradation and remodeling in development and disease, Cold Spring Harb. Perspect. Biol. 3 (12) (2011) 1−24.
[93] G. Vlahovic, J. Crawford, Activation of tyrosine kinases in cancer, Oncology 8 (6) (2003) 531−538.

[94] K. Chrabaszcz, K. Kochan, A. Fedorowicz, A. Jasztal, E. Buczek, L.S. Leslie, et al., FT-IR- and Raman-based biochemical profiling of the early stage of pulmonary metastasis of breast cancer in mice, Analyst 143 (9) (2018) 2042–2050.
[95] Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, Peter Walter, Molecular Biology of the Cell, fourth ed., Garlands Science, New York, 2002.
[96] R.P. Rava, J.J. Baraga, M.S. Feld, Near infrared fourier transform Raman spectroscopy of human artery, Spectrochim. Acta Part A Mol. Spectrosc. 47 (3–4) (1991) 509–512.
[97] B.E. Sumpio, J. Timothy Riley, A. Dardik, Cells in focus: endothelial cell, Int. J. Biochem. Cell Biol. 34 (12) (2002) 1508–1512.
[98] B.G. Frushour, J.L. Koenig, Raman scattering of collagen, gelatin, and elastin, Biopolymers 14 (2) (1975) 379–391.
[99] E. Green, R. Ellis, P. Winlove, The molecular structure and physical properties of elastin fibers as revealed by Raman microspectroscopy, Biopolymers 89 (11) (2008) 931–940.
[100] J.J. Baraga, M.S. Feld, R.P. Rava, In situ optical histochemistry of human artery using near infrared Fourier transform Raman spectroscopy, Proc. Natl Acad. Sci. U.S.A. 89 (8) (1992) 3473–3477.
[101] S. Chlopicki, R.J. Gryglewski, Angiotensin coverting enzyme (ACE) and hydroxy methyl glutaryl-CoA (HMG-CoA) reductase inhibitors in the forefront of pharmacology of endothelium, Pharmacol. Rep. 57 (Suppl.) (2005) 86–96.
[102] R. Manoharan, J.J. Baraga, M.S. Feld, R.P. Rava, Quantitative histochemical analysis of human artery using Raman spectroscopy, J. Photochem. Photobiol. B Biol. 16 (2) (1992) 211–233.
[103] L. Silveira, S. Sathaiah, R.A. Zângaro, M.T.T. Pacheco, M.C. Chavantes, C.A.G. Pasqualucci, Correlation between near-infrared Raman spectroscopy and the histopathological analysis of atherosclerosis in human coronary arteries, Lasers Surg. Med. 30 (4) (2002) 290–297.
[104] C.H. Liu, W.B. Wang, A. Alimova, V. Sriramoju, V. Kartazayev, R.R. Alfano, Monitoring changes of proteins and lipids in laser welded aorta tissue using Raman spectroscopy and basis biochemical component analyses, Proc. SPIE 7175 (2009) 717504.
[105] A. Lattermann, et al., Characterization of atherosclerotic plaque depositions by Raman and FTIR imaging, J. Biophotonics 6 (1) (2013) 110–121.
[106] C. Matthäus, et al., In vivo characterization of atherosclerotic plaque depositions by Raman-probe spectroscopy and in vitro coherent anti-stokes Raman scattering microscopic imaging on a rabbit model, Anal. Chem. 84 (18) (2012) 7845–7851.
[107] K. Majzner, A. Kaczor, N. Kachamakova-Trojanowska, A. Fedorowicz, S. Chlopicki, M. Baranska, 3D confocal Raman imaging of endothelial cells and vascular wall: perspectives in analytical spectroscopy of biomedical research, Analyst 138 (2) (2013) 603–610.
[108] M. Pilarczyk, et al., A novel approach to investigate vascular wall in 3D: combined Raman spectroscopy and atomic force microscopy for aorta en face imaging, Vib. Spectrosc. 75 (2014) 39–44.
[109] M. Pilarczyk, A. Rygula, L. Mateuszuk, S. Chlopicki, M. Baranska, A. Kaczor, Multi-methodological insight into the vessel wall cross-section: Raman and AFM imaging combined with immunohistochemical staining, Biomed. Spectrosc. Imaging 2 (3) (2013) 191–197.
[110] M.Z. Pacia, L. Mateuszuk, S. Chlopicki, M. Baranska, A. Kaczor, Biochemical changes of the endothelium in the murine model of NO-deficient hypertension, Analyst 140 (7) (2015) 2178–2184.
[111] M.Z. Pacia, et al., 3D Raman imaging of systemic endothelial dysfunction in the murine model of metastatic breast cancer, Anal. Bioanal. Chem. 408 (13) (2016) 3381–3387.
[112] M. Pilarczyk, et al., Endothelium in Spots – High-Content Imaging of Lipid Rafts Clusters in db/db Mice, PLoS One 9 (8) (2014) e106065.
[113] M. Maase, et al., Combined Raman- and AFM-based detection of biochemical and nanomechanical features of endothelial dysfunction in aorta isolated from ApoE/LDLR −/− mice, Nanomed. Nanotechnol. Biol. Med. 16 (2019) 97–105.
[114] K.M. Marzec, et al., Visualization of the biochemical markers of atherosclerotic plaque with the use of Raman, IR and AFM, J. Biophotonics 7 (9) (2014) 744–756.

[115] A. Rygula, et al., Identification of a biochemical marker for endothelial dysfunction using Raman spectroscopy, Analyst 140 (7) (2015) 2185–2189.
[116] M.Z. Pacia, et al., Rapid biochemical profiling of endothelial dysfunction in diabetes, hypertension and cancer metastasis by hierarchical cluster analysis of Raman spectra, J. Raman Spectrosc. 47 (11) (2016) 1310–1317.
[117] E.H.C. Tang, P.M. Vanhoutte, Endothelial dysfunction: a strategic target in the treatment of hypertension? Pflügers Arch. Eur. J. Physiol. 459 (6) (2010) 995–1004.
[118] H.D. Intengan, E.L. Schiffrin, Vascular remodeling in hypertension, Hypertension 38 (3) (2001) 581–587.
[119] M.M. Steed, N. Tyagi, U. Sen, D.A. Schuschke, I.G. Joshua, S.C. Tyagi, Functional consequences of the collagen/elastin switch in vascular remodeling in hyperhomocysteinemic wild-type, eNOS −/−, and iNOS −/− mice, Am. J. Physiol. Cell. Mol. Physiol 299 (3) (2010) L301–L311.
[120] J.E. Wagenseil, R.P. Mecham, Elastin in large artery stiffness and hypertension, J. Cardiovasc. Transl. Res. 5 (3) (2012) 264–273.
[121] H. Wolinsky, Effects of hypertension and its reversal on the thoracic aorta of male and female rats, Circ. Res. 28 (6) (1971) 622–637.
[122] F. Severcan, G. Gorgulu, S.T. Gorgulu, T. Guray, Rapid monitoring of diabetes-induced lipid peroxidation by Fourier transform infrared spectroscopy: evidence from rat liver microsomal membranes, Anal. Biochem. 339 (1) (2005) 36–40.
[123] F. Severcan, N. Kaptan, B. Turan, Fourier transform infrared spectroscopic studies of diabetic rat heart crude membranes, Spectroscopy 17 (2,3) (2003) 569–577.
[124] D.A. Scott, et al., Diabetes-related molecular signatures in infrared spectra of human saliva, Diabetol. Metab. Syndr. 2 (2010) 48.
[125] F. Severcan, O. Bozkurt, R. Gurbanov, G. Gorgulu, FT-IR spectroscopy in diagnosis of diabetes in rat animal model, J. Biophotonics 3 (8–9) (2010) 621–631.
[126] F. Severcan, N. Toyran, N. Kaptan, B. Turan, Fourier transform infrared study of the effect of diabetes on rat liver and heart tissues in the CH region, Talanta 53 (1) (2000) 55–59.
[127] E. Goormaghtigh, J.-M. Ruysschaert, V. Raussens, Evaluation of the information content in infrared spectra for protein secondary structure determination, Biophys. J. 90 (8) (2006) 2946–2957.
[128] P. Lasch, W. Haensch, D. Naumann, M. Diem, Imaging of colorectal adenocarcinoma using FT-IR microspectroscopy and cluster analysis, Biochim. Biophys. Acta Mol. Basis Dis. 1688 (2) (2004) 176–186.
[129] M.J. Walsh, A. Kajdacsy-Balla, S.E. Holton, R. Bhargava, Attenuated total reflectance Fourier-transform infrared spectroscopic imaging for breast histopathology, Vib. Spectrosc. 60 (2012) 23–28.
[130] E. Gazi, et al., A correlation of FTIR spectra derived from prostate cancer biopsies with gleason grade and tumour stage, Eur. Urol. 50 (2006) 750–761.
[131] B.R. Wood, L. Chiriboga, H. Yee, M.A. Quinn, D. McNaughton, M. Diem, Fourier transform infrared (FTIR) spectral mapping of the cervical transformation zone, and dysplastic squamous epithelium, Gynecol. Oncol. 93 (1) (2004) 59–68.
[132] M.J. Baker, E. Gazi, M.D. Brown, J.H. Shanks, P. Gardner, N.W. Clarke, FTIR-based spectroscopic analysis in the identification of clinically aggressive prostate cancer, Br. J. Cancer 99 (11) (2008) 1859–1866.
[133] M.J. Baker, E. Gazi, M.D. Brown, J.H. Shanks, N.W. Clarke, P. Gardner, Investigating FTIR based histopathology for the diagnosis of prostate cancer, J. Biophotonics 2 (1–2) (2009) 104–113.
[134] N. Amharref, et al., Brain tissue characterisation by infrared imaging in a rat glioma model, Biochim. Biophys. Acta 1758 (7) (2006) 892–899.
[135] R. Bhargava, Towards a practical Fourier transform infrared chemical imaging protocol for cancer histopathology, Anal. Bioanal. Chem. 389 (4) (2007) 1155–1169.
[136] A.H. Colagar, M.J. Chaichi, T. Khadjvand, Fourier transform infrared microspectroscopy as a diagnostic tool for distinguishing between normal and malignant human gastric tissue, J. Biosci. 36 (4) (2011) 669–677.

[137] R.P. Joosten, et al., A series of PDB related databases for everyday needs, Nucleic Acids Res. 39 (Database Issue) (2011) D411–D419.
[138] Ca Orengo, et al., Classifying a protein in the CATH database of domain structures, Acta Crystallogr. Sect. D. Biol. Crystallogr. 54 (6) (1998) 1155–1167.
[139] K.A. Oberg, J. Ruysschaert, E. Goormaghtigh, Rationally selected basis proteins: a new approach to selecting proteins for spectroscopic secondary structure analysis, Protein Sci. 12 (9) (2003) 2015–2031.
[140] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, Biopolymers 22 (12) (1983) 2577–2637.
[141] J.T. Pelton, L.R. McLean, Spectroscopic methods for analysis of protein secondary structure, Anal. Biochem. 277 (2) (2000) 167–176.
[142] K. Griebenow, A.M. Klibanov, Lyophilization-induced reversible changes in the secondary structure of proteins, Proc. Natl Acad. Sci. U.S.A 92 (24) (1995) 10969–10976.
[143] H. R. Costantino, K. Griebenow, P. Mishra, R. Langer, and A. M. Klibanov. Fourier-transform infrared spectroscopic investigation of protein stability in the lyophilized form. Biochim. Biophys. Acta Protein Struct. Mol. Enzymol.Bio, 1253 (1995) 69–74.
[144] T.P. Wrobel, K. Majzner, M. Baranska, Protein profile in vascular wall of atherosclerotic mice analyzed ex vivo using FT-IR spectroscopy, Spectrochim. Acta. A. Mol. Biomol. Spectrosc. 96 (2012) 940–945.

CHAPTER 8

Interpretation of vibrational optical activity spectra of proteins

Jana Hudecová[1,2] **and Petr Bouř**[1]

[1]*Institute of Organic Chemistry and Biochemistry, Academy of Sciences, Prague, Czech Republic,*
[2]*Department of Optics, Palacký University Olomouc, Olomouc, Czech Republic*

8.1 Introduction

Peptides and proteins probably form the most versatile class of biomolecules, adopting many functions and serving as building blocks in all living organisms. Determination of their structure and consecutive clarification of its relation to the functionality is one of the main tasks of molecular biology and biochemistry. An immense amount of information on this field was obtained by the methods of X-ray crystallography [1] and nuclear magnetic resonance (NMR) [2]. Both methods can provide data with up to atomic resolution, but they cannot be applied universally. For X-ray crystallography, the growing of protein crystals is often a very challenging task, and the "native" structure can change during this process. Except for small molecules such crystals must contain "heavy" atoms facilitating the phase determination. NMR spectroscopy is more flexible in that it can be applied both for liquids and solids, but spectral interpretation in terms of the structure is even more difficult, typically requiring laborious cross-peak assignments and parallel molecular dynamics (MD) modeling [3]. Recent advances of imaging techniques and cryomicroscopy promise to complement the classical "high-resolution" approaches also for proteins [4].

Vibrational optical activity (VOA) typically provides a lower resolution than X-ray and NMR, but can conveniently complement the mainstream approaches. For example, absolute configuration can be determined by VOA, while basic NMR is blind to it [5]. VOA spectra of conformer mixtures are sums of individual components, whereas NMR often provides an averaged signal only [6]. VOA and other optical spectroscopies usually make it easier to vary experimental conditions, such as pH, temperature, concentration, and the solvent [7]. Interpretation of VOA spectra used to be a problem in the past, but as documented below advances in computational technology make them applicable almost routinely for peptide and protein studies.

Vibrational Spectroscopy in Protein Research.
DOI: https://doi.org/10.1016/B978-0-12-818610-7.00008-6

Before going to concrete applications we should mention that VOA is only one of many chiroptical techniques that explore the interaction of left- and right-circularly polarized (LCP, RCP) light with molecules [8]. Circularly or linearly polarized light is used as a chiral probe. The first observation of rotation of the polarization plane by Luis Pasteur gave birth to optical rotation and optical rotatory dispersion [9–11]. These methods are regularly used, for example, by chemists to characterize chiral compounds [12]. Later, the absorption difference of LCP and RCP visible or ultraviolet light, electronic circular dichroism, was found more convenient for many applications including proteins [13,14]. A rarer technique is circularly polarized luminescence, measuring differences in emission of the LCP and RCP light [15,16]. It was observed for the first time in the 1940s, on chiral crystals [17].

VOA started to be vigorously promoted in the 1970s, because vibrational bands are typically more numerous, narrower, and better resolved than the electronic ones. Hence they contain more information, and effectively make possible "higher-resolution" experiments, also in terms of molecular structure. Vibrational transitions are more localized than the electronic ones, and easier to calculate, requiring electronic ground state only [18]. Today, under VOA we usually understand two techniques, Raman optical activity (ROA), first measured in 1973 [19], and vibrational circular dichroism (VCD), first observed in 1974 [20].

Measured quantities include difference in Raman scattering intensities of RCP and LCP light for ROA [7]

$$I^{\mathrm{ROA}} = I(R) - I(L) \tag{8.1}$$

and difference in absorption of LCP and RCP infrared light for VCD [18]

$$\Delta A(\nu) = A_L(\nu) - A_R(\nu). \tag{8.2}$$

This small inconsistency in definition (right−left for ROA versus left−right for VCD) appeared historically, as these two phenomena were expected to be complementary and often opposite, as given by their different physical nature.

Similarly to the absorption

$$A(\nu) = \frac{1}{2}[A_R(\nu) + A_L(\nu)], \tag{8.3}$$

CD spectra can be normalized to concentration (c, in mol/L) and optical path length (l, in cm) and expressed as the differential molar absorption coefficient

$$\Delta\varepsilon(\nu) = \varepsilon_L(\nu) - \varepsilon_R(\nu) = \frac{\Delta A}{cl}. \tag{8.4}$$

Very useful dimensionless quantities are also the anisotropy ratio [21]

$$g = \frac{\Delta A}{A} = \frac{2[A_L - A_R]}{A_R + A_L} \tag{8.5}$$

and the circular intensity difference [22]

$$\Delta = \frac{I^{\mathrm{ROA}}}{I^{\mathrm{RAM}}} = \frac{I(R) - I(L)}{I(R) + I(L)}. \tag{8.6}$$

Their typical value for organic molecules is about 10^{-4}. Both VCD and ROA signals are thus much weaker than the unpolarized parent signals. This causes many practical problems, such as artifacts, false signals stemming from polarization optics, etc. The artifacts can be even larger than those originated in the investigated sample. A good way to avoid them is to measure both enantiomers, ideally giving "mirror image" signals, although this is often practically not possible [23].

VOA ideally combines the sensitivity of chiroptical methods to the spatial arrangement of atoms in molecules with the richness of vibrational spectra. Other advantages are similar as those for the parent Raman and IR spectroscopies. The usual VCD spectral range, however, is narrower than for IR, typically from 800 to 4000 cm^{-1}. The ROA span is broader and the newest instruments seem to be able to measure the "full" range, approximately from 40 to 4000 cm^{-1} [24]. This makes possible, for example, to study low-frequency vibrationsin proteins, such as amide modes IV−VII, helix breathing, torsions, and skeletal deformations [25].

As for Raman, water is an ideal solvent for ROA. This spectroscopy is thus particularly well suited for biomolecular applications [7]. For VCD, water (H_2O) is more problematic, because its 1650 cm^{-1} absorption band overlaps with the main protein "amide I" band (largely C = O bond stretching). Therefore proteins are regularly "deuterated" for VCD studies, and the measurement runs in D_2O. This applies to other systems as well and deuterated organic solvents are frequently employed to reduce the absorption within the spectroscopic window. On the other hand, VCD is not so much affected by fluorescent impurities in the sample. These can totally prevent ROA measurement.

Apart from proteins [26], VOA spectroscopy provides valuable information on carbohydrates [27,28], polypeptide and carbohydrate components of intact glycoproteins [29], and nucleic acids [30,31]. Typically, the regular structure of biopolymers (peptides/proteins, nucleic acids) provides intense and well-recognizable spectral patterns [32].

8.2 Theory and calculations

Experimental spectra can be interpreted in many ways, such as by comparison to similar compounds or using numerical analysis of a set of spectra [33,34]. Nevertheless,

quantum-chemical simulations have become established as the most universal method to link spectral shapes to properties of studied systems. Therefore let us summarize some basic theoretical foundations of VOA. It is true that calculations of chiroptical spectra can nowadays be done at a user level, using free or commercial computer programs (Gaussian [35], Turbomole [36], Dalton [37], ADF [38], CADPAC [39], etc.), but basic knowledge of the underlying principles is necessary for a good understanding and application.

We may start from the Schrödinger equation. Stationary wave functions (states) and energies of a molecule are predicted by

$$H\psi = E\psi, \tag{8.7}$$

where H is the Hamiltonian, $\psi = \psi(\boldsymbol{r}, \boldsymbol{R}, t)$ is the wave function dependent on positions of electrons, $\boldsymbol{r}$, and nuclei, $\boldsymbol{R}$, and time t, and E is the energy.

For a relativistic case, the Dirac equation needs to be used instead, nevertheless the formalism is for our purpose similar. To apply the Schrödinger equation (Eq. 8.7), for "big" molecules many simplifications have to be adopted. One of them is the Born−Oppenheimer approximation, when the wave function is assumed to be a product of an electronic and nuclear part

$$\psi(\boldsymbol{r}, \boldsymbol{R}) = \varphi(\boldsymbol{r}, \boldsymbol{R})\Phi(\boldsymbol{R}). \tag{8.8}$$

The electronic wave function $\varphi(\boldsymbol{r}, \boldsymbol{R})$ depends also on nuclear positions $\boldsymbol{R}$, but only in a parametric way. We may do this because the atomic nuclei are much heavier (> 1800 times) than electrons. Except for a small error, the Schrödinger equation can then be separated into two, one for electrons and one for the nuclei:

$$H_e\varphi(\boldsymbol{r}, \boldsymbol{R}) = \varepsilon(\boldsymbol{R})\varphi(\boldsymbol{r}, \boldsymbol{R}), \tag{8.9}$$

$$(T_n + \varepsilon(\boldsymbol{R}))\Phi(\boldsymbol{R}) = E\Phi(\boldsymbol{R}), \tag{8.10}$$

where H_e is the "electronic" Hamiltonian and T_n is the nuclear kinetic energy. The electronic energy $\varepsilon(\boldsymbol{R})$ thus becomes a potential for the nuclei. Typically, vibrational energies (E) are much smaller than the electronic ones, including energy differences (this is also the reason why we can talk about "vibrational" spectroscopy).

Determination of the vibrational energies is notably simplified if we assume that the molecule is nearly rigid. Then the nuclei move around their equilibrium positions and the potential can be approximately written as

$$\varepsilon(\boldsymbol{R}) = \frac{1}{2}\sum_{I=1}^{3M}\sum_{J=1}^{3M}\left(\frac{\partial^2 \mathrm{E}}{\partial R_I \partial R_J}\right)_{R^0}\left(R_I - R_I^0\right)\left(R_J - R_J^0\right), \tag{8.11}$$

where the matrix of the second energy derivatives, called also "force field" or "Hessian", completely determines molecular vibrational energies. In this form, the vibrational Schrödinger equation (Eq. 8.10) for any molecule can be transformed to

$$H_n = T_n + \varepsilon = \sum_{j=1}^{3M} \frac{1}{2}\left(\dot{Q}_j^2 + \Lambda_{jj} Q_j^2\right), \tag{8.12}$$

where Q_j are the vibrational normal mode coordinates. They are obtained by a linear transformation from the Cartesian ones so that

$$\left(R_{A\beta} - R_{A\beta}^0\right) = \sum_j S_{A\beta,j} Q_j, \tag{8.13}$$

where the $S_{A\beta,j}$ elements belong to the so-called "S-matrix." Because the nuclear Hamiltonian in Eq. (8.12) has a form of a sum of harmonic oscillators, molecular vibrational energies in the harmonic approximation are obtained as

$$E = \sum_{j=1}^{3M} \left(\frac{1}{2} + n_j\right)\hbar\omega_j, \tag{8.14}$$

where ω_j are vibrational normal mode frequencies, $\hbar$ is the reduced Planck constant, and n_j are the quantum numbers.

The electronic part (Eq. 8.9) also needs some simplifications to be usable in practice. Similarly to the harmonic approximation for the vibrational problem, the Hartree−Fock method is probably the simplest reasonable approach for the electronic problem. Today, it is usually not considered accurate enough for VOA simulations, and more advanced methods are needed, such as the Møller−Plesset perturbation approach (MP2, MP4, etc.) [40], configurational interaction, or coupled cluster theory [41−43]. These "traditional" wave function−based approaches can be quite accurate, but really precise computations may require excessive computational times.

From about the 1990s, the density functional theory (DFT) started to dominate for electronic quantum-chemical computations on applied chemical problems [44,45]. DFT is in principle also based on the Schrödinger equation, but the electron density is used instead of the wave function as the basic mathematical object, which enables formulations leading to faster computations. The DFT formalism is commonly based on the Kohn−Sham method [46,47] analogous to HF theory, but incorporating higher degrees of electron correlation. The price to pay is that a usable "Schrödinger equation" for the density does not exist and inter-electron forces are described by potentials−functionals dependent on empirical parameters.

Choosing appropriate functional requires some experience. Harmonic force fields and vibrational frequencies of reasonable quality can be obtained, for example, by "hybrid" B3LYP and B3PW91 functionals [48,49] combining the HF and "pure" DFT/Kohn−Sham formalisms. Similarly, attention must be paid to choosing a reasonable basis of atomic orbitals, as for any quantum-chemical computations.

To describe the interaction of molecules with the light (electromagnetic radiation) we would need the Maxwell equations. They were derived for electric (electric intensity **E**, and induction **D**) and magnetic quantities (magnetic intensity **H**, and induction **B**) by James Clerk Maxwell (1831−79). They are also relativistically invariant and compatible with the quantum theory. We will obviously not derive spectroscopically usable equations in detail. But similarly as for the Schrödinger equation, practical computations imply some approximations that may be useful to mention. For example, the electromagnetic field is supposed not to change much over the molecule, that is the size of the studied system is supposed to be much smaller than the wavelength of the light [7]. Then the molecule can be represented by multipoles and the minimalistic interaction Hamiltonian (to be added to the Schrödinger equation) suitable for the prediction of VOA is

$$H^{int} = -\boldsymbol{\mu}\cdot\boldsymbol{E} - \mathbf{m}\cdot\boldsymbol{B} - \frac{1}{3}\boldsymbol{\Theta}\cdots\nabla\boldsymbol{E}, \tag{8.15}$$

where $\boldsymbol{\mu}$, **m**, and $\boldsymbol{\Theta}$ are electric and magnetic dipole and electric quadrupole moments, respectively, defined as

$$\mu_\alpha = \sum_i q_i r_{i\alpha}, \tag{8.16}$$

$$m_\alpha = \sum_i \frac{q_i}{2m_i}\epsilon_{\alpha\beta\gamma} r_{i\beta} p_{i\gamma}, \tag{8.17}$$

$$\Theta_{\alpha\beta} = \sum_i \frac{q_i}{2}\left(3r_{i\alpha}r_{i\beta} - r_i^2\delta_{\alpha\beta}\right), \tag{8.18}$$

where particle i has position vector r_i, charge q_i, mass m_i, and momentum p_i. The Einstein summation convention is assumed for the Greek indices, $\delta_{\alpha\beta}$ is the Kronecker delta, and $\epsilon_{\alpha\beta\gamma}$ is the Levi−Civita antisymmetric tensor.

From these theoretical foundations we can find parameters that are important in order to interpret experimental VOA spectra. For isotropic samples, for example, we can define the dipole strength of a $j \rightarrow n$ transition as

$$D_{jn} = \langle n|\mu|j\rangle\cdot\langle j|\mu|n\rangle \tag{8.19}$$

It is related to the integral (area) of an absorption band in a spectrum according to

$$D = 9.184\times10^{-3}\int_{\text{ABSband}}\frac{\varepsilon(\lambda)}{\lambda}d\lambda. \tag{8.20}$$

Similarly, for CD we need the rotatory (or "rotational") strength

$$R_{jn} = \text{Im}\left(\langle n|\mu|j\rangle\cdot\langle j|\text{m}|n\rangle\right) \tag{8.21}$$

related to spectral intensity $\Delta\varepsilon$ as

$$R = 2.296 \times 10^{-3} \int_{\text{CDband}} \frac{\Delta\varepsilon(\lambda)}{\lambda} d\lambda. \tag{8.22}$$

Eqs. (8.20) and (8.22) are valid for the usual units of L/mol/cm for ε and $\Delta\varepsilon$; the strengths D and R are in debye^2 (1 debye is 10^{-18} esu cm or 3.33546 $\times 10^{-30}$ C m).

Within the harmonic approximation only the linear term of expansion of the dipole moment around the equilibrium nuclear position as a function of Q_i is taken into account

$$\mu_\alpha^i = \sqrt{\frac{\hbar}{2\omega_i}} \left(\frac{\partial \mu_\alpha}{\partial Q_i} \right)_0 = \sqrt{\frac{\hbar}{2\omega_i}} \sum_{A,\beta} P_{\beta\alpha}^A S_{A\beta,i}, \tag{8.23}$$

where $S_{A\beta,i}$ is the S-matrix (Eq. 8.13) and $P_{\beta\alpha}^A$ is the atomic polar tensor (APT) defined as

$$P_{\beta\alpha}^A = \left(\frac{\partial \mu_\alpha}{\partial R_{A\beta,}} \right)_0 = E_{\beta\alpha}^A + Z_A e \delta_{\beta\alpha}, \tag{8.24}$$

where Z_A is the atomic number. While evaluation of the nuclear contribution $Z_A e\delta_{\beta\alpha}$ is relatively simple, the electronic contribution $E_{\beta\alpha}^A$ needs to be obtained from perturbation theory [50].

The magnetic dipole moment is obtained in a similar way as

$$m_\alpha^i = \sqrt{2\hbar^3\omega_i} \left(\frac{\partial m}{\partial \Pi_i} \right)_0 = \sqrt{2\hbar^3\omega_i} \sum_{A,\beta} M_{\beta\alpha}^A S_{A\beta,i}, \tag{8.25}$$

where Π_i is the nuclear momentum and $M_{\beta\alpha}^A$ is the atomic axial tensor (AAT) defined as

$$M_{\beta\alpha}^A = \left(\frac{\partial m_\beta}{\partial \Pi_\alpha^A} \right)_0 = I_{\beta\alpha}^A + \frac{Z_A e}{2M_A} \varepsilon_{\beta\alpha\gamma} R_{A\gamma}^0, \tag{8.26}$$

where M_A is nuclear mass, and Π_α^A is nuclear momentum. Slightly different definitions of AAT can be found in the literature. The nuclear part (second term in Eq. 8.26) is again easy to evaluate. But the electronic contribution to the magnetic dipole moment $I_{\beta\alpha}^A$ vanishes in the Born−Oppenheimer approximation. This has been overcome in the vibronic coupling theory based on electronic excited states (wave functions) [51]. More comfortable calculations based on the electronic ground state were enabled by the magnetic field perturbation formalism of Stephens [52]. Its analytical implementation by Amos first appeared in CADPAC [53].

The molecular theory of ROA was started by works of Barron and Atkins [54], Buckingham [22], and others. In analogy to the transition dipole moments for VCD, the central quantities here are transition polarizabilities. We can imagine that the electromagnetic field (light) induces oscillating moments in the molecule. For example, the electric dipole moment is

$$\mu_\alpha = \alpha_{\alpha\beta}E_\beta + \frac{1}{\omega}G'_{\alpha\beta}\frac{\partial \mathrm{B}_\beta}{\partial \mathrm{t}} + \frac{1}{3}A_{\alpha,\beta\gamma}\nabla_\beta E_\gamma + \cdots, \tag{8.27}$$

where ω is the angular frequency of the incident laser beam, and the polarizabilities are defined as

$$\alpha_{\alpha\beta} = \frac{2}{\hbar}\sum_{j\neq n}\frac{\omega_{jn}}{\omega_{jn}^2 - \omega^2}\mathrm{Re}\left(\langle n|\mu_\alpha|j\rangle\langle j|\mu_\beta|n'\rangle\right), \tag{8.28}$$

$$G'_{\alpha\beta} = -\frac{2}{\hbar}\sum_{j\neq n}\frac{\omega}{\omega_{jn}^2 - \omega^2}\mathrm{Im}\left(\langle n|\mu_\alpha|j\rangle\langle j|m_\beta|n'\rangle\right), \tag{8.29}$$

$$A_{\alpha,\beta\gamma} = \frac{2}{\hbar}\sum_{j\neq n}\frac{\omega_{jn}}{\omega_{jn}^2 - \omega^2}\mathrm{Re}\left(\langle n|\mu_\alpha|j\rangle\langle j|\Theta_{\beta\gamma}|n'\rangle\right), \tag{8.30}$$

where $\alpha_{\alpha\beta}$ is the electric-dipole−electric-dipole, $G'_{\alpha\beta}$ is the electric-dipole−magnetic-dipole and $A_{\alpha,\beta\gamma}$ is the electric-dipole−electric-quadrupole polarizability. In Eqs. (8.28)−(8.30) we defined the transition polarizabilities needed for the Raman scattering, where n, j, and n' denote the initial, excited, and final state of a molecule, $\omega_{jn} = \omega_j - \omega_n$ is the angular frequency difference, and $\hbar$ is the reduced Planck constant. These expressions are valid for a nonresonance case where $\omega_{jn} \neq \omega$ [18].

Raman and ROA scattering intensities of isotropic samples (gases, liquids, solutions) are then dependent on tensor invariants [55]

$$\alpha = \frac{1}{3}\alpha_{\alpha\alpha} \tag{8.31}$$

$$\alpha(G') = \frac{1}{3}G'_{\alpha\alpha} \tag{8.32}$$

$$\beta(\alpha)^2 = \frac{1}{2}\left[3\alpha_{\alpha\beta}\alpha_{\alpha\beta} - \alpha_{\alpha\alpha}\alpha_{\beta\beta}\right] \tag{8.33}$$

$$\beta(G')^2 = \frac{1}{2}\left[3\alpha_{\alpha\beta}G'_{\alpha\beta} - \alpha_{\alpha\alpha}G'_{\beta\beta}\right] \tag{8.34}$$

$$\beta(\mathrm{A})^2 = \frac{\omega}{2}\alpha_{\alpha\beta}\epsilon_{\alpha\gamma\delta}A_{\gamma,\delta\beta} \tag{8.35}$$

so that

$$I^{\mathrm{RAM}} = I(R) + I(L) = 4K\left[D_1\alpha^2 + D_2\beta(\alpha)^2\right], \tag{8.36}$$

$$I^{\mathrm{ROA}} = I(R) - I(L) = \frac{8K}{c}\left[D_3\alpha G' + D_4\beta(G')^2 + D_5\beta(A)^2\right], \tag{8.37}$$

where the coefficients D_{1-5} depend on both the scattering geometry and modulation scheme used for the ROA experimental setup [7]. Depending on the scattering angle between directions of the incident and scattered light, one can encounter three most frequent geometries, forward (0 degree), right-angle (90 degrees), and backward (180 degrees) scattering. The backscattering geometry is currently considered to be the best experimental strategy for proteins due to high signal intensity and low artifact level [56].

Polarizations of the incident and scattered light allow for another three basic schemes. Originally, incident circular polarization (ICP) [19], when the incident laser beam switches its polarization between RCP and LCP, and unpolarized scattered radiation is detected, was used. The only broadly available commercial instrument so far (Biotools) uses scattered circular polarization (SCP) [57], that is, unpolarized incident light enters the sample and the difference in RCP and LCP components of the scattered light is measured. The SCP instrument is considered to be more stable, and within the far-from-resonance approximation the ICP and SCP spectra are the same [58] and the coefficients in Eqs. (8.36) and (8.37) are for the backscattering geometry: $D_1 = 45$, $D_2 = 7$, $D_3 = 0$, $D_4 = 12$, and $D_5 = 4$ [7]. The two remaining forms, which modulate polarization in both the incident and scattered radiation, either in a synchronous (DCPI) or out-of-phase (DCPII) way, provide different ROA spectra [55,59].

To evaluate the transition polarizabilities for fundamental vibrational transitions we may employ the Placzek's approximation [60] and find that derivatives of electronic polarizabilities with respect to the normal mode coordinates are needed to simulate the Raman and ROA spectra.

The first calculation of ROA spectra appeared in 1989 [61,62], using a static limit ($\omega = 0$) of the $\left(1/\omega\right)\mathrm{G}'(\omega)$ ratio [63] and numerical differentiation. Analytical implementation appeared in 2007 within the HF approach [64]. Later it was improved to be computationally mode efficient [65] and implemented into the Gaussian09 program [66] by Cheeseman and coworkers [67]. It can be readily used also within DFT.

Nevertheless, users may be also aware that tensors AAT, A, and G′ depend on the choice of the coordinate system, and so may computed VCD and ROA intensities. In common software, this dependence is often removed by implementing the gauge-independent atomic orbitals (GIAOs, also called London or field-dependent AOs) [50,68–72]. They are obtained from the usual atomic basis sets by multiplication with a phase factor dependent on the magnetic field

$$\phi_\lambda^{\mathrm{GIAO}}(\boldsymbol{r}) = \exp\left[-i\frac{e}{\hbar}\boldsymbol{A}_\lambda\cdot\boldsymbol{r}\right]\phi_\lambda(\boldsymbol{r}), \tag{8.38}$$

where $A_\lambda = B \times R_\lambda/2$ is the vector potential located at nucleus λ.

To summarize this chapter, for theoretical VOA spectra we need to calculate at least the force field, giving us the harmonic frequencies, and the intensity tensors (APT, AAT, and derivatives of α, G', and A), determining band intensities. In practical computations users usually want to improve the accuracy by using more advanced DFT or wave function methods. At the same time, they need to treat big systems, such as molecules, including environment/solvent. This may be often contradictory, on the other hand in many cases the quantum-chemical methods can be combined or replaced by MD simulations or less rigorous approaches. Some of them are described in part 8.5.

8.3 Small molecules

Rigid "small" molecules are certainly interesting also for themselves. With respect to protein ROA and VCD, they serve as convenient model systems, making possible thorough theoretical investigations. For example, vibrations of *N*-methylacetamide (NMA), a single amide "blocked" with the methyl groups, can be directly related to those in proteins (Fig. 8.1).

Kubelka and Keiderling investigated the dependence of amide I NMA frequency on the solvent and basis set [73]. DFT and modified basis sets were used to model the IR frequencies with a range of explicit and continuum solvent models (Onsager [74] or CPCM reaction field [75]). These calculations also showed that the continuum models cannot replace "explicit" hydrogen-bonded water molecules to the carboxyl group to explain in full the frequency shift when NMA is transferred from vacuum to water environment. However, a combination of three water molecules embedded in a continuum dielectric already provided good agreement with experiment. This concept was further investigated later [76],

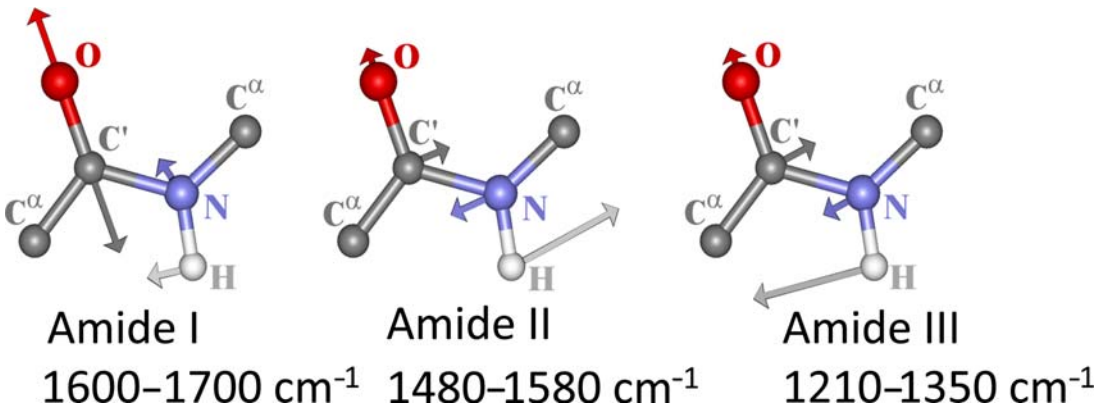


Figure 8.1
Selected vibrational modes of the amide group (amide I, II, and III), which are considered most important for peptide and protein secondary structure analyses.

and DFT results obtained on NMA-water clusters were generalized to give an empirical correction to the force field and transition dipole moment of the amide I mode. Realistic band widths could be then obtained for combined DFT/MD simulations. This method was extended for a general solvated chromophore. For an α-helical peptide, the model significantly improved the simulated spectra and was able to explain many features observed experimentally, including the isotopic deuterium-exchange effects [77].

8.3.1 Flexible molecules, Boltzmann averaging

Except for the solvent, the interpretation of VOA spectra of even small molecules is complicated by their flexibility [78]. In fact, shorter peptides are often more flexible than bigger ones stabilized by many intermolecular noncovalent interactions. Populations of individual conformers are obtained as their Boltzmann weights

$$B_j = \frac{e^{-\Delta E_j/kT}}{\sum_j e^{-\Delta E_j/kT}}, \tag{8.39}$$

where k is the Boltzmann constant, E_j is the energy of the jth conformer, and T is temperature.

For molecules with a limited number of degrees of freedom, the most energetically favored conformers can be identified by the potential energy scan. In a "relaxed" scan the "scanned" coordinates are usually fixed in equidistant increments, while others are optimized by energy minimization. For larger molecules, the number of possible conformations grows rapidly. Then MD simulations may be used to identify the relevant conformations. Typically, MD allows for a better involvement of the solvent, but QM/DFT method provides more reliable conformer energies.

Already early studies showed that the molecular flexibility is very much reflected in VOA spectra. In general, the flexibility makes the chiral signal weaker and individual bands broader. ROA spectra of rigid molecules exhibit sharper bands and more ROA features than those of flexible ones. A good example of this is shown in Fig. 8.2 for rigid Gly-Pro and flexible Pro-Gly dipeptide [79].

8.3.2 Solvent models, clusters

The neglect of the solvent can obviously lead to large errors in the simulations, especially for the polar peptide-like systems [23,80]. The solvent balances partial atomic charges, participates in hydrogen bonding, may cause an aggregation of the solute molecules, etc. Additionally, for some low-frequency modes, it is difficult to separate the solvent and solute spectral signal [81]. The solvent can be "implicitly" added to quantum-chemical computations as a continuous medium, which usually results only in a small increase of

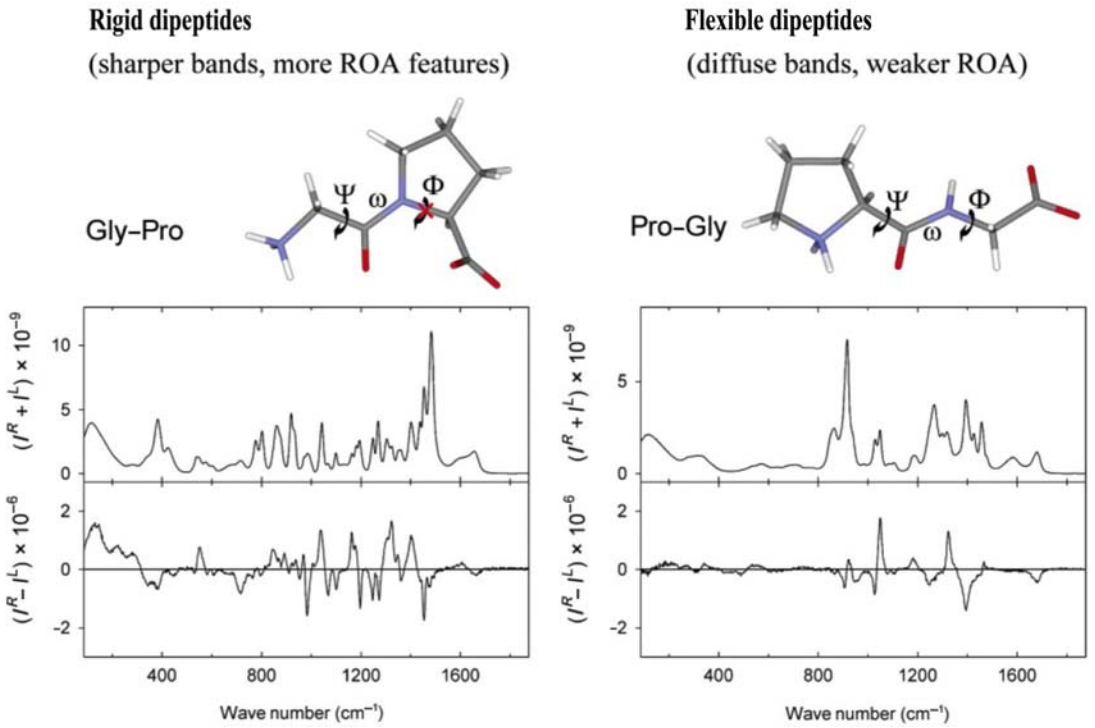


Figure 8.2

Comparison of Raman and ROA spectra of a more rigid Gly-Pro and a flexible Pro-Gly dipeptide. Source: *Reproduced with permission from J. Kapitán, V. Baumruk, V. Kopecký, Jr., P. Bouř, Conformational flexibility of L-alanine zwitterion determines shapes of Raman and Raman optical activity spectral bands, J. Phys. Chem. A 110 (2006) 4689–4696 [79]. Copyright 2006 American Chemical Society.*

computational time. More accurate "explicit" consideration requires direct involvement of solvent molecules in the computations. Then averaging of solvent positions is often needed. The implicit and explicit approach may be combined, when distant solvent molecules are considered as a dielectric continuum [82–84].

As an example of successful implicit solvent model, the conductor-like screening model (COSMO) was originally based on the assumption that the solvent may be represented by a conductor [85]. The polarizable continuum models (PCM) represent the solvent by a polarizable medium characterized by the dielectric constant [86]. A cavity for the solute is defined by a sphere centered on each solute atom. The surface of the cavity is divided into small areas possessing charges induced by the solute. In early implementations this fragmentation caused large instabilities in practical computations; however, these were removed in the later software. In the Gaussian program package [35], the COSMO model is implemented in the PCM framework as "CPCM" [75,87].

Solvent−solute clusters can be obtained ad hoc (educated guess) or generated from MD simulations. A more complicated treatment is possible, for example, with a layered quantum mechanics/molecular mechanics (QM/MM) approach, where the solute molecule is described by QM, explicit solvent molecules create the MM part, and the whole system can be enclosed in a CPCM cavity [88,89]. MD geometries obtained at a certain temperature are far from the energy minima, and a direct simulation of the vibrational spectra based on the harmonic approximation is not practical. Therefore partial geometry optimization must be done. A full optimization is usually avoided, to preserve some geometry dispersion obtained from MD. For example, some torsional angles in a peptide may be fixed.

In the past, we promoted the partial optimization in normal mode coordinates [90,91]. It provides better numerical stability than the optimization in conventional or redundant internal coordinates [92], especially for weakly bonded molecular complexes with noncovalent interactions such as solvent−solute clusters. No selection of coordinates is needed, as this is regulated by definition of the frequency cutoff (typically 200−300 cm^{-1}) under which the normal modes are fixed [93].

MD with ab initio (DFT) potentials is also possible; however this treatment of explicit hydration was found to be computationally very demanding, as shown by Hopmann et al. [94] (Fig. 8.3). In this study, the CPCM model with conformational averaging provided reasonable Raman profiles, but ROA profited more from the inclusion of explicit water molecules. A significant amount of clusters was needed to achieve good accuracy of the ROA spectrum (the ROA/Raman error ratio converged approximately as $\sim 1/\sqrt{N}$). It was concluded that 50 and 500 clusters were needed for a reliable simulation of the Raman and ROA intensities. About 10 times fewer clusters are needed for averaging of solute spectral properties to provide the same accuracy as the plain averaging, if the most probable cluster geometries are preselected, by the parallel variable selection of MD clusters proposed by Kessler et al. [95]. The mass, charge, or atomic density MD distributions are used as a secondary variable.

The treatment of the solvent is particularly important for zwitterionic peptides, because the NH_3^+ and COO^- groups in vacuum after geometry optimization often combine into unphysical structures. In the past, many single amino acids including zwitterionic species were used as convenient models for solvated proteins. Experimental and theoretical attention was paid to ROA studies of alanine [96−98]. Other studies focused on proline, as a model for secondary amine [81]. The secondary structure of a peptide chain is significantly affected by the conformational rigidity of proline [99].

Short oligopeptides represent an important link between properties of individual amino acids and longer peptides. The Ala-Ala peptide is very popular in the simulations. It is well suited as a minimum model of the peptide linkage and the amide group, and it is quite flexible [80,100−104]. Conformer weights can be obtained by a decomposition of

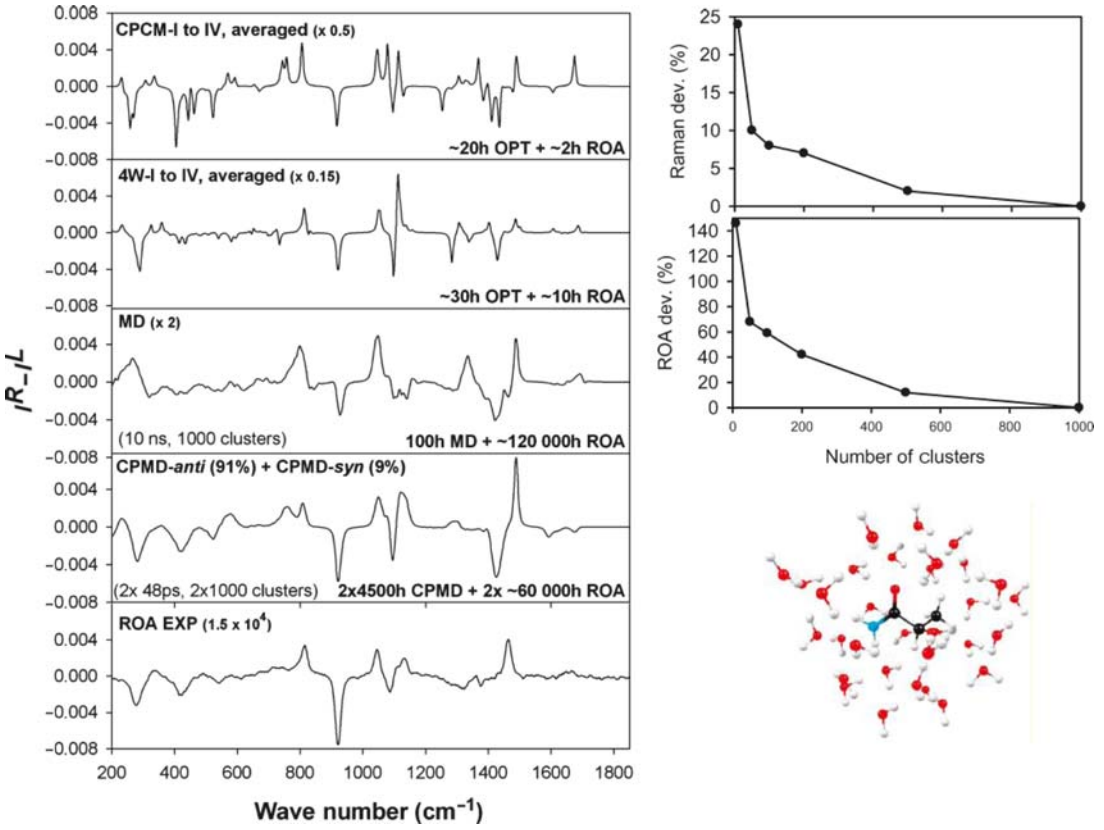


Figure 8.3

Left: ROA spectra of (R)-lactamide computed with different hydration models, and the experiment. Right: Normalized errors of Raman and ROA spectra as dependent on the number of clusters. Source: *Reproduced with permission from K.H. Hopmann, K. Ruud, M. Pecul, A. Kudelski, M. Dračínský, P. Bouř, Explicit versus implicit solvent modeling of Raman optical activity spectra, J. Phys. Chem. B 115 (2011) 4128–4137 [94]. Copyright 2011 American Chemical Society.*

experimental VOA spectra into simulated subspectra of individual conformers [105]. In a later work we investigated the accuracy and mathematical stability of such decomposition [106]. The results were verified in multiple experiments with protonated, zwitterionic, and deprotonated Ala-Ala forms, and natural and d_2- and d_8-isotopically labeled dipeptides. An example for the zwitterionic form is shown in Fig. 8.4. In this study, we supposed that the deuteration leaves the electronic structure nearly unaltered, and its effect on the Ala-Ala conformation is negligible.

8.4 Large molecules

At present, a direct application of DFT to big molecules may not be possible, because of excessive demands on computer time and memory. The largest peptide for which full ab

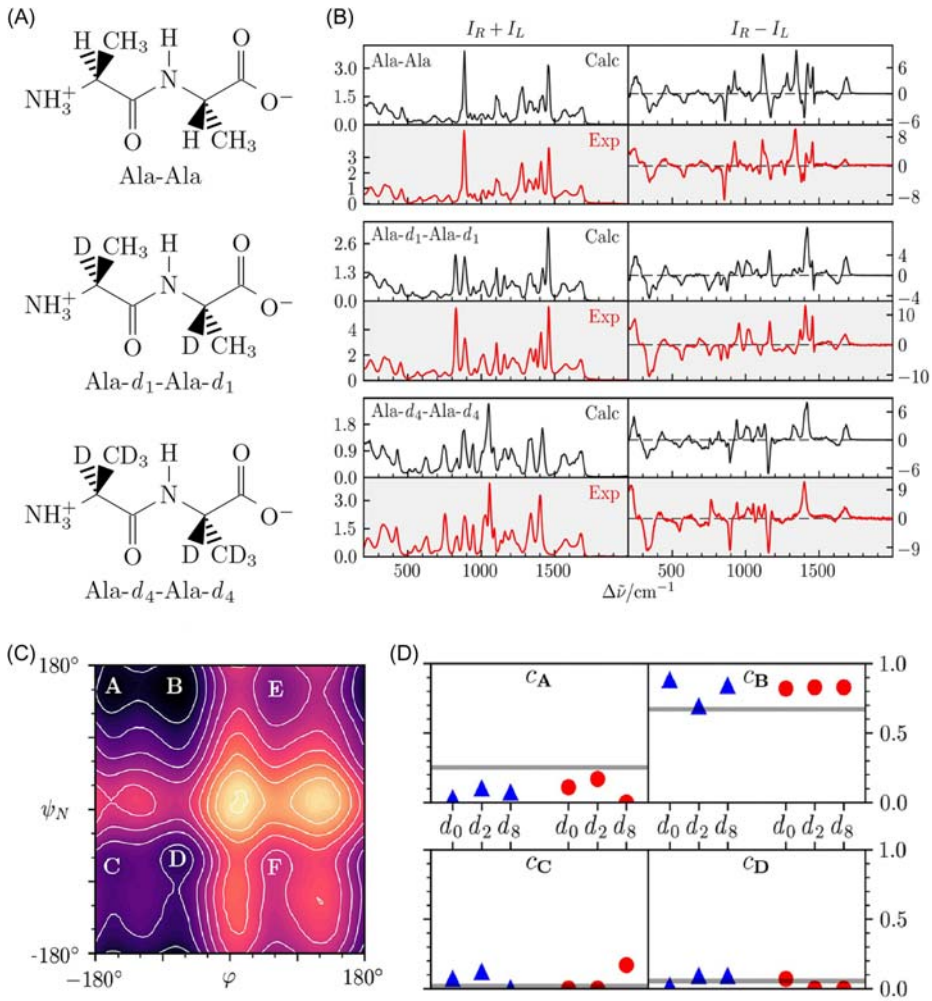


Figure 8.4

Isotopically labeled Ala-Ala dipeptides (zwitterionic form, pH 7), (B) simulated and experimental Raman and ROA spectra, (C) potential energy as a function of φ and ψ_N angles as obtained from MD simulations for the zwitterionic Ala-Ala, (D) conformer populations obtained for three isotopic isomers (d_0, d_2, d_8) of zwitterionic Ala-Ala dipeptide by the WHAM method (*gray lines*) and by the decompositions of Raman (*triangles*) and ROA (*circles*) experimental spectra. Source: *Reproduced with permission from J. Jungwirth, J. Šebestík, M. Šafařík, J. Kapitán, P. Bouř, Quantitative determination of Ala-Ala conformer ratios in solution by decomposition of Raman optical activity spectra, J. Phys. Chem. B 121 (2017) 8956−8964 [106]. Copyright 2017 American Chemical Society.*

initio ROA calculations have been performed, is probably the β domain of rat metallothionein, a cysteine-rich irregular protein with more than 400 atoms (31 amino acids, Fig. 8.5) [107]. The calculation has been possible at the BP86/TZVP level, due to the employment of density-fitting techniques (resolution of identity) and a massive parallelization [108].

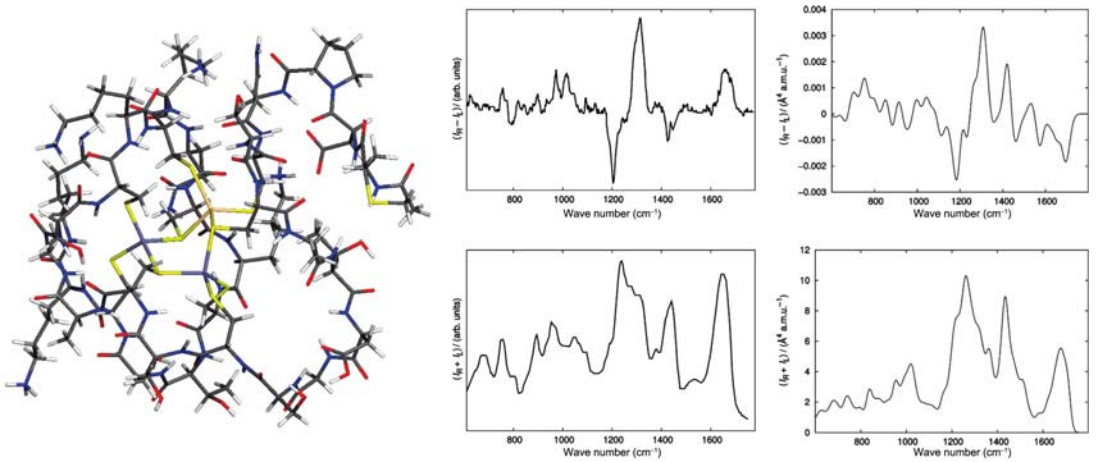


Figure 8.5
Optimized geometry, experimental (middle) and calculated (right, BP86/RI/TZVP) ROA (top) and Raman (bottom) spectra of the β-domain of rat metallothionein. Source: *Reproduced with permission from S. Luber, M. Reiher, Theoretical Raman optical activity study of the β domain of rat metallothionein, J. Phys. Chem. B 114 (2010) 1057–1063 [107]. Copyright 2010 American Chemical Society.*

8.5 Semiempirical approaches

8.5.1 Transition dipole coupling

Semiempirical methods may not be always considered inferior to direct quantum-chemical computation. They often provide a deep insight into the physics and chemistry of studied systems, and depending on the implementation they may be equally or even more useful than "blind" ab initio approaches in interpreting the experimental results.

One of the oldest models used for VOA simulations is the coupled oscillator mechanism (or transition dipole coupling, TDC). Chromophores, for example, the amide groups in a peptide, are replaced by oscillators(transition electric dipoles) interacting with each other and with the circularly polarized light [109–112]. The TDC model was used to simulate oligopeptide IR and VCD spectra by Schellman and later by Diem [113–115]. While some spectral features were predicted correctly (e.g., α-helical amide I signal), the TDC model failed for some other vibrations. Its performance against more advanced computation and the experiment could be estimated for a dipeptide held in various conformations in Fig. 8.6 [116].

8.5.2 Cartesian coordinate tensor transfer

Another computational trick making simulations of VOA of large molecules more efficient is the Cartesian coordinate tensor transfer (CCT) of the tensors needed for VOA

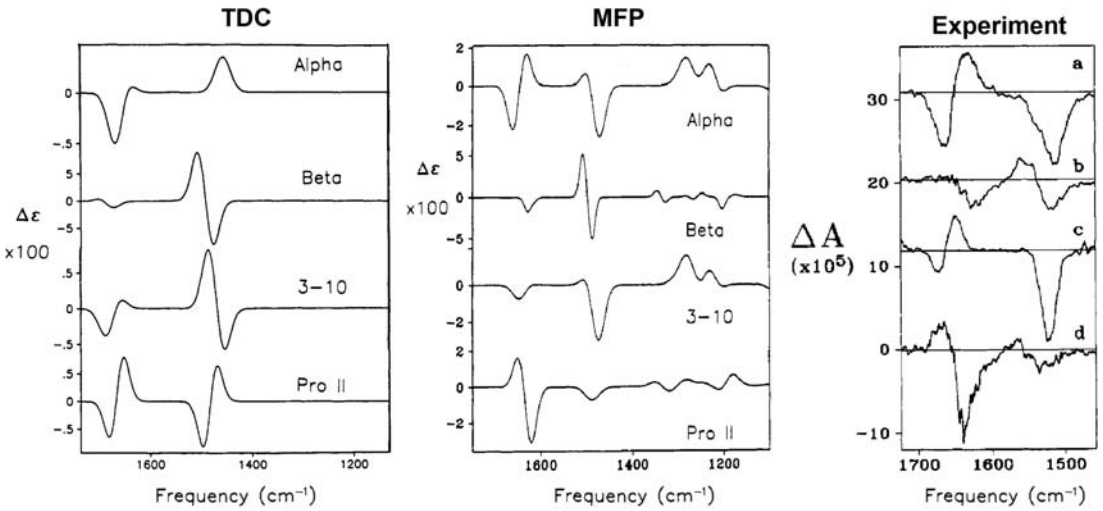


Figure 8.6

VCD spectra from the TDC and a quantum-chemical MFP calculations on a dipeptide, and experimental spectra of larger proteins. Source*: Reproduced with permission from P. Bouř, T.A. Keiderling, Ab initio simulation of the vibrational circular dichroism of coupled peptides, J. Am. Chem. Soc. 115 (1993) 9602–9607 [116]. Copyright 1993 American Chemical Society.*

simulations. The tensors (force field, APT, AAT, and α, G', and A derivatives) are computed by DFT for smaller molecular fragments, and transferred back to the system of interest. This methodology often leads to spectral simulations of nearly ab initio quality, but much lower calculation costs [117–121]. In the latest implementations of CCT, a "big" molecular system, protein, can be automatically divided into smaller fragments (Fig. 8.7).

This approach was applied, for example, in a set of globular proteins, and provided very satisfactory precision in frequencies and spectral intensities if compared with experimental spectra [122–124]. Typically, X-ray protein geometries were divided into overlapping fragments containing four amides, and the fragments were partially optimized using the normal mode optimization technique [90,91]. Then, vibrational spectroscopic parameters (force field, Raman, IR, ROA, and VCD intensity tensors) were calculated at the B3PW91/6-31++G**/CPCM level for each fragment and transferred back to the protein. Technically, the CCT procedure comprises a tensor rotation. The rotation matrix is obtained from the best fit of the fragments if overlapped with the parent molecule. Individual rotation matrices are constructed, in case of the force field (second energy derivatives) for each pair of atoms.

By CCT, many experimental differences between VOA spectra of mainly α-helical human serum albumin (585 amino acids) and concanavalin A (237 amino acids) containing mainly β-sheets were reproduced (Fig. 8.8). In the case of ROA, human serum albumin provided

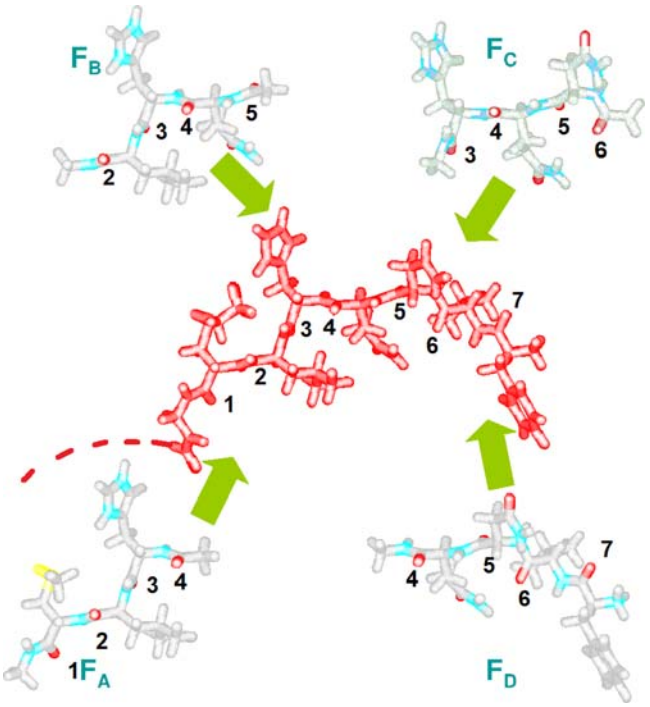


Figure 8.7
The CCT method: The peptide main chain (*red*) is divided into overlapping fragments ($F_A \cdots F_D$), containing four amide residues. Source: *Reproduced from J. Kessler, J. Kapitán, P. Bouř, First-principles predictions of vibrational Raman optical activity of globular proteins, J. Phys. Chem. Lett. 6 (16) (2015) 3314–3319 [122].*

typical α-helical features: positive band at $\sim$1344 cm^{-1} (calculated at 1369 cm^{-1}), a "−/+" couplet at $\sim$1646/1667 cm^{-1} (calculated at 1668/1728 cm^{-1}), and a positive signal within $\sim$870–950 cm^{-1}. Predominantly β-sheets concanavalin A had a negative band at $\sim$1242 cm^{-1} (calculated at 1273 cm^{-1}) and a "−/+" couplet at $\sim$1661 and 1680 cm^{-1} (calculated at 1728 cm^{-1}). For VCD, the α-helical "−/+" amide I signal changed to an approximately opposite "+/−" one for β-sheet; however, the positive part was further split (1732/1715 and 1681/1659 cm^{-1} in simulation and experiment, respectively). The results indicated that the high- and low-frequency amide I vibrations differ in the sensitivity to coupling with distant vibrations.

We were also able to reproduce tiny spectral differences between hen and human lysozymes (Fig. 8.9). However, many questions about the link between protein structure and spectral shapes remain unanswered. In VCD the computations suggest an effect of longer-range interactions, normally omitted by the CCT method. Partially, they can be included as the TDC correction.

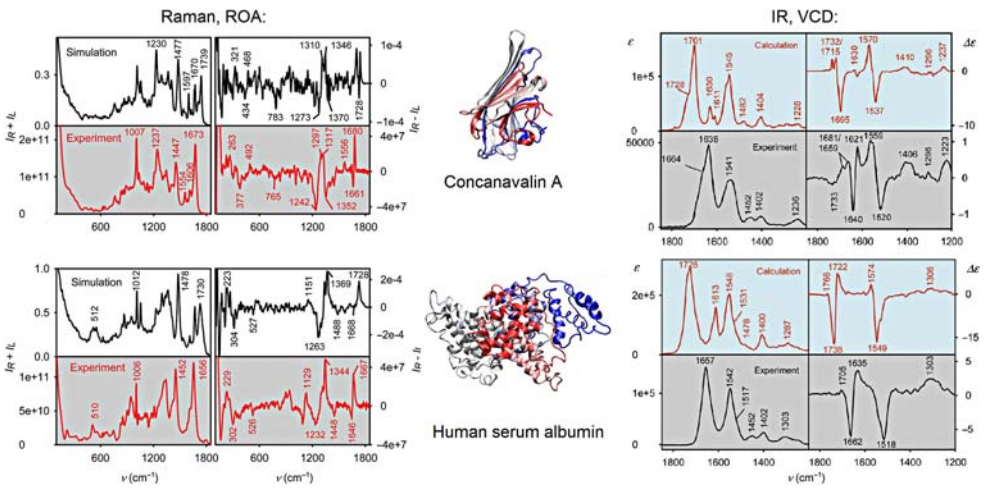


Figure 8.8

Calculated and experimental Raman/ROA (left) and IR/VCD (right) spectra of concanavalin A (top) and human serum albumin (bottom). Source: *Adapted from J. Kessler, J. Kapitán, P. Bouř, First-principles predictions of vibrational Raman optical activity of globular proteins, J. Phys. Chem. Lett. 6 (16) (2015) 3314–3319 [122] and J. Kessler, V. Andrushchenko, J. Kapitán, P. Bouř, Insight into vibrational circular dichroism of proteins by density functional modeling. Phys. Chem. Chem. Phys. 20 (2018) 4926-4935 [124].*

The CCT approach coupled with the TDC model has also been successfully used to model VCD patterns of protein β-sheet structures including fibrils. In Ref. [125], layered β-sheet structures consisting of assembled five-stranded Ala_9 peptides were studied using CCT-based VCD, IR, and Raman spectral simulations, changing the number of layers, interstrand and stacking arrangements. DFT parameters for each sheet were combined with the TDC approach. Force field elements for distant atoms i and j (which were not possible to obtain by DFT or CCT) were calculated by the TDC formalism as

$$f_{i\alpha,j\beta} = \frac{\partial^2 E}{\partial r_{i\alpha}\partial r_{j\beta}} = \frac{1}{4\pi\varepsilon_r\varepsilon_0}\frac{\boldsymbol{P}_{i\alpha}\cdot\boldsymbol{P}_{j\beta}r_{ij}^2 - 3\boldsymbol{P}_{i\alpha}\cdot r_{ij}\boldsymbol{P}_{j\beta}\cdot r_{ij}}{r_{ij}^5}, \tag{8.40}$$

where ε_r and ε_0 are the relative and vacuum dielectric constants, $r_{ij} = r_i - r_j$ is the distance vector between the atoms, $\boldsymbol{P}_{i\alpha} = \partial\mu/\partial r_{i\alpha}$ is the APT. In this case the TDC approximation was not limited to the amide I, but included all vibrations.

The TDC model was also used to generate VCD and ROA spectra of protein fibrils [126]. Optical activity of such aggregates is often significantly stronger than for isolated molecules [127]. Because of the size of the aggregates, theoretical foundations of the chirality

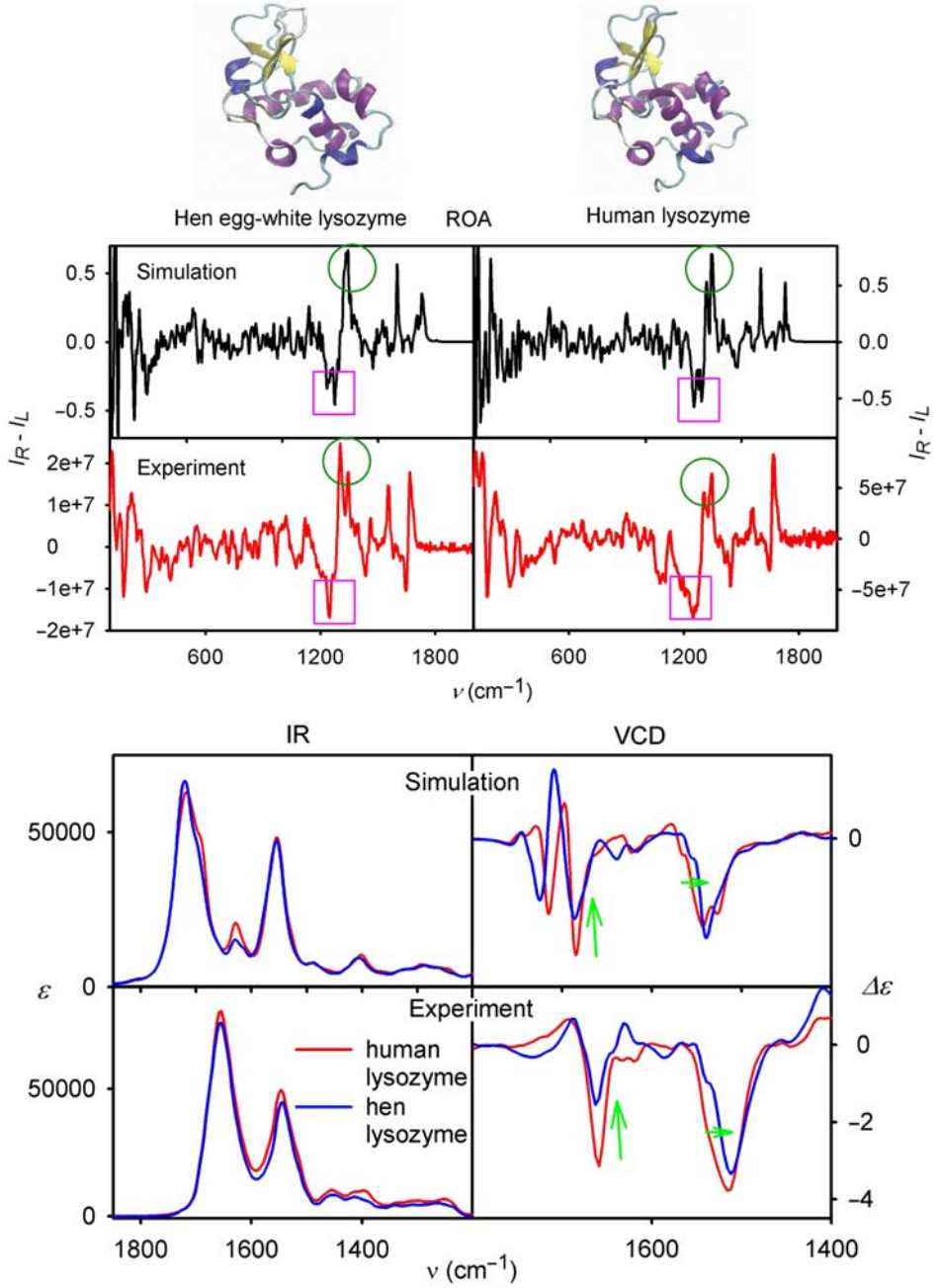


Figure 8.9

Top: Comparison of simulated and experimental ROA spectra of hen egg-white lysozyme (left) and human lysozyme (right). Bottom: Comparison of IR (left) and VCD (right) spectra of the two lysozymes. Some differences in spectral shapes are indicated by the circles, boxes and arrows. Source*: Redrawn from J. Kessler, J. Kapitán, P. Bouř, First-principles predictions of vibrational Raman optical activity of globular proteins, J. Phys. Chem. Lett. 6 (16) (2015) 3314−3319 [122] and J. Kessler, V. Andrushchenko, J. Kapitán, P. Bouř, Insight into vibrational circular dichroism of proteins by density functional modeling. Phys. Chem. Chem. Phys. 20 (2018) 4926-4935 [124].*

enhancement are rather incomplete. Within the TDC model these extensive molecular systems can be treated relatively cheaply in terms of computer time and memory. The planar stacking of the chromophores was identified as the most significant factor responsible for the observable enhancements.

Fibril ROA was modeled for insulin [123]. In solution, insulin probably adopts a relaxed crystal-like geometry [128]. Measurement of ROA spectra of mature insulin fibrils was not possible because of instrumental artifacts. However, spectra of prefibrillar stages provided spectral changes that could be assigned to the formation of the fibrils (Fig. 8.10). They could be attributed to an α-helix−β-sheet secondary structure transition. These changes were observed in the amide I, extended amide III, and the lowest wave number ($\sim$300 cm^{-1}) spectral regions. The CCT methodology was coupled with MD simulations to take into account also the flexibility of the molecule and conformer equilibria in solution, and many snapshots were averaged.

A theoretical model of VCD of dense fibrils was developed in Ref. [129] for molecular assemblies containing repeating β-strands of $(Glu)_n$. We could explain the sign of the major amide I band and the signal of the glutamate side chain, although the modeling did not provide correctly absolute VCD intensities.

8.5.3 Molecules in molecules

Quite a similar idea as for the CCT is exploited in the molecules-in-molecules (MIM) fragment-based method introduced by Mayhall and Raghavachari [130]. MIM is more advanced in that the fragmentation is done at the DFT stage of computation. The MIM is a general, extrapolated ONIOM-like [131,132] methodology for a multilevel fragmentation energy approach (Fig. 8.11). The main advantage of the MIM methodology is the possibility to include all interactions (also long-range interactions neglected in the CCT method) into calculations, albeit, at the low-level of theory. Computational requirements are significantly reduced if compared to a plain DFT computation.

The MIM approach involves four steps: (1) a large molecule is divided into nonoverlapping small fragments. The initial fragments and subsystems can be generated in a systematic manner by cutting single bonds in the large molecule without any assumptions or symmetry restrictions. (2) Overlapping primary subsystems are formed from the local interactions between fragments. (3) A subsystem is made to account for the overcounting from the overlapping regions via the inclusion−exclusion principle. All the dangling bonds in the subsystems are saturated with hydrogen link atoms. Later, the contribution of all of the link atoms in the subsystems are projected back onto the corresponding supporting and host atoms through a Jacobi projection method. (4) The energy of large molecules is evaluated by summation of independent energies of the

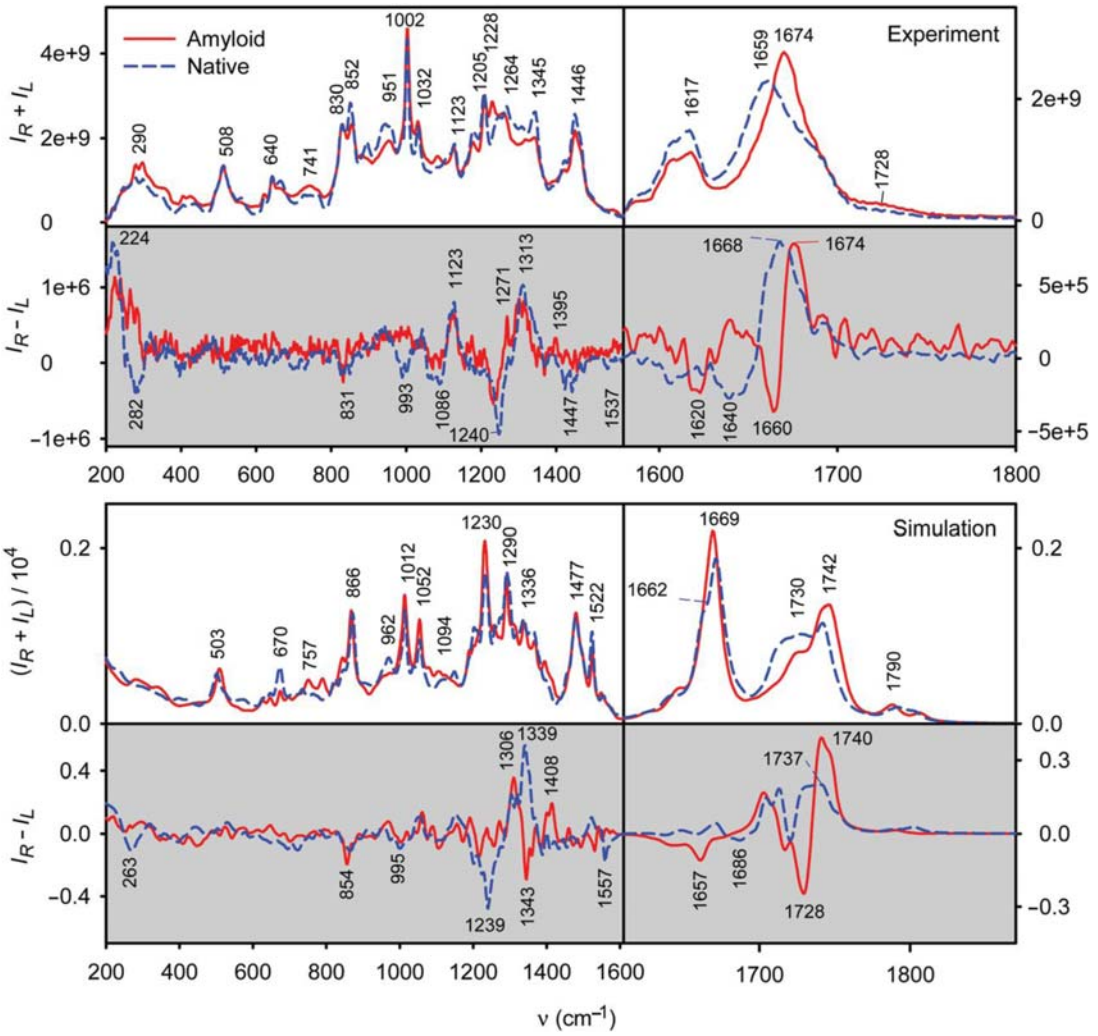


Figure 8.10

Experimental (top) and simulated (bottom) Raman and ROA spectra of native and fibrous insulin; the amide I vibrational region is expanded on the right-hand side. The calculated spectra were obtained as averages from 865 MD snapshot geometries. Source*: Reproduced with permission from J. Kessler, S. Yamamoto, P. Bouř, Establishing the link between fibril formation and Raman optical activity spectra of insulin, Phys. Chem. Chem. Phys. 19 (2017) 13614–13621 [123].*

individual subsystems, taking into account the signs for the energies of the subsystems. For example, for the two-layer MIM2 model, where the missing long-range interfragment interactions are accounted for using a computationally cheaper lower level of theory [134]:

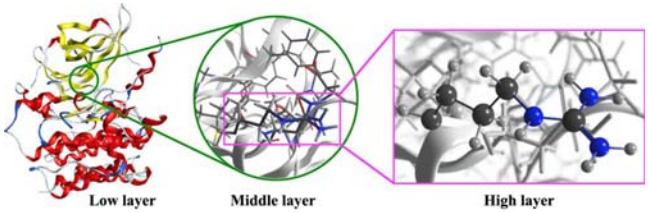


Figure 8.11
Fragmentation of a studied system into subsystems treated at different levels in the three-layer MIM method (MIM3). While subsystems of four to five covalently bonded fragments are computed at high level of theory, a DFT lower level with a substantially smaller basis set is used for dimers and dispersion-corrected semiempirical methods are used for the unfragmented molecule. Source: *Reproduced with permission from B. Thapa, D. Beckett, K.V.J. Jose, K. Raghavachari, Assessment of fragmentation strategies for large proteins using the multilayer molecules-in-molecules approach, J. Chem. Theory Comput. 14 (3) (2018) 1383−1394 [133]. Copyright 2018 American Chemical Society.*

$$E_{\text{Total}} = E_{rl} - \left\{ \sum_i E^i_{ml} \cdots + (-1)^{n-1} \sum_n E^{i \cap j \cap k \cdots \cap n}_{ml} \right\} + \left\{ \sum_i E^i_{mh} \cdots + (-1)^{n-1} \sum_n E^{i \cap j \cap k \cdots \cap n}_{mh} \right\}, \tag{8.41}$$

where E_{rl} is the energy of the real (i.e., full) system calculated at the low level, and E_{ml} and E_{mh} are the energies of the model-low and model-high regions calculated at the low and high levels of theory. In principle, the method can also be extended for multiple low-level correction layers [133].

The relevant higher energy derivatives of the parent molecule needed for calculations of VOA spectra are assembled from the corresponding derivatives of smaller fragment subsystems. Independent fragment calculations are performed in the global coordinate frame; hence there are no errors resulting from the rotation and translation of the tensor elements while constructing the full molecular property tensors.

The MIM method has been recently adapted to evaluate IR [135], Raman [136], and NMR [137] spectra of large molecules. The MIM-VCD and MIM-ROA methods were benchmarked on a set of 10 pairs of carbohydrate enantiomers with strong intramolecular interactions [138,139]. In a later work [134], the MIM approach was used also to explore evolution of the chiroptical spectroscopic characteristics of 3_{10}-helix, α-helix, β-hairpin, γ-turn, and β-extended conformers of gas phase polyalanine (chain length $n = 6-14$). Also isotopic substitutions and continuum solvation effects could be incorporated.

8.6 Conclusions

We were trying to show that the interpretations of VOA spectra based on thorough simulations lead to a deeper understanding of molecular structure, dynamics, and

interactions. While small molecules enable precise calculations, attention still should be paid to molecular flexibility and interactions with the solvent. For large molecular systems, useful approximations are available, such as the MIM, combination of the CCT method with DFT, or the TDC model.

Acknowledgment

The work was supported by the Czech Grant Agency (18-05770 S) and Ministry of Education (LTC17012 and CZ.02.1.01/0.0/0.0/16_019/0000729) and computational grants of CESNET (LM2015042) and the CERIT-SC (LM2015085).

References

[1] J.M. Bijvoet, A.F. Peerdeman, A.J. van Bommel, Determination of the absolute configuration of optically active compounds by means of X-rays, Nature 168 (1951) 271–272.
[2] J. Keeler, Understanding NMR Spectroscopy, second ed., John Wiley & Sons, 2010.
[3] M. Dračínský, H.M. Moller, T.E. Exner, Conformational sampling by ab initio molecular dynamics simulations improves NMR chemical shift predictions, J. Chem. Theory Comput. 9 (8) (2013) 3806–3817.
[4] W. Kühlbrandt, Cryo-EM enters a new era, Elife 3 (2014).
[5] R.A. Harris, C.J. Jameson, A note on chirality in NMR spectroscopy, J. Chem. Phys. 124 (2006) 096101.
[6] S. Yamamoto, H. Watarai, P. Bouř, Monitoring the backbone conformation of valinomycin by Raman optical activity, ChemPhysChem 12 (2011) 1509–1518.
[7] L.D. Barron, Molecular Light Scattering and Optical Activity, second ed., Cambridge University Press, Cambridge, 2009.
[8] P.L. Polavarapu, Renaissance in chiroptical spectroscopic methods for molecular structure determination, Chem. Rec. 7 (2) (2007) 125–126.
[9] L. Pasteur, Théses de Chimie et de Physique, ed. Faculté des sciences de paris. Bachelier, Paris, 1847.
[10] T.M. Lowry, Optical Rotatory Power, Longmans, Green and Co, London, 1935.
[11] C. Djerassi, Optical Rotatory Dispersion: Applications to Organic Chemistry, McGraw-Hill & Company, New York, 1960.
[12] P.L. Polavarapu, Protocols for the analysis of theoretical optical rotations, Chirality 18 (2006) 348–356.
[13] G.D. Fasman, Circular Dichroism and the Conformational Analysis of Biomolecules, Plenum, New York, 1996.
[14] W.C. Johnson, Analyzing protein circular dichroism spectra for accurate secondary structures, Proteins 35 (1999) 307–312.
[15] J.P. Riehl, F.S. Richardson, Circularly polarized luminicence spectroscopy, Chem. Rev. 86 (1985) 1–16.
[16] J.P. Riehl, G. Muller, Circularly polarized luminescence spectroscopy and emission-detected circular dichroism, in: N. Berova, et al. (Eds.), Comprehensive Chiroptical Spectroscopy, Volume 1: Instrumentation, Methodologies, and Theoretical Simulations, John Wiley & Sons, Hoboken NJ, 2012, pp. 65–90.
[17] B.N. Samoilov, Spektry Pogloshcheniya i Lyuminestsentsii Uranilovykh Solei pri Temperature Zhidkogo Geliya, Zh. Eksp. Teor. Fiz. 18 (1948) 1030–1040.
[18] L. Nafie, Vibrational Optical Activity: Principles and Applications, Wiley, Chichester, 2011.
[19] L.D. Barron, M.P. Bogaard, A.D. Buckingham, Raman scattering of circularly polarized light by optically active molecules, J. Am. Chem. Soc. 95 (1973) 603–605.

[20] G. Holzwarth, E.C. Hsu, H.S. Mosher, T.R. Faulkner, A. Moscowitz, Infrared circular dichroism of carbon-hydrogen and carbon-deuterium stretching modes. Observations, J. Am. Chem. Soc. 96 (1) (1974) 251–252.
[21] P.J. Stephens, F.J. Devlin, J.R. Cheeseman, VCD Spectroscopy for Organic Chemists, CRC Press, 2012.
[22] L.D. Barron, A.D. Buckingham, Rayleigh and Raman scattering from optically active molecules, Mol. Phys. 20 (6) (1971) 1111–1119.
[23] P.L. Polavarapu, Molecular structure determination using chiroptical spectroscopy: where we may go wrong? Chirality 24 (2012) 909–920.
[24] V. Profant, M. Pazderková, T. Pazderka, P. Maloň, V. Baumruk, Relative intensity correction of Raman optical activity spectra facilitates extending the spectral region, J. Raman Spectrosc. 45 (7) (2014) 603–609.
[25] A.T. Tu, Peptide backbone conformation and microenvironment of protein side chains, in: R.J.H. Clark, R.E. Hester (Eds.), Spectroscopy of Biological Systems, Wiley, Chichester, 1986, pp. 47–112.
[26] C. Mensch, L.D. Barron, C. Johannessen, Ramachandran mapping of peptide conformation using a large database of computed Raman and Raman optical activity spectra, Phys. Chem. Chem. Phys. 18 (46) (2016) 31757–31768.
[27] N.R. Yaffe, A. Almond, E.W. Blanch, A new route to carbohydrate secondary and tertiary structure using Raman spectroscopy and Raman optical activity, J. Am. Chem. Soc. 132 (31) (2010) 10654–10655.
[28] A. Melcrová, J. Kessler, P. Bouř, J. Kaminský, Simulation of Raman optical activity of multi-component monosaccharide samples, Phys. Chem. Chem. Phys. 18 (2016) 2130–2142.
[29] F. Zhu, N.W. Isaacs, L. Hecht, L.D. Barron, Polypeptide and carbohydrate structure of an intact glycoprotein from Raman optical activity, J. Am. Chem. Soc. 127 (17) (2005) 6142–6143.
[30] A.F. Bell, L. Hecht, L.D. Barron, Vibrational Raman optical activity of DNA and RNA, J. Am. Chem. Soc. 120 (1998) 5820–5821.
[31] A.F. Bell, L. Hecht, L.D. Barron, Vibrational Raman optical activity as a probe of polyribonucleotide solution stereochemistry, J. Am. Chem. Soc. 119 (26) (1997) 6006–6013.
[32] L.D. Barron, The development of biomolecular Raman optical activity spectroscopy, Biomed. Spectrosc. Imaging 4 (2015) 223–253.
[33] F. Zhu, G.E. Tranter, N.W. Isaacs, L. Hecht, L.D. Barron, Delineation of protein structure classes from multivariate analysis of protein Raman optical activity data, J. Mol. Biol. 363 (2006) 19–26.
[34] V. Baumruk, P. Pančoška, T.A. Keiderling, Predictions of secondary structure using statistical analyses of electronic and vibrational circular dichroism and Fourier transform infrared spectra of proteins in H_2O, J. Mol. Biol. 259 (1996) 774–791.
[35] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, et al., Gaussian 16, Wallingford, CT, 2016.
[36] R. Ahlrichs, M. Bar, H.-P. Baron, R. Bauernschmitt, S. Bocker, M. Ehrig, et al., Turbomole, Version 6, Quantum Chemistry Group, University of Karlsruhe, Karlsruhe, 1998.
[37] K. Aidas, C. Angeli, K.L. Bak, V. Bakken, R. Bast, L. Boman, et al., The Dalton quantum chemistry program system, Wiley Interdiscip. Rev. Comput. Mol. Sci. 4 (3) (2014) 269–284.
[38] E.J. Baerends, T. Ziegler, J. Autschbach, D. Bashford, A. Bérces, F.M. Bickelhaupt, et al., ADF2013, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, 2013.
[39] R.D. Amos, CADPAC: The Cambridge Analytic Derivative Package, SERC Laboratory, Daresbury, 1995.
[40] C. Møller, M.S. Plesset, Note on an approximation treatment for many-electron systems, Phys. Rev. 46 (1934) 618–622.
[41] J. Čížek, J. Paldus, Coupled cluster approach, Phys. Scr. 21 (1980) 251–254.
[42] A. Szabo, N.S. Ostlund, Modern Quantum Chemistry. Introduction to Advanced Electronic Structure Theory, Dover, New York, 1989.
[43] F. Jensen, Introduction to Computational Chemistry, second ed., John Wiley & Sons Ltd, 2007.
[44] H. Koch, M.C. Holthausen, Chemist's Guide to Density Functional Theory, second ed., Wiley-VCH, Weinheim, 2015.

[45] R.G. Parr, W. Yang, Density-Functional Theory of Atoms and Molecules, Oxford University Press, New York, 1994.
[46] P. Hohenberg, W. Kohn, Inhomogeneous electron gas, Phys. Rev. B 136 (1964) 864–871.
[47] W. Kohn, L.J. Sham, Self-consistent equations including exchange and correlation effects, Phys. Rev. A 140 (1965) 1133–1138.
[48] A.D. Becke, Density-functional thermochemistry. III. The role of exact exchange, J. Chem. Phys. 98 (1993) 5648–5652.
[49] J.P. Perdew, J.A. Chevary, S.H. Vosko, K.A. Jackson, M.R. Pederson, D.J. Singh, et al., Atoms, molecules, solids, and surfaces: applications of the generalized gradient approximation for exchange and correlation, Phys. Rev. B 46 (1992) 6671–6687.
[50] M.J. Frisch, M. Head-Gordon, J.A. Pople, Direct analytic second derivatives and electric field properties, Chem. Phys. 141 (1990) 189–196.
[51] L.A. Nafie, T.B. Freedman, Vibronic coupling theory of infrared vibrational transitions, J. Chem. Phys. 78 (1983) 7108–7116.
[52] P.J. Stephens, Theory of vibrational circular dichroism, J. Phys. Chem. 89 (1985) 748–752.
[53] R.D. Amos, N.C. Handy, K.J. Jalkanen, P.J. Stephens, Efficient calculation of vibrational magnetic dipole transition moments and rotational strengths, Chem. Phys. Lett. 133 (1987) 21.
[54] P.W. Atkins, L.D. Barron, Rayleigh scattering of polarized photons by molecules, Mol. Phys. 16 (5) (1969) 453–466.
[55] H. Li, L.A. Nafie, Simultaneous acquisition of all four forms of circular polarization Raman optical activity: results for α-pinene and lysozyme, J. Raman Spectrosc. 43 (1) (2012) 89–94.
[56] L.D. Barron, A.D. Buckingham, Vibrational optical activity, Chem. Phys. Lett. 492 (2010) 199–213.
[57] W. Hug, S. Kint, G.F. Bailey, J.R. Schere, Raman circular intensity differential spectroscopy. The spectra of (-)-α-pinene and (+)-α-phenylethylamine, J. Am. Chem. Soc. 97 (5589–5590) (1975).
[58] L. Nafie, Theory of Raman scattering and Raman optical activity: near resonance theory and levels of approximation, Theor. Chem. Acc. 119 (2008) 39–55.
[59] L.A. Nafie, T.B. Freedman, Dual circular polarization Raman optical activity, Chem. Phys. Lett. 154 (1989) 260–266.
[60] G. Placzek, The Rayleigh and the Raman scattering, in: E. Marx (Ed.), Handbuch der Radiologie, Akademische Verlagsgesellschaft, Leipzig, 1934, p. 205.
[61] P.K. Bose, L.D. Barron, P.L. Polavarapu, Ab initio and experimental vibrational Raman optical activity in (+)-(R)-methylthiirane, Chem. Phys. Lett. 155 (4–5) (1989) 423–429.
[62] P.L. Polavarapu, Ab initio vibrational Raman and Raman optical activity spectra, J. Phys. Chem. 94 (1990) 8106–8112.
[63] R.D. Amos, Electric and magnetic properties of CO, HF, HCI, and CH3F, Chem. Phys. Lett. 87 (1982) 23–26.
[64] V. Liegeois, K. Ruud, B. Champagne, An analytical derivative procedure for the calculation of vibrational Raman optical activity spectra, J. Chem. Phys. 127 (2007) 204105.
[65] K. Ruud, J. Thorvaldsen, Theoretical approaches to the calculation of Raman optical activity spectra, Chirality 21 (1E) (2009) E54–E67.
[66] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, et al., Gaussian 09, Gaussian, Inc., Wallingford, CT, 2009.
[67] J.R. Cheeseman. Calculation of molecular chiroptical properties using density functional theory, CD 2007, Groningen, University of Groningen, 2007.
[68] F. London, Théorie Quantique des Courants Interatomiques dans les Combinaisons Aromatiques, J. Phys. Radium 8 (1937) 397–409.
[69] R. Ditchfield, Self-consistent perturbation theory of diamagnetism I. A gauge-invariant LCAO method for N. M. R. chemical shifts, Mol. Phys. 27 (1974) 789–807.
[70] K.L. Bak, P. Jorgensen, T. Helgaker, K. Ruud, Basis set convergence and correlation effects in vibrational circular dichroism calculations using London orbitals, Farad. Discuss. 99 (1994) 121–129.

[71] J.R. Cheeseman, M.J. Frisch, F.J. Devlin, P.J. Stephens, Ab initio calculation of atomic axial tensors and vibrational rotational strengths using density functional theory, Chem. Phys. Lett. 252 (1996) 211–220.
[72] K. Ruud, T. Helgaker, P. Bouř, Gauge-origin independent density-functional theory calculations of vibrational Raman optical activity, J. Phys. Chem. A 106 (32) (2002) 7448–7455.
[73] J. Kubelka, T.A. Keiderling, Ab initio calculation of amide carbonyl stretch vibrational frequencies in solution with modified basis sets. 1. N-Methyl acetamide, J. Phys. Chem. A 105 (48) (2001) 10922–10928.
[74] L. Onsager, Electric moments of molecules in liquids, J. Am. Chem. Soc. 58 (1936) 1486–1493.
[75] M. Cossi, N. Rega, G. Scalmani, V. Barone, Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model, J. Comput. Chem. 24 (6) (2002) 669–681.
[76] P. Bouř, T.A. Keiderling, Empirical modeling of the peptide amide I band IR intensity in water solution, J. Chem. Phys. 119 (2003) 11253–11262.
[77] P. Bouř, D. Michalík, J. Kapitán, Empirical solvent correction for multiple amide group vibrational modes, J. Chem. Phys. 122 (2005) 144501.
[78] B. Mennucci, C. Cappelli, R. Cammi, J. Tomasi, Modeling solvent effects on chiroptical properties, Chirality 23 (9) (2011) 717–729.
[79] J. Kapitán, V. Baumruk, V. Kopecký Jr., P. Bouř, Conformational flexibility of L-alanine zwitterion determines shapes of Raman and Raman optical activity spectral bands, J. Phys. Chem. A 110 (2006) 4689–4696.
[80] K.J. Jalkanen, R.M. Nieminen, M. Knapp-Mohammady, S. Suhai, Vibrational analysis of various isotopomers of L-alanyl-L-alanine in aqueous solution: vibrational absorption, vibrational circular dichroism, Raman, and Raman optical activity spectra, Int. J. Quantum Chem. 92 (2) (2003) 239–259.
[81] J. Kapitán, V. Baumruk, V. Kopecký Jr., R. Pohl, P. Bouř, Proline zwitterion dynamics in solution, glass and crystalline state, J. Am. Chem. Soc. 128 (41) (2006) 13451–13462.
[82] J.R. Cheeseman, M.S. Shaik, P.L.A. Popelier, E.W. Blanch, Calculation of Raman optical activity spectra of methyl-β-d-glucose incorporating a full molecular dynamics simulation of hydration effects, J. Am. Chem. Soc. 133 (13) (2011) 4991–4997.
[83] E. Deplazes, W. van Bronswijk, F. Zhu, L.D. Barron, S. Ma, L.A. Nafie, et al., A combined theoretical and experimental study of the structure and vibrational absorption, vibrational circular dichroism, Raman and Raman optical activity spectra of the L-histidine zwitterion, Theor. Chem. Acc. 119 (2008) 155–176.
[84] S. Luber, Solvent effects in calculated vibrational Raman optical activity spectra of α-helices, J. Phys. Chem. A 117 (2013) 2760–2770.
[85] A. Klamt, G. Schuurmann, COSMO: a new approach to dielectric screening in solvent with explicit expression for the screening energy and its gradient, J. Chem. Soc. Perkin Trans. 2 (1993) 799–805.
[86] J. Tomasi, B. Mennucci, R. Cammi, Quantum mechanical continuum solvation models, Chem. Rev. 105 (2005) 2999–3093.
[87] V. Barone, M. Cossi, Quantum calculations of molecular energies and energy gradients in solution by a conductor solvent model, J. Phys. Chem. A 102 (11) (1998) 1995–2001.
[88] H. Urago, T. Suga, T. Hirata, H. Kodama, M. Unno, Raman optical activity of a cyclic dipeptide analyzed by quantum chemical calculations combined with molecular dynamics simulations, J. Phys. Chem. B 118 (2014) 6767.
[89] S.T. Mutter, F. Zielinski, P.L.A. Popelier, E.W. Blanch, Calculation of Raman optical activity spectra for vibrational analysis, Analyst 140 (9) (2015) 2944–2956.
[90] P. Bouř, T.A. Keiderling, Partial optimization of molecular geometry in normal coordinates and use as a tool for simulation of vibrational spectra, J. Chem. Phys. 117 (2002) 4126–4132.
[91] P. Bouř, Convergence properties of the normal mode optimization and its combination with molecular geometry constraints, Collect. Czech. Chem. Commun. 70 (2005) 1315–1340.
[92] P. Pulay, G. Fogarasi, Geometry optimization in redundant internal coordinates, J. Chem. Phys. 96 (1992) 2856–2860.
[93] J. Hudecová, K.H. Hopmann, P. Bouř, Correction of vibrational broadening in molecular dynamics clusters with the normal mode optimization method, J. Phys. Chem. B 116 (2012) 336–342.

[94] K.H. Hopmann, K. Ruud, M. Pecul, A. Kudelski, M. Dračínský, P. Bouř, Explicit versus implicit solvent modeling of Raman optical activity spectra, J. Phys. Chem. B 115 (2011) 4128–4137.
[95] J. Kessler, M. Dračínský, P. Bouř, Parallel variable selection of molecular dynamics clusters as a tool for calculation of spectroscopic properties, J. Comput. Chem. 34 (2013) 366–371.
[96] L.D. Barron, A.R. Gargaro, L. Hecht, P.L. Polavarapu, Experimental and ab initio theoretical vibrational Raman optical activity of alanine, Spectrochim. Acta. A 47 (8) (1991) 1001–1016.
[97] E. Tajkhorshid, K.J. Jalkanen, S. Suhai, Structure and vibrational spectra of the zwitterion L-alanine in the presence of explicit water molecules: a density functional analysis, J. Phys. Chem. B 102 (1998) 5899–5913.
[98] K.J. Jalkanen, R.M. Nieminen, K. Frimand, J. Bohr, H. Bohr, R.C. Wade, et al., A comparison of aqueous solvent models used in the calculation of the Raman and ROA spectra of L-alanine, Chem. Phys. 265 (2) (2001) 125–151.
[99] V. Profant, V. Baumruk, X. Li, M. Šafařík, P. Bouř, Tracking of the polyproline folding by density functional computations and Raman optical activity spectra, J. Phys. Chem. B 115 (50) (2011) 15079–15089.
[100] M.R. Oboodi, C. Alva, M. Diem, Solution phase Raman studies of alanyl dipeptides and various isotopomers. A reevaluation of accepted amide vibrational frequencies, J. Phys. Chem. 88 (3) (1984) 501–505.
[101] J. Šebek, B. Gyurscik, J. Šebestík, Z. Kejík, L. Bernárová, P. Bouř, Interpretation of synchrotron radiation circular dichroism spectra of anionic, cationic, and zwitterionic dialanine forms, J. Phys. Chem. A 111 (2007) 2750–2760.
[102] P. Bouř, J. Kapitán, V. Baumruk, Simulation of the Raman optical activity of L-alanyl-L-alanine, J. Phys. Chem. A 105 (2001) 6362–6368.
[103] V. Sychrovský, M. Buděšínský, L. Benda, V. Špirko, Z. Vokáčová, J. Šebestík, et al., Dependence of the L-alanyl-L-alanine conformation on molecular charge determined from ab initio computations and NMR spectra, J. Phys. Chem. B 112 (2008) 1796–1805.
[104] J. Šebek, J. Kapitán, J. Šebestík, V. Baumruk, P. Bouř, L-Alanyl-L-alanine conformational changes induced by pH as monitored by the Raman optical activity spectra, J. Phys. Chem. A 113 (2009) 7760–7768.
[105] V. Parchaňský, J. Kapitán, J. Kaminský, J. Šebestík, P. Bouř, Ramachandran plot for alanine dipeptide as determined from Raman optical activity, J. Phys. Chem. Lett. 4 (2013) 2763–2768.
[106] J. Jungwirth, J. Šebestík, M. Šafařík, J. Kapitán, P. Bouř, Quantitative determination of Ala-Ala conformer ratios in solution by decomposition of Raman optical activity spectra, J. Phys. Chem. B 121 (2017) 8956–8964.
[107] S. Luber, M. Reiher, Theoretical Raman optical activity study of the β domain of rat metallothionein, J. Phys. Chem. B 114 (2010) 1057–1063.
[108] S. Luber, M. Reiher, Raman optical activity spectra of chiral transition metal complexes, Chem. Phys. 346 (2008) 212–223.
[109] G. Holzwarth, I. Chabay, Optical activity of vibrational transitions: a coupled oscillator model, J. Chem. Phys. 57 (4) (1972) 1632–1635.
[110] Y.N. Chirgadze, N.A. Nevskaya, Infrared spectra and resonance interaction of amide I vibration of the antiparallel-chain pleated sheet, Biopolymers 15 (1976) 607–625.
[111] Y.N. Chirgadze, N.A. Nevskaya, Infrared spectra and resonance interaction of amide-I vibration of the parallel-chain pleated sheet, Biopolymers 15 (1976) 627–636.
[112] N.A. Nevskaya, Y.N. Chirgadze, Infrared spectra and resonance interactions of amide-I and II vibrations of a-helix, Biopolymers 15 (1976) 637–648.
[113] J. Snir, R.A. Frankel, J.A. Schellman, Optical activity of polypeptides in the infrared. predicted CD of the amide I and amide II bands, Biopolymers 14 (1975) 173–196.
[114] S.S. Birke, I. Agbaje, M. Diem, Experimental and computational infrared CD studies of prototypical peptide conformations, Biochemistry 31 (1992) 450.

[115] H.R. Wyssbrod, M. Diem, Infrared vibrational CD of peptide beta-turns: a theoretical and experimental study of cyclo(Gly-Pro-Gly-D-Ala-Pro), Biopolymers 31 (1992) 1237.
[116] P. Bouř, T.A. Keiderling, Ab initio simulation of the vibrational circular dichroism of coupled peptides, J. Am. Chem. Soc. 115 (1993) 9602−9607.
[117] P. Bouř, J. Sopková, L. Bednárová, P. Maloň, T.A. Keiderling, Transfer of molecular property tensors in Cartesian coordinates: a new algorithm for simulation of vibrational spectra, J. Comput. Chem. 18 (1997) 646−659.
[118] P. Bouř, J. Kubelka, T.A. Keiderling, Quantum mechanical models of peptide helices and their vibrational spectra, Biopolymers 65 (2002) 45−69.
[119] J. Hilario, J. Kubelka, T.A. Keiderling, Optical spectroscopic investigations of model beta-sheet hairpins in aqueous solution, J. Am. Chem. Soc. 125 (25) (2003) 7562−7574.
[120] S. Yamamoto, X. Li, K. Ruud, P. Bouř, Transferability of various molecular property tensors in vibrational spectroscopy, J. Chem. Theory Comput. 8 (3) (2012) 977−985.
[121] P. Bouř, CCT, Cartesian Tensor Transfer, Academy of Sciences, Prague, 1997-2009.
[122] J. Kessler, J. Kapitán, P. Bouř, First-principles predictions of vibrational Raman optical activity of globular proteins, J. Phys. Chem. Lett. 6 (16) (2015) 3314−3319.
[123] J. Kessler, S. Yamamoto, P. Bouř, Establishing the link between fibril formation and Raman optical activity spectra of insulin, Phys. Chem. Chem. Phys. 19 (2017) 13614−13621.
[124] J. Kessler, V. Andrushchenko, J. Kapitán, P. Bouř, Insight into vibrational circular dichroism of proteins by density functional modeling, Phys. Chem. Chem. Phys. 20 (2018) 4926−4935.
[125] W.R.W. Welch, J. Kubelka, T.A. Keiderling, Infrared, vibrational circular dichroism, and Raman spectral simulations for β-sheet structures with various isotopic labels, interstrand, and stacking arrangements using density functional theory, J. Phys. Chem. B 117 (36) (2013) 10343−10358.
[126] J. Průša, P. Bouř, Transition dipole coupling modeling of optical activity enhancements in macromolecular protein systems, Chirality 30 (2017) 55−64.
[127] T. Measey, R. Schweitzer-Stenner, Vibrational circular dichroism as a probe of fibrillogenesis: the origin of the anomalous intensity enhancement of amyloid-like fibrils, J. Am. Chem. Soc. 133 (4) (2011) 1066−1076.
[128] S. Yamamoto, H. Watarai, Raman optical activity study on insulin amyloid and prefibril intermediate, Chirality 24 (2) (2012) 97−103.
[129] J. Kessler, T.A. Keiderling, P. Bouř, Arrangement of fibril side chains studied by molecular dynamics and simulated infrared and vibrational circular dichroism spectra, J. Phys. Chem. B 118 (2014) 6937−6945.
[130] N.J. Mayhall, K. Raghavachari, Molecules-in-molecules: an extrapolated fragment-based approach for accurate calculations on large molecules and materials, J. Chem. Theor. Comput. 7 (2011) 1336−1343.
[131] S. Dapprich, I. Komáromi, K.S. Byun, K. Morokuma, M.J. Frisch, A. New ONIOM, Implementation in Gaussian 98. 1. The calculation of energies, gradients and vibrational frequencies and electric field derivatives, J. Mol. Struct. 462 (1999) 1−21.
[132] T. Vreven, K. Mokoruma, in: B. Mennucci, R. Cammi (Eds.), Continuum Solvation Models in Chemical Physics: From Theory to Applications, Wiley, 2008.
[133] B. Thapa, D. Beckett, K.V.J. Jose, K. Raghavachari, Assessment of fragmentation strategies for large proteins using the multilayer molecules-in-molecules approach, J. Chem. Theory Comput. 14 (3) (2018) 1383−1394.
[134] K.V.J. Jose, K. Raghavachari, Molecules-in-molecules fragment-based method for the calculation of chiroptical spectra of large molecules: vibrational circular dichroism and Raman optical activity spectra of alanine polypeptides, Chirality 28 (12) (2016) 755−768.
[135] K.V.J. Jose, K. Raghavachari, Evaluation of energy gradients and infrared vibrational spectra through molecules-inmolecules fragment-based approach, J. Chem. Theory Comput. 11 (2015) 950−961.
[136] K.V.J. Jose, K. Raghavachari, Molecules-in-molecules fragment-based method for the evaluation of Raman spectra of large molecules, Mol. Phys. 113 (2015) 3057−3066.

[137] K.V.J. Jose, K. Raghavachari, Fragment-based approach for the evaluation of NMR chemical shifts for large biomolecules incorporating the effects of the solvent environment, J. Chem. Theor. Comput. 13 (3) (2017) 1147−1158.
[138] K.V.J. Jose, D. Beckett, K. Raghavachari, Vibrational circular dichroism spectra for large molecules through molecules-in-molecules fragment-based approach, J. Chem. Theory Comput. 11 (9) (2015) 4238−4247.
[139] K.V.J. Jose, K. Raghavachari, Raman optical activity spectra for large molecules through molecules-in-molecules fragment-based approach, J. Chem. Theory Comput. 12 (2) (2016) 585−594.

CHAPTER 9

Nanoscale analysis of protein self-assemblies

Ewelina Lipiec

Department of Physics of Nanostructures and Nanotechnology, Institute of Physics, Jagiellonian University, Krakow, Poland

9.1 Introduction

The world's aging population is living currently much longer and therefore the number of cases of debilitating and deadly neurodegenerative diseases, such as Parkinson's, Huntington's, and Alzheimer's increases continuously [1–6]. Over 50 million people worldwide live with dementia and the cost of treatment already has exceeded 1% of the global gross domestic product (GDP) [7,8]. These diseases not only affect individuals and their families but they also have significant influence on healthcare systems [6]. Despite the best scientific efforts, unfortunately neurodegenerative diseases are still incurable. The neurodegenerative processes related to these diseases are caused by structurally abnormal peptides called amyloids, characterized by very high cytotoxicity [1–3]. The structural disordering and abnormal aggregation are associated with the cytotoxic properties of specific proteins. It was proven that prefibrillar intermediates, called oligomers, are more toxic than mature fibrils [9,10]. The high cytotoxicity of oligomers may be related to their small size and round shape. The cellular membrane is much more accessible to smaller features such as those oligomers in comparison to the longer fibrils [11]. A barrier to the development of successful treatment strategies is related to the lack of understanding of the aggregation pathways [12,13]. An understanding of how native, water-soluble peptides are able to form very stable, cytotoxic amyloid fibrils, that are implicated in these neurodegenerative diseases is crucial for treatment strategies [12,13]. Abnormal peptide aggregation is strictly related to conformational transition, the β-sheet content increases with the ongoing fibrillation process [14]. Therefore all experimental approaches allowing for verification of protein secondary structure are very significant for research into abnormal protein aggregation. Various analytical techniques have been used in studies of the conformation of neurodegenerative protein, including crystallography [15], circular dichroism [16], Fourier-transform infrared spectroscopy (FTIR) [17], and nuclear magnetic

Vibrational Spectroscopy in Protein Research.
DOI: https://doi.org/10.1016/B978-0-12-818610-7.00009-8

resonance (NMR) [18,19]. All the methods listed above require significant sample volumes in order to achieve a sufficient signal-to-noise ratio. Therefore they all deliver only information about the protein secondary structure that is averaged from many aggregates. Samples of aggregating proteins are very heterogeneous (Fig. 9.1) because the aggregation and related structural transition from α-helices, turns, and unstructured coils to β-sheet conformation is a long-term process, which does not occur simultaneously in all molecules. In Fig. 9.1 this heterogeneity can easily be seen. On the atomic force microscopy (AFM) topography of $A\beta_{1\text{-}42}$ aggregating for 36 hours before being deposited on gold-coated mica the reader can distinguish at least three forms of amyloids: globular oligomers (marked by blue arrows), granular protofibrils (marked by green profiles 4, 5), and smooth mature fibrils (marked by cross sections 1−3). Additionally, the sample presented in Fig. 9.1 also

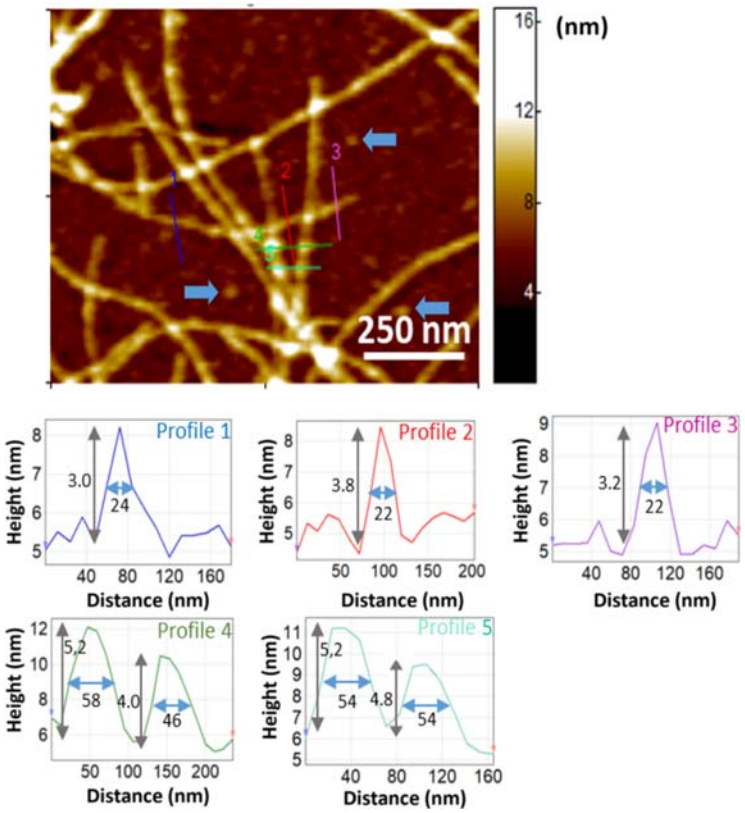


Figure 9.1

AFM topography image of aggregating $A\beta_{1-42}$ incubated at 37°C for 36 h and then deposited on gold substrate, with profiles extracted crosswise mature fibril (profiles 1−3) and segments of protofibrils. Source: *Reproduced with permission from E. Lipiec, D. Perez-Guaita, J. Kaderli, B.R. Wood, R. Zenobi, Direct nanospectroscopic verification of the amyloid aggregation pathway, Angew. Chem. Int. Ed. 57 (28) (2018) 8519−8524. © Wiley-VCH Verlag GmbH & Co. KGaA, 2018.*

contains small aggregates such as dimers, tetramers, and not yet aggregated monomers, which could be too small to be observed by AFM. Based on precise AFM measurements of the thickness of oligomers, fibrils, and granular segments of protofibrils (as presented in Fig. 9.1), these amyloids forms could be classified.

Conventional analytical techniques, do not allow the investigation of each amyloid form separately. Despite this limitation, current knowledge about aggregation processes is very broad. For example for amyloid β, related to Alzheimer's disease, several hypothetic aggregation pathways were proposed (Fig. 9.2) based on experimental data [20–22] and theoretical calculations [23]. Structural data is consistent just for homogenous samples at the beginning of aggregation (because samples do not yet contain aggregated peptide) and at the final stage of aggregation, after a long incubation period (because all peptide is already aggregated).

The possible pathways of primary aggregation are summarized and schematically presented in Fig. 9.2. The aggregation always starts from a merging of monomers with predominantly unstructured coils and turns of secondary structures. Aggregating monomeric precursors may form metastable oligomers with disordered β-sheets, which could slowly rearrange β-sheets into a parallel ordered structure [23,24]. This structural rearrangement may occur

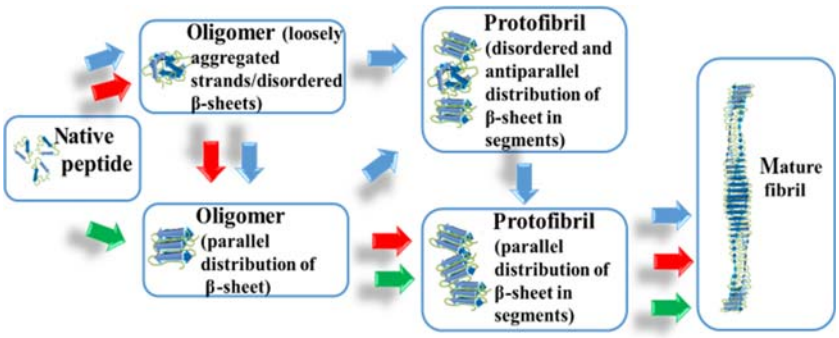


Figure 9.2

Hypothetical representation of three possible pathways of $A\beta_{1-42}$ aggregation: Blue aggregation pathway: native peptide → merging → oligomer (loosely aggregated strands/disordered β-sheets) → partial structural reorganization → oligomer (parallel distribution of β-sheet) → merging → protofibril (disordered and antiparallel distribution of β-sheet in segments) → partial structural reorganization → protofibril (parallel distribution of β-sheet in segments) → ordering of a protofibril core → protofibril → mature fibril. Red aggregation pathway: native protein → merging → oligomer (disordered distribution of β-sheet) → structural reorganization → oligomer (parallel distribution of β-sheet) → merging → protofibril (parallel distribution of β-sheet in segments) → ordering of a protofibril core → protofibril → mature fibril. Green aggregation pathway: native protein → one-by-one monomeric conversion → oligomer (parallel distribution of β-sheet) → merging → protofibril (parallel distribution of β-sheet in segments) → ordering of a protofibril core → protofibril → mature fibril.

at the oligomeric stage (according to the blue pathway, Fig. 9.2) or at the protofibril stage (according to the red pathway, Fig. 9.2). The one-by-one monomeric conversion to the already ordered β-sheet structure is also possible [23,24] (the green pathway, Fig. 9.2). In this pathway, the orientation of β-sheet is ordered in a parallel manner and oriented perpendicular to the fibril axis like in protofilaments and mature fibrils. The most common type of cross-β structure of mature $A\beta_{(1-42)}$ fibrils is parallel β-sheets that are in the registry [20,23–25]. The primary aggregation described above is followed by secondary nucleation, which leads to an increase in fibril diameter [26].

In summary, an application of technologies that provide information about chemical composition and molecular conformation at the nanoscale is required to verify each step of the abnormal protein aggregation and develop methods to control and stop it. Recent advances in applications of molecular nanospectroscopic techniques to following self-assemblies formation will be discussed in this chapter.

9.2 Analysis of microscopic steps of abnormal protein aggregation at the nanoscale—a comparison of various experimental approaches

To follow conformational transition in single peptide aggregates, nondestructive analytical techniques characterized by chemical selectivity, high sensitivity, and nanometric spatial resolution are required. All methods that could possibly meet these criteria are summarized and briefly described in Table 9.1.

9.2.1 Thioflavin-T–based kinetic measurements

Thioflavin-T (ThT)–based kinetic measurements are commonly applied for analysis of individual microscopic steps of amyloids aggregation. Recent advances to data analysis and its interpretation deliver highly robust data [27–30]. This methodology allows the following of the kinetics of protein aggregation based on the dramatic enhancement of fluorescence from the dye ThT, when bound to fibrils. However, fully quantitative data cannot be obtained due to the lack of molecular details of the binding of ThT to amyloid fibrils. Another serious limitation is the high sensitivity to the presence of impurities or amorphous protein aggregates, which cause perturbations in fluorescence intensity [31].

9.2.2 Scanning probe microscopy

An application of label-free nanoscale scanning probe microscopy (SPM): AFM [23] and scanning tunneling microscopy (STM) [32] can follow each single step of the amyloid aggregation, including the formation of oligomers, protofibrils, and fibrils. Recently high-speed AFM measurements have brought new insights into our understanding of the

Table 9.1: Nondestructive analytical methods to probe molecular structure at the nanoscale.

			Method	Achieved information	Resolution (nm)	Advantages	Disadvantages/limitations
Nanospectroscopy and nanoimaging	Chemical marking		Superres. Imaging (dSTORM, BALM)	Subdiffractional localization of particular functional groups previously chemically marked	20–50	High chemical, selectivity, 3D analysis	Chemical marking required, limited information about chemical composition
	No chemical marking	Imaging	NSOM/ SNOM	Subdiffractional probing of optical properties of samples (reflectivity transparency, scattering and fluorescence if stained)	15–100	No or very little out-of-focus light contributions	Limited information about chemical composition and conformation, short working distance, long scan times
			X-ray tomography	Distribution of X-rays absorption–sample density	<1000	3D analysis	Limited information about chemical composition and conformation
		Molecular analysis	TERS	Molecular structure: conformation and composition	Sub-nm–20	High sensitivity (single molecules!)	Unstable enhancement, exceptionally high Raman scattering cross section for carbonaceous contamination, high risk of sample thermal decomposition
			Nano-FTIR	Molecular structure: conformation and composition	30	High sensitivity	Unstable enhancement, long scan times
			AFM-IR	Molecular structure: conformation and composition	10–100	Highly reproducible, robustness	Relatively low sensitivity (in comparison with nano-FTIR and TERS), long scan times
			SIMS	Molecular structure: conformation and composition	50	Highly reproducible, well established	Invasive, only in UHV
			Nano NMR, nano-MRI	Optical measurements of nitrogen-vacancy color centers	< 100	Specific chemical information	Not applicable yet for complex biological samples
		Elemental analysis	SEM, TEM Cryo-EM	Detailed information about the chemical structure, the distribution of single atoms	Sub-nm	Cryo-EM detailed chemical structure	Usually in UHV (except environmental SEM). no molecular information, Cryo-EM requires homogenous samples
			X-ray fluorescence	Elemental composition	<100	Simultaneous detection of a plethora of elements	Limited information about molecular structure and conformation
			Nano XANES	Chemical environment and valence state of selected element	< 100	High chemical selectivity	Limited information about molecular structure and conformation

aggregation mechanisms. This advanced technique allows for the monitoring of the structural dynamics of amyloid $\beta_{(1-42)}$ aggregates and the observation of two distinct "straight" and "spiral" morphemes. Additionally switching between these two structures was also visualized [33]. While SPM including AFM and STM can provide high spatially resolved images of surface topology and visualize each form of amyloids separately, they are chemically blind and provide no information about molecular structure and composition of investigated samples.

9.2.3 Superresolving fluorescence microscopy

Recently of great importance in amyloid studies is single molecule localization used in superresolving fluorescence microscopic techniques. Several various approaches were already successfully applied. Direct stochastic optical reconstruction microscopy (dSTORM) uses the stochastic nature of emissions from isolated fluorophore molecules localized with nanometric accuracy [34,35]. DSTORM has been used to obtain high-resolution morphology of different types of amyloids such as amyloid-β [36,37], α-synuclein [38–40], and huntingtin protein [41–43]. A spatial resolution DSTORM reaches 20 nm and it could be applied in vivo as well as in fixed cells. Binding-activated localization microscopy (BALM) provides even lower nanometer accuracy. In BALM, molecules of interest become highly fluorescent upon interaction with the applied dye and therefore they can be precisely localized. BALM was applied in buffer solution by activating α-synuclein amyloid fibrils with a low concentration of fluorescent target-binding probes [44]. Superresolution microscopy relies on either immunolabeling of the amyloid fibrils or the covalent binding of monomeric proteins by fluorescent organic dyes. Despite the growing importance of superresolution imaging techniques in the amyloid field, none of those approaches could be appropriate for structural analysis/imaging [45–47].

9.2.4 Cryoelectron microscopy

NMR [48] and cryoelectron microscopy [49] were recently used to measure experimentally an atomic resolution structure of $A\beta_{(1-42)}$ [48,49]. Both methods provide the detailed structure of analytes at subnanometric scale. However, both methods require homogeneous samples because the signal in both methods is averaged from thousands of amyloids [49].

9.2.5 Nanoscale nuclear magnetic resonance

Highly localized chemical information could be provided by nanoscale NMR spectroscopy and imaging (nano-NMR, nano-MRI). This methodology applies optical measurements of nitrogen-vacancy color centers. Nano-NMR/nano-MRI could possibly be applied in studies of individual amyloid forms due to a combination of magnetic field sensitivity/chemical

selectivity and nanoscale spatial resolution. However nano-NMR and nano-MRI techniques until now were used for very simple samples like patterns of 19 F [50,51]. Therefore it seems that this technique is still very far from being applied as an analytical method for biologically relevant and complex samples.

9.2.6 X-ray spectroscopy

Due to high X-ray spatial confinement related to their short wavelength several analytical techniques employing properties of X-rays, such as diffraction (nano-XRD), emission (nano-X-ray fluorescence), and absorption (nano-XANES, nano-EXAFS), can probe investigated samples at the nanoscale. However, these techniques deliver very limited information about molecular conformation, which is crucial for investigation into amyloid secondary structures. Nanofocused X-ray diffraction is not applicable to sensitive biological objects, due to the invasive nature of the focused X-ray beam. Until now only inorganic samples such as GaAs-GaInP core−shell nanowires [52] have been successfully measured by nano-X-ray diffraction.

9.2.6.1 Infrared nanospectroscopy

A combination of the AFM and infrared (IR) spectroscopy, gives us the possibility to obtain information about physical and chemical properties of the sample at nanometric spatial resolution [53−59]. This can be implemented by the application of two different techniques, which employees (1) scattering of infrared light at a metal/metalized scanning probe tip [53,57] (nano-FTIR), or (2) an application of an AFM cantilever to detect sample thermal expansion upon absorption of IR light [54−56,59] (AFM-IR).

9.2.7 Nano-Fourier-transform infrared spectroscopy

In the first approach, the AFM probe is typically metallized, therefore it may act as a nanoantenna, converting a freely propagating incident field into localized energy and vice versa at the probe apex for local probing of molecular vibrations [53,57,58]. A tunable quantum cascade laser (QCL) or a broadband mid-infrared laser beam are used for sample illumination. The spectral signal is based on interferometric detection of the scattered infrared light. Fourier-transform spectroscopy of the scattered light yields nanometric infrared spectra. For scattering-type scanning near-field optical microscopy (s-SNOM) imaging the tip-scattered light is usually detected by operating the interferometer in pseudoheterodyne mode [60]; infrared amplitude and phase images are obtained by demodulating the detector signal at a higher harmonic of the tip oscillation frequency.

To the best of author's knowledge until now there is only one application of nano-FTIR in amyloid studies reported in the literature [53]. The pioneering work of Hillenbrand and coworkers involves an in-depth investigation into the secondary structure of insulin

aggregates. Achieved spectra were interpreted in terms of insulin α-helical and/or β-sheet structure and indicated a high content of α-helical in 3-nm-thick fibril-like aggregates, as presented in Fig. 9.3. The secondary structure was determined by deconvolution of amide I mode (Fig. 9.3D) Researchers also have proven the extremely high sensitivity of the nano-FTIR technique, reaching about 0.8×10^{-18} g (1 attogram) of the investigated individual protein complex.

9.2.8 Atomic force microscopy–infrared

AFM-IR approach is based on the detection of the thermal expansion of the sample induced upon absorption of infrared light. Each AFM-IR instrument is equipped with a tunable IR laser to irradiate the sample. Two types of lasers are used: QCL and optical parametric oscillator. The absorbed pulse of IR light at particular wavelengths causes rapid thermal expansion of the sample. This expansion is detected by the AFM cantilever and registered by a photodiode as a change in deflection. The amplitude of the Fourier transform from the lever deflection signal (AFM-IR signal) is proportional to the IR absorption of the sample. Local absorption spectra are generated by the amplitude variation of the cantilever oscillations as a function of the particular absorbed wavelength from the mid-infrared spectral range [54].

AFM-IR is a very efficient tool in studies of abnormal protein aggregation. Several systems have been already investigated [55,56,59,61,62], including the first exon of the Huntingtin protein [55], the Alzheimer's amyloid β [56], supramolecular aggregates of diphenylalanine, the core recognition module of β-amyloid peptide and its derivative Boc-diphenylalanine [61], and (ERα17p) peptide issued from the hinge region of the estrogen receptor α involved in the activation of the breast carcinoma cell proliferation [62].

AFM-IR allows also for chemical and structural characterization of the oligomeric and fibrillar species formed along the aggregation of the Josephin domain of ataxin 3 [59]. AFM-IR spectra were acquired from individual amyloid forms. The obtained data was analyzed by multivariate data analysis—principal component analysis (PCA). The results were correlated with an independent measurement of the evolution of the intrinsic stiffness of single peptide aggregates.

In this work, an in-depth investigation into the conformational transition from native spheroidal oligomers, through misfolded oligomers, to the final amyloid fibrils was performed. The secondary structure of the Josephin aggregates was investigated with a spatial resolution down to 50 nm. High sensitivity of AFM-IR allowed monitoring at the nanoscale an α–β transition in single amyloid forms. The obtained results confirmed that before incubation the secondary structure of Josephin inside the spheroidal oligomers is clearly native. PCA divided AFM-IR spectra into three groups, corresponding to three types of aggregates: native oligomers, misfolded oligomers, and amyloid fibrils. The main spectral

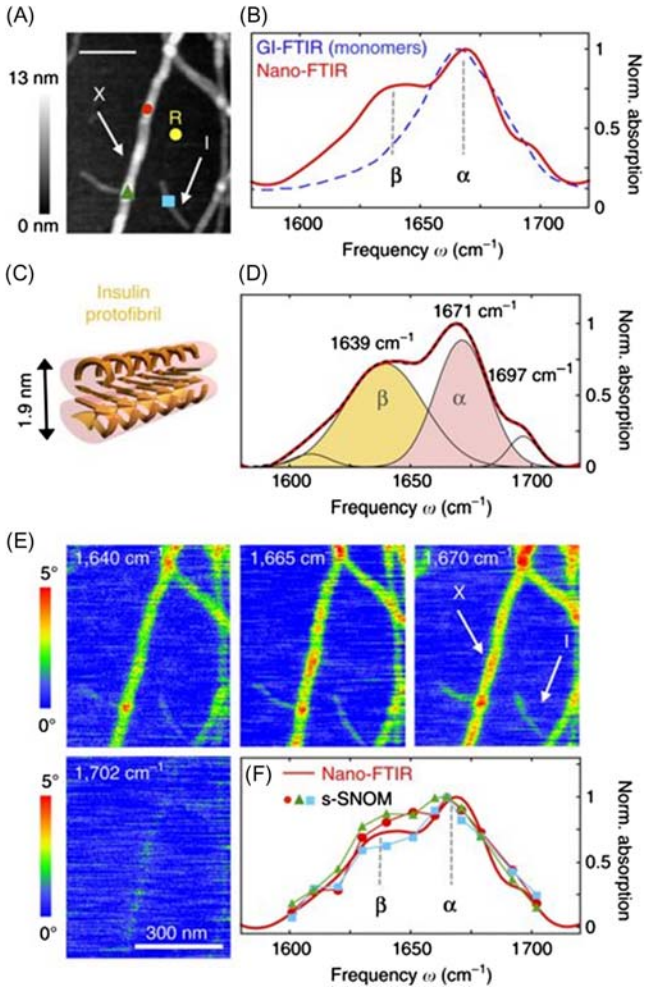


Figure 9.3

Infrared nanospectroscopy and nanoimaging of secondary structure in individual insulin fibrils. (A) Topography of insulin fibrils on a silicon substrate. Scale bar, 200 nm. The arrows indicate a type I fibril (I) and a 9-nm-thick fibril composed of several protofilaments (X), respectively. (B) Nano-FTIR spectrum of a 9-nm-thick insulin fibril (*red*, average of 154 interferograms; 8 cm^{-1} resolution; $\times 4$ zero filling) recorded at the position marked by the red symbol in (A). The position marked by the yellow dot and R indicates where the reference spectrum was recorded. The dashed blue line shows for comparison a GI-FTIR spectrum of insulin monomers on a gold substrate. (C) Illustration of the structure of an amyloid-like insulin protofibril (a type I fibril consists of two protofibrils). (D) Band decomposition of the nano-FTIR spectrum (*red curve*) based on five absorption bands (*thin black curves*). The dashed black curve shows the resulting fit. (E) s-SNOM phase images of the fibrils shown in (A). Scale bar, 300 nm. (F) Local infrared absorption spectra (*symbols*) depicting the normalized imaginary part of the near-field signal at the positions marked in (A). The data points were extracted from 12 near-field amplitude and phase images and were normalized to the imaginary part of the near-field signal on the silicon substrate. Four of the phase images are shown in (E). For comparison, the nano-FTIR spectrum of (B) is depicted by the red thick curve. All spectra are normalized to their maximum value. Source*: Reprinted by permission from Springer Nature [I. Amenabar, S. Poly, W. Nuansing, E.H. Hubrich, A.A. Govyadinov, F. Huth, R. Krutokhvostov, L. Zhang, M. Knez, J. Heberle, A.M. Bittner, R. Hillenbrand, Structural analysis and mapping of individual protein complexes by infrared nanospectroscopy, Nat. Commun. 4 (2013) 2890] Copyright (2013).*

differences between these three groups of spectra involves wave numbers related to the amide I in the spectral range from 1710 to 1680 cm^{-8}, to the COO^- vibration around 1430 cm^{-3}, and to the amide III at 1285 cm^{-8}. PCA has confirmed a decrease in random coil and α-helical structure (1655 cm^{-5}), an increase in antiparallel β-sheet content (169 cm^{-1}), and also a decrease in the native β-sheet content as a function of fibrillization maturity.

QI force–volume proved that the intrinsic stiffness (Young's modulus) of the investigated peptide increases as a function of the aggregation time (Fig. 9.4). Both independent

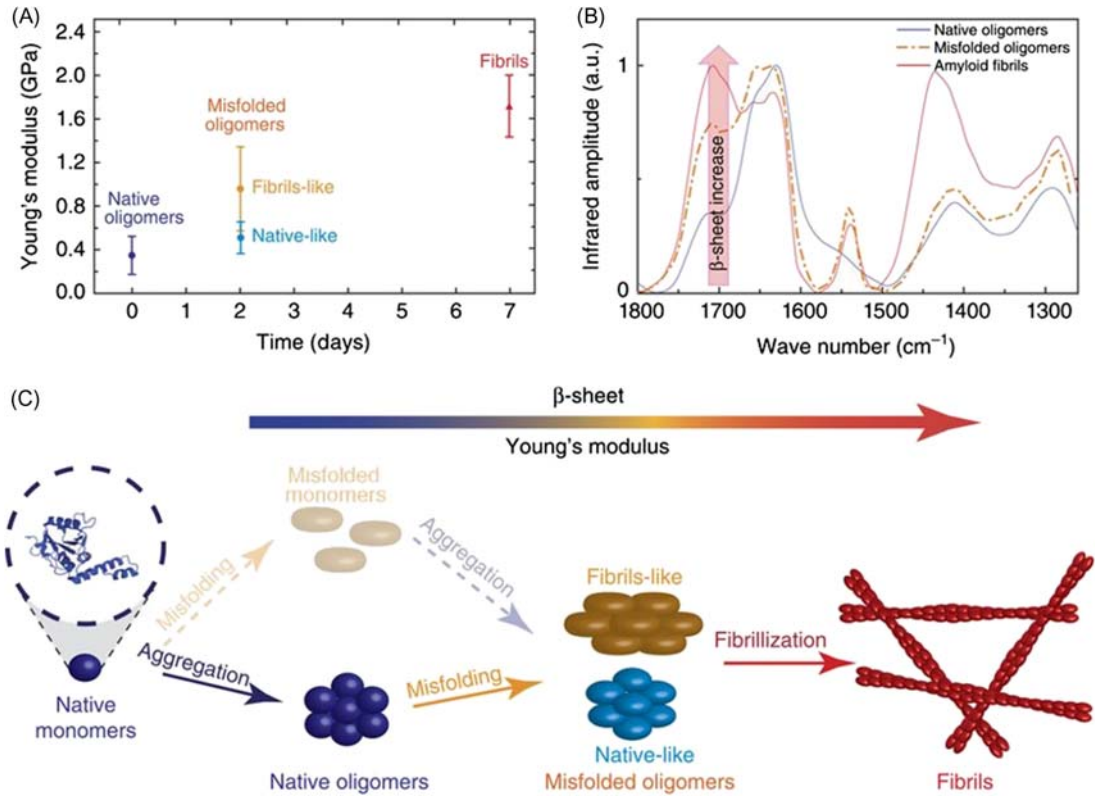


Figure 9.4

(A) The increase in the Young's modulus as a function of fibrillization maturity (the *error bars* are defined as the standard deviation of the distribution of the stiffness values of the aggregates). (B) Spectra of native oligomers at 0 day (*blue*), misfolded oligomers at 2 days (*orange and light blue*) and amyloid fibrils at 7 days (*red*). The red arrow indicates the increase in the content of β-sheet secondary structure. (C) Model of the possible pathways of Josephin aggregation: the transparent model refers to the generally accepted model "first-misfolding-then-aggregation," while the solid one refers to the new suggested model of "first-aggregation-then-misfolding." Source*: Reproduced from F.S. Ruggeri, et al., Infrared nanospectroscopy characterization of oligomeric and fibrillar aggregates during amyloid formation, Nat. Commun. 6 (2015) 7831. doi: 10.1038/ncomms8831, used under CC BY 4.0, http://creativecommons.org/licenses/by/4.0/.*

measurements, AFM-IR and stiffness, have proven the intermediate nature of the oligomers, having secondary structures partially more similar to native monomers, and partially to fibrils. The results confirmed that the content of β-sheets in the aggregates during the fibrillization process is a significant parameter affecting the nanomechanical properties of neurodegenerative peptides and determines the increase of amyloids' intrinsic stiffness.

Summarizing, the group of Prof. Dietler from EPFL and Prof. Pastore from King's College suggested the alternative aggregation mechanism of Josephin, which starts from the monomers' aggregation, forming spheroidal intermediates within a native structure. Successively, these intermediates evolve into misfolded aggregates and then into the fibrils.

A very important aspect of studies of the fibrillation process is the destructive influence of dehydration on the secondary structure of protein [63]. Delicate, biochemical structures such as folded peptides should be studied under native conditions, otherwise they may lose their conformational properties, which are crucial for their biochemical activity [27,63]. Therefore recent studies of Centrone et al. [61] deserve special attention from the reader. Researchers for the first time applied AFM-IR technique to a neurodegenerative peptide system immersed in liquid, and obtained comparable SNR and lateral resolution to IR spectra and maps acquired in air and water [61].

9.2.8.1 TERS

TERS, similar to infrared nanospectroscopy, combines the nanometric spatial resolution of AFM or STM and the chemical selectivity of molecular spectroscopy [64]. A metal nanostructure, which is accumulated at the apex of an AFM probe (AFM-TERS) or STM probe (STM-TERS) apex itself, modifies the electromagnetic field of an incident (Raman) laser enhancing the Raman scattering cross section from the extremely small amount of sample located close to the tip. In summary, TERS exploits the near-field enhancement generated by laser excitation of surface plasmons resulting from a strong electromagnetic field generated at the laser-irradiated probe apex of a metal or metallized tip, as schematically presented in Fig. 9.5 [64–66]. In other words, the basis of TERS is very similar to that of nano-FTIR, because in TERS the probe also works as a nanoantenna and converts the electric field of an incident laser into localized energy [67] (improving special resolution down to a few nanometers in vacuum and low temperature [68,69]) and converting the near field from the sample to a far field, accessible to the objective (increasing sensitivity to the single molecule level [70]). TERS was used in various measurements of biological samples including lipids [70], nucleic acids [65,66,71–73], and proteins [74–86].

As already mentioned, TERS' configuration simultaneously delivers highly resolved topographic information and therefore it seems to be ideal for studying proteins' conformations and their chemical composition [14,76–81]. The significant achievements of

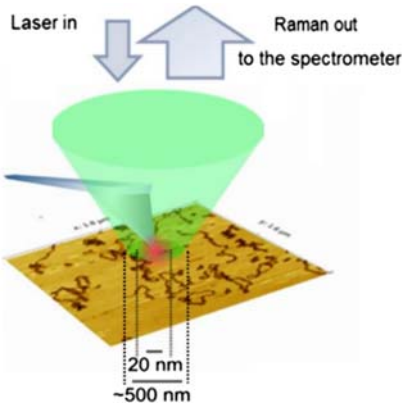


Figure 9.5
Schematic representation of the TERS setup: incident laser green light is focused on the sample, atomic force microscopy (AFM) tip modified with plasmonic nanostructures in contact with the sample surface, leads to an enhancement of the Raman scattered light. The TER signal is generated from a very small spot, localized at the probe apex, improving the spatial resolution by more than two orders of magnitude.

Prof. Deckert's group from Jena University in applications of TERS in studies of protein secondary structures should be highlighted [76–81]. Researchers from Jena investigated the molecular structure of human islet amyloid polypeptide fibrils surface [76], probed the chemical composition of insulin fibrils [77], as well as the distinct morphology, stiffness, and chemical structure of other amyloid fibril polymorphs, including β-lactoglobulin and insulin amyloids fibrils [78,81]. In recent studies, the group from Jena applied TERS combined with immune-gold labeling in studies of albumin/hemoglobin nanofibers. Such protein nanofibers might be used as universal building blocks for the bottom-up preparation of functional materials with individual properties. An application of a nanospectroscopic technique such as TERS allowed the demonstration of the heterogenic nature of the hybrid protein nanofibers presented in Fig. 9.6. This study allowed exploration of the high albumin/hemoglobin composition ratio. This work proved that the self-assembly approach supported by ethanol-induced protein denaturation is an efficient method to create hybrid protein nanofibers composed of two plasma proteins. The fast self-assembling kinetics of albumin was also proven in this research [80].

9.2.9 Hyperspectral nanospectroscopic mapping

Current applications of TERS and infrared nanospectroscopy mainly involve reports on only single-point spectra at selected locations [53–62,71–81] except for two articles. Paulite et al. for the first time successfully applied STM-TERS to map Amyloid $\beta_{1\text{-}40}$. On the basis

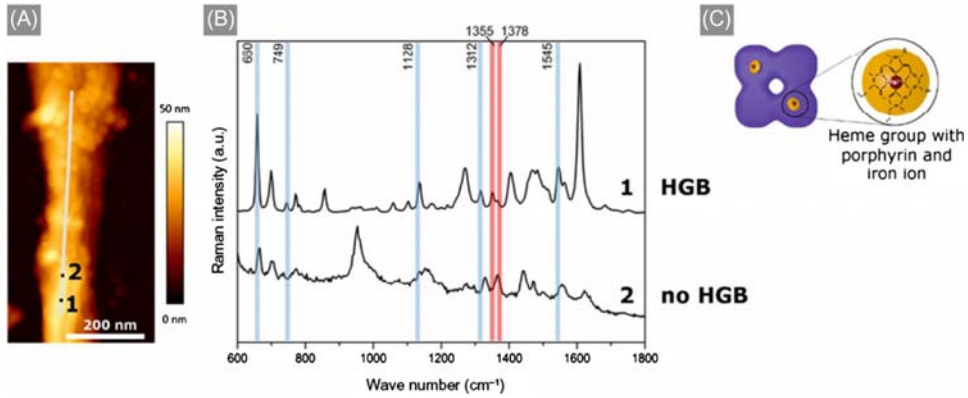


Figure 9.6

(A) AFM topography image of a single HSA-HGB fiber. The white line represents the TERS measurement area and the two points where the spectra in (B) were taken. (B) Two representative TERS spectra with and without characteristic HGB bands. For a better comparison, the intensity of spectrum 1 was lowered. In the shown spectra, corresponding bands for porphyrin (660, 749, 1128, 1312,1545 cm^{-1}) are marked blue and for Fe^{2+} (1355 cm^{-1}) and Fe^{3+} (1378 cm^{-1}) are marked red (excitation at 532 nm, incident laser power 360 μW, acquisition time: 1 s). The band around 1640–1679 cm^{-1} corresponds to the amide I mode of the peptide bonds of the proteins. Note, the amide I band can be suppressed, most likely due to the protein backbone, large amino acid side chains or protein orientation. (C) A schematic representation of an HGB molecule with two heme groups containing porphyrin rings and the iron cation of the heme group. Source: *Adapted with permission from C. Helbing, T. Deckert-Gaudig, I. Firkowska-Boden, G. Wei, V. Deckert, K.D. Jandt, Protein handshake on the nanoscale: how albumin and hemoglobin self-assemble into nanohybrid fibers, ACS Nano 12 (2) (2018) 1211–1219. Copyright 2013 American Chemical Society.*

of the integration of the phenylalanine ring breathing mode the first hyperspectral STM-TERS maps of amyloid fibrils were achieved [75].

The reproducible and intense peak from phenylalanine at 1004 cm^{-1} dominates each STM-TER spectrum, therefore it was possible to obtain TER maps corresponding to topography features presented in Fig. 9.7A–C. Other spectral features characteristic for proteins, such as amide modes or C–H/N–H motions, allowing for conformation detection based on spectral positions are mostly not well-resolved in the STM-TER spectra, as shown in Fig. 9.7D. The high enhancement of phenyl ring vibrations in STM-TERS (gap mode) has been discussed in depth by Blum et al. [83]. Recently, nanospectroscopic hyperspectral mapping of amyloid β was achieved also by AFM-TERS [14]. This research allowed for localization of the β-sheet conformation distribution in oligomers and along fibrils and protofibrils. It was pioneering proof that TERS is an efficient tool in studies of amyloids and allows for direct following of their aggregation pathways. Lipiec et al. found a link between the distribution of the β-sheet secondary structure in single amyloid-β forms (such as oligomers, protofibrils, fibrils) and a particular scheme of amyloid-β aggregation.

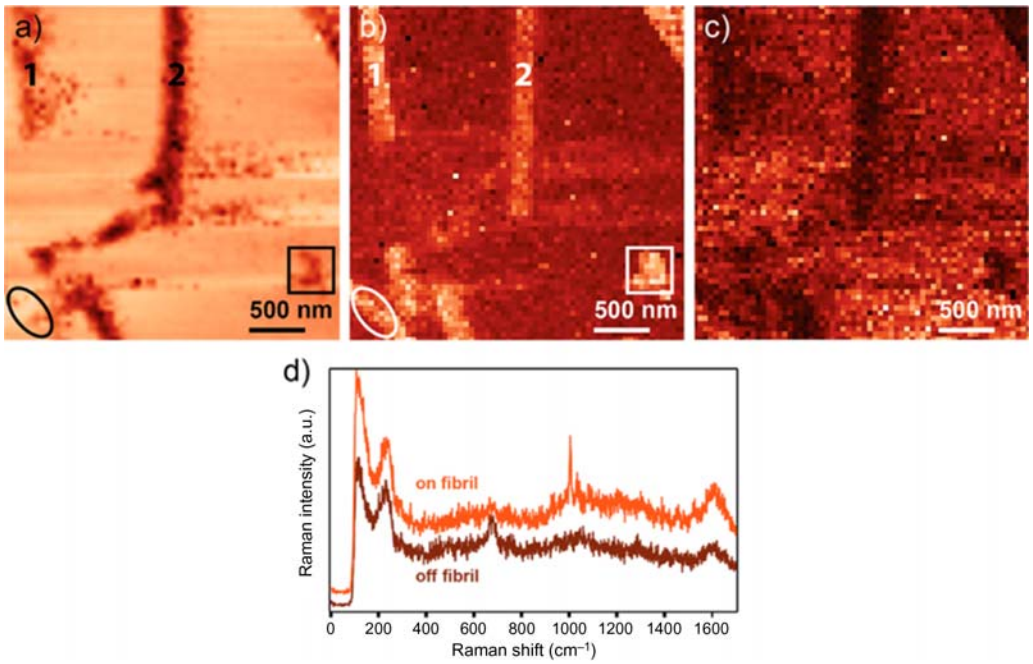


Figure 9.7

Simultaneously acquired (A) STM and (B,C) TERS (acquisition time 1 s/pixel, 2 mW incident power) images of individual nanotapes with $3 \times 3\ \mu m^2$ scan size and 64×64 pixel resolution. The color-coded TER images display (B) the intensity of the aromatic ring breathing marker band (1004 cm^{-1}, value of the peak integral) and (C) the band at 675 cm^{-1} that can be seen in the brown spectrum in (D). (D) Representative TER spectra taken from the full spectrum map that the TER images in (B,C) are based on. Nanotapes 1 and 2 and the circle and square, respectively, highlighting the different STM contrasts and TERS intensities in each region, strongly giving evidence of observed structural and/or chemical heterogeneity between the nanostructures.

Source: *Adapted with permission from M. Paulite, C. Blum, T. Schmid, L. Opilik, K. Eyer, G.C. Walker, R. Zenobi, Full spectroscopic tip-enhanced Raman imaging of single nanotapes formed from β-amyloid (1–40) peptide fragments, ACS Nano 7 (2013) 911. Copyright 2013 American Chemical Society.*

Fig. 9.7 demonstrates the AFM topography of a fibril sample (gray color scale) and its zoomed area, which is partially overlapped by the TER map in central part. Two types of TER spectra (green and blue Fig. 9.8), were collected in the area of the fibril and the distribution of each of those spectra was superimposed on AFM topography. These spectra were extracted by an application of multivariate data analysis including unsupervised hierarchical cluster analysis and PCA, as marker spectra β-sheet and turns/unstructured coils secondary structure of amyloid-β. The difference between these two types of spectra acquired in the fibril area is a significant shift of the amide III band spectral position from 1260 cm^{-1} (green spectrum) to 1250 cm^{-1} (blue spectrum). Based on the spectral position of conformation marker band amide III, the peptide secondary structure such as

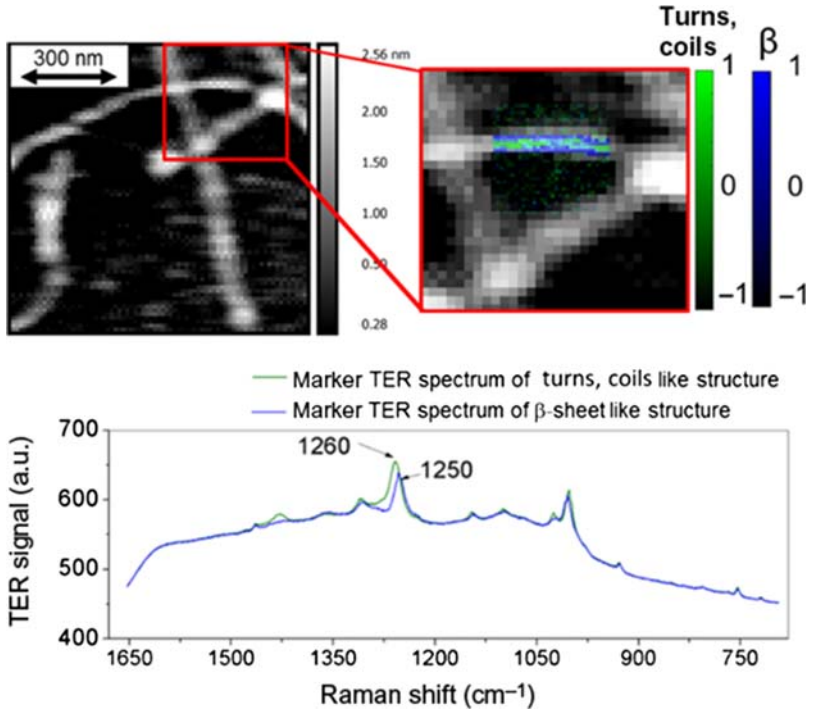


Figure 9.8

AFM-TERS mapping of $A\beta_{(1-42)}$ fixed on gold: AFM topography with overlapped TER map in zoomed area showing the distribution of β-sheets (*blue*) and turns/unstructured coils (*green*) with corresponding averaged TER spectra, characteristic for both conformations. Source: *Reproduced with permission from E. Lipiec, D. Perez-Guaita, J. Kaderli, B.R. Wood, R. Zenobi, Direct nanospectroscopic verification of the amyloid aggregation pathway. Angew. Chem. Int. Ed. 57 (28) (2018) 8519–8524. © Wiley-VCH Verlag GmbH & Co. KGaA, 2018.*

turns/random coil from β-sheet could be assigned [14,79,87,88]. Additionally, a slightly higher intensity of the Cα−H/N−H bending mode at 1364 cm^{-1} [87,88], which was also detected in the blue spectrum, is considered as a Raman marker of β-sheet secondary structure. These studies have proven that nanoscale chemical mapping of single amyloid-β forms at various stages of aggregation allow for capturing the fibril formation process. The aggregation of amyloid-β, which starts from the merging of monomers to metastable oligomers, which then rearrange to ordered β-sheets, is already at the oligomeric or protofibrillar stage [14].

9.3 Conclusions

Molecular nanospectroscopic methods, including TERS, AFM-IR, and nano-FTIR, are noninvasive, label-free techniques, which deliver substantial information about chemical

structures and composition of samples at the nanoscale. Therefore they brought new insights into amyloid research. Several neurodegenerative peptide systems were studied, including amyloid-β, Josephin, β-lactoglobulin, insulin, and many others. Deep nanoscale investigation into secondary structure of amyloids allowed the monitoring of single steps of their aggregation. Despite the high capability of infrared nanospectroscopy in studies of amyloids, including measurements in liquids, so far only single-point spectra at selected locations or single-frequency mapping have been reported. The reason for that is related to the unstable enhancement, which is wavelength dependent, and therefore for nano-FTIR this instability could be more problematic than for TERS, which uses monochromatic laser light. AFM-IR is fairly reproducible, but the sensitivity is relatively low in comparison with plasmonic techniques such as nano-FTIR and TERS, and therefore obtaining sufficient signal-to-noise ratio in spectra from such thin samples (several nm height or less) is very difficult. TER hyperspectral maps were recorded due to high sensitivity (down to single molecules), the noninvasive nature of the AFM, which can be operated in tapping mode, and a shorter time of spectra acquisition [14]. Each of these three described techniques exhibits unique advantages, and the choice of appropriate methodology should be determined by research project goals and the properties of the investigated sample, primarily thickness/size, Raman scattering/infrared absorption cross sections, and thermomechanical properties.

References

[1] R. Riek, D.S. Eisenberg, The activities of amyloids from a structural perspective, Nature 539 (7628) (2016) 227–235.
[2] T.P. Knowles, M. Vendruscolo, C.M. Dobson, The amyloid state and its association with protein misfolding diseases, Nat. Rev. Mol. Cell Biol. 15 (6) (2014) 384–396.
[3] M. Jucker, L.C. Walker, Self-propagation of pathogenic protein aggregates in neurodegenerative diseases, Nature 501 (7465) (2013) 45–51.
[4] J. Hardy, D.J. Selkoe, The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics, Science 297 (5580) (2002) 353–356.
[5] M. Goedert, Nature familial Parkinson's disease, Awakening Alpha-Synuclein 388 (6639) (1997) 232–233.
[6] H.Y. Zoghbi, H.T. Orr, Glutamine repeats and neurodegeneration, Annu. Rev. Neurosci. 23 (2000) 217–247.
[7] https://www.alz.co.uk/research/statistics
[8] A. Wimo, M. Guerchet, G.-C. Ali, Y.-T. Wu, A.M. Prina, B. Winblad, et al., The worldwide costs of dementia 2015 and comparisons with 2010, Alzheimer's Dement. 13 (1) (2017) 1–7.
[9] L.C. Walker, H. LeVine, The cerebral proteopathies, Neurobiol. Aging 21 (4) (2000) 559–561.
[10] J. Greenwald, R. Riek, Biology of amyloid: structure, function, and regulation, Structure 18 (10) (2010) 1244–1260.
[11] V. Uversky, D. Eliezer, Biophysics of Parkinson's disease: structure and aggregation of alpha-synuclein, Cur. Protein Pept. Sci. 10 (5) (2009) 483–499.
[12] C.A. Ross, M.A. Poirier, Protein aggregation and neurodegenerative disease, Nat. Med. 10 (2004) 10–17.
[13] C.A. Ross, M.A. Poirier, What is the role of protein aggregation in neurodegeneration? Nat. Rev. Mol. Cell Biol. 6 (11) (2005) 891–898.

[14] E. Lipiec, D. Perez-Guaita, J. Kaderli, B.R. Wood, R. Zenobi, Direct nanospectroscopic verification of the amyloid aggregation pathway, Angew. Chem. Int. Ed. 57 (28) (2018) 8519–8524.
[15] E.D. Eanes, G.G. Glenner, X-ray diffraction studies on amyloid filaments, J. Histochem. Cytochem. 16 (11) (1968) 673–677.
[16] N.J. Greenfield, Using circular dichroism spectra to estimate protein secondary structure, Nat. Protoc. 1 (6) (2006) 2876–2890.
[17] A. Barth, Infrared spectroscopy of proteins, Biochim. Biophys. Acta (BBA) Bioenerg. 1767 (9) (2007) 1073–1101.
[18] J.J. Meiler, PROSHIFT: protein chemical shift prediction using artificial neural networks, J. Biomol. NMR 26 (2003) 25–37.
[19] L. Bertini, C. Gonnelli, J. Luchinat, A. Mao, A. Nesi, New structural model of Aβ40 fibrils, J. Am. Chem. Soc. 133 (40) (2011) 16013–16022.
[20] S.L. Bernstein, N.F. Dupuis, N.D. Lazo, T. Wyttenbach, M.M. Condron, G. Bitan, et al., Amyloid-β protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease, Nat. Chem. 1 (4) (2009) 326–331.
[21] M. Ahmed, J. Davis, D. Aucoin, T. Sato, S. Ahuja, et al., Amyloid-β protein oligomerization and the importance of tetramers and dodecamers in the aetiology of Alzheimer's disease, Nat. Struct. Mol. Biol. 17 (2010) 561–567.
[22] T. Petkova, W.M. Yau, R. Tycko, Experimental constraints on quaternary structure in Alzheimer's beta-amyloid fibrils, Biochemistry. 45 (2) (2006) 498–512.
[23] M. Cheon, C.K. Hall, I. Chang, Structural conversion of Aβ17–42 peptides from disordered oligomers to U-shape protofilaments via multiple kinetic pathways, PLoS Comput. Biol. 11 (2015) 5–27.
[24] M. Margittai, R. Langen, Fibrils with parallel in-register structure constitute a major class of amyloid fibrils: molecular insights from electron paramagnetic resonance spectroscopy, Rev. Biophys. 41 (3–4) (2008) 265–297.
[25] T. Luhrs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Dobeli, et al., 3D structure of Alzheimer's amyloid-beta(1-42) fibrils, Proc. Natl. Acad. Sci. 102 (48) (2005) 17342–17347.
[26] J.S. Jeong, A. Ansaloni, R. Mezzenga, H.A. Lashuel, G. Dietler, Novel mechanistic insight into the molecular basis of amyloid polymorphism and secondary nucleation during amyloid formation, J. Mol. Biol. 425 (10) (2013) 1765–1781.
[27] J. Habchi, P. Arosio, M. Perni, A.R. Costa, M. Yagi-Utsumi, P. Joshi, et al., An anticancer drug suppresses the primary nucleation reaction that initiates the production of the toxic Aβ42 aggregates linked with Alzheimer's disease, Sci. Adv. 2 (2) (2016) e1501244.
[28] J. Habchi, S. Chia, C. Galvagnion, T.C.T. Michaels, M.M.J. Bellaiche, F.S. Ruggeri, et al., Cholesterol catalyses Aβ42 aggregation through a heterogeneous nucleation pathway in the presence of lipid membranes, Nat. Chem. 10 (6) (2018) 673–683.
[29] E. Hellstrand, B. Boland, D.M. Walsh, S. Linse, Amyloid β-protein aggregation produces highly reproducible kinetic data and occurs by a two-phase process, ACS Chem. Neurosci. 1 (1) (2010) 13–18.
[30] T.P. Knowles, C.A. Waudby, G.L. Devlin, S.I. Cohen, A. Aguzzi, M. Vendruscolo, et al., An analytical solution to the kinetics of breakable filament assembly, Science 326 (5959) (2009) 1533–1537.
[31] P. Arosio, T.P.J. Knowles, S. Linse, On the lag phase in amyloid fibril formation, Phys. Chem. Chem. Phys. 17 (2015) 7606–7618.
[32] Z. Wang, C. Zhou, C. Wang, L. Wan, X. Fang, C. Bai, AFM and STM study of beta-amyloid aggregation on graphite, Ultramicroscopy 97 (1-4) (2003) 73–79.
[33] T. Watanabe-Nakayama, K. Ono, M. Itami, R. Takahashi, D.B. Teplow, M. Yamada, High-speed atomic force microscopy reveals structural dynamics of amyloid β1–42 aggregates, Proc. Natl. Acad. Sci. 113 (21) (2016) 5835–5840.
[34] M.J. Rust, M. Bates, X. Zhuang, Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM), Nat. Methods 3 (2006) 793–795.

[35] E. Betzig, et al., Imaging intracellular fluorescent proteins at nanometer resolution, Science 313 (5793) (2006) 1642–1645.
[36] G.S. Kaminski-Schierle, et al., In situ measurements of the formation and morphology of intracellular β-amyloid fibrils by super-resolution fluorescence imaging, J. Am. Chem. Soc. 133 (33) (2011) 12902–12905.
[37] E.K. Esbjorner, et al., Direct observations of amyloid β self-assembly in live cells provide insights into differences in the kinetics of Aβ(1–40) and Aβ(1–42) aggregation, Chem. Biol. 21 (6) (2014) 732–742.
[38] D. Pinotsi, et al., Direct observation of heterogeneous amyloid fibril growth kinetics via two-color super-resolution microscopy, Nano Lett. 14 (1) (2014) 339–345.
[39] M.J. Roberti, et al., Imaging nanometer-sized α-synuclein aggregates by superresolution fluorescence localization microscopy, Biophys. J. 102 (7) (2012) 1598–1607.
[40] M.M. Apetri, et al., Direct observation of α-synuclein amyloid aggregates in endocytic vesicles of neuroblastoma cells, PLoS One 11 (2016).
[41] W.C. Duim, B. Chen, J. Frydman, W.E. Moerner, Sub-diffraction imaging of huntingtin protein aggregates by fluorescence blink-microscopy and atomic force microscopy, Chem. Phys. Chem 12 (13) (2011) 2387–2390.
[42] S.J. Sahl, L.E. Weiss, W.C. Duim, J. Frydman, W.E. Moerner, Cellular inclusion bodies of mutant huntingtin exon 1 obscure small fibrillar aggregate species, Sci. Rep. 2 (2012) 895.
[43] W.C. Duim, Y. Jiang, K. Shen, J. Frydman, W.E. Moerner, Super-resolution fluorescence of huntingtin reveals growth of globular species into short fibers and coexistence of distinct aggregates, ACS Chem. Biol. 9 (12) (2014) 2767–2778.
[44] J. Ries, et al., Superresolution imaging of amyloid fibrils with binding-activated probes, ACS Chem. Neurosci. 4 (7) (2013) 1057–1061.
[45] R. Sabaté, S.J. Saupe, Thioflavin-T excimer formation upon interaction with amyloid fibers, Biochem. Biophys. Res. Commun. 360 (2007) 135–138.
[46] M.R.H. Krebs, E.H.C. Bromley, A.M. Donald, The binding of thioflavin-T to amyloid fibrils: localisation and implications, J. Struct. Biol. 149 (1) (2005) 30–37.
[47] J. Duboisset, P. Ferrand, W. He, X. Wang, H. Rigneault, S. Brasselet, Thioflavine-T and congo red reveal the polymorphism of insulin amyloid fibrils when probed by polarization-resolved fluorescence microscopy, J. Phys. Chem. B 117 (3) (2013) 784–788.
[48] M.T. Colvin, R. Silvers, Q.Z. Ni, T.V. Can, I. Sergeyev, M. Rosay, et al., Atomic resolution structure of monomorphic Aβ42 amyloid fibrils, J. Am. Chem. Soc. 138 (30) (2016) 9663–9674.
[49] L. Gremer, D. Schölzel, C. Schenk, E. Reinartz, J. Labahn, R.B.G. Ravelli, et al., Cryo-EM structure of an amyloid fibril from systemic amyloidosis, Science 6 (2017) 116–119.
[50] J.M. Taylor, P. Cappellaro, L. Childress, L. Jiang, D. Budker, P.R. Hemmer, et al., High-sensitivity diamond magnetometer with nanoscale resolution, Nat. Phys. 4 (2008) 810–816.
[51] T. Häberle, D. Schmid-Lorch, F. Reinhard, J. Wrachtrup, Nanoscale nuclear magnetic imaging with chemical contrast, Nat. Nanotechnol. 10 (2) (2015) 125–128.
[52] J. Wallentin, Dl Jacobsson, M. Osterhoff, M.T. Borgstro, T. Salditt, Bending and twisting lattice tilt in strained core − shell nanowires revealed by nanofocused X-ray diffraction, Nano Lett. 17, 2017, pp. 4143–4150.
[53] I. Amenabar, S. Poly, W. Nuansing, E.H. Hubrich, A.A. Govyadinov, F. Huth, et al., Structural analysis and mapping of individual protein complexes by infrared nanospectroscopy, Nat. Commun. 4 (2013) 2890.
[54] Dazzi, C.B. Prater, Q. Hu, D.B. Chase, J.F. Rabolt, C. Marcott, AFM-IR: combining atomic force microscopy and infrared spectroscopy for nanoscale chemical characterization, Appl. Spectrosc. 66 (12) (2012) 1365–1384.
[55] F.S. Ruggeri, S. Vieweg, U. Cendrowska, G. Longo, A. Chiki, H.A. Lashuel, et al., Nanoscale studies link amyloid maturity with polyglutamine diseases onset, Sci. Rep. 6 (2016) 31155.

[56] D. Galante, F.S. Ruggeri, G. Dietler, F. Pellistri, E. Gatta, A. Corsaro, et al., A critical concentration of N-terminal pyroglutamylated amyloid beta drives the misfolding of Ab1-42 into more toxic aggregates, Int. J. Biochem. Cell Biol. 79 (2016) 161–170.
[57] F. Huth, M. Schnell, J. Wittborn, N. Ocelic, R. Hillenbrand, Infrared-spectroscopic nanoimaging with a thermal source, Nat. Mater. 10 (2011) 352–356.
[58] F. Keilman, R. Hillenbrand, in: D. Richards, A. Zayats (Eds.), Nano-Optics and Near-Field Optical Microscopy, Artech House, Boston, MA/London, 2009.
[59] F.S. Ruggeri, G. Longo, S. Faggiano, E. Lipiec, A. Pastore, G. Dietler, Infrared nanospectroscopy characterization of oligomeric and fibrillar aggregates during amyloid formation, Nat. Comm. 8 (2015) 7831.
[60] N. Ocelic, A. Huber, R. Hillenbrand, Pseudoheterodyne detection for background-free near-field spectroscopy, Phys. Lett. 89 (2006) 101124.
[61] G. Ramer, F.S. Ruggeri, A. Levin, T.P.J. Knowles, A. Centrone, Determination of polypeptide conformation with nanoscale resolution in water, ACS Nano. 12 (7) (2018) 6612–6619.
[62] F.S. Ruggeri, C. Byrne, L. Khemtemourian, G. Ducouret, G. Dietler, Y. Jacquot, Concentration-dependent and surface-assisted self-assembly properties of a bioactive estrogen receptor α-derived peptide, J. Pept. Sci. 21 (2) (2015) 95–104.
[63] G. Maurstad, M. Prass, L.C. Serpell, P. Sikorski, Dehydration stability of amyloid fibrils studied by AFM, Eur. Biophys. J. 38 (2009) 1135–1140.
[64] R.M. Stöckle, Y.D. Suh, V. Deckert, R. Zenobi, Nanoscale chemical analysis by tip-enhanced Raman spectroscopy, Chem. Phys. Lett. 318 (2000) 131–136.
[65] E. Lipiec, R. Sekine, J. Bielecki, W.M. Kwiatek, B.R. Wood, Molecular characterization of DNA double strand breaks with tip-enhanced Raman scattering, Angew. Chem. 53 (1) (2014) 169–172.
[66] E. Lipiec, B. Wood, Encyclopedia of Analytical Chemistry, Online © 2006–2017, John Wiley & Sons, Ltd, 2017.
[67] L. Novotny, N. van Hulst, Antennas for light nat, Photonics 5 (2) (2011) 83.
[68] M. Liao, S. Jiang, C. Hu, R. Zhang, Y. Kuang, Y. Zhu, et al., Tip-enhanced Raman spectroscopic imaging of individual carbon nanotubes with subnanometer resolution, Nano Lett. 16 (7) (2016) 4040.
[69] R. Zhang, Y. Zhang, Z.C. Dong, S. Jiang, C. Zhang, L.G. Chen, et al., Chemical mapping of a single molecule by plasmon-enhanced Raman scattering, Nature 498 (2013) 82–86.
[70] L. Opilik, T. Bauer, T. Schmid, J. Stadlera, R. Zenobi, Nanoscale chemical imaging of segregated lipid domains using tip-enhanced Raman spectroscopy, Phys. Chem. Chem. Phys. 13 (2011) 9978.
[71] A. Japaridze, D. Vobornik, E. Lipiec, A. Cerreta, J. Szczerbinski, R. Zenobi, et al., Macromolecules, toward an effective control of DNA's submolecular conformation on a surface, Macromolecules 49 (2016) 643–652.
[72] E. Lipiec, A. Japaridze, J. Szczerbiński, G. Dietler, R. Zenobi, Preparation of well-defined DNA samples for reproducible nanospectroscopic measurements, Small 12 (35) (2016) 4821–4829.
[73] S. Najjar, D. Talaga, L. Schué, Y. Coffinier, S. Szunerits, R. Boukherroub, et al., Tip-enhanced Raman spectroscopy of combed double-stranded DNA bundles, J. Phys. Chem. C. 118 (2014) 1174–1181.
[74] S. Bonhommeau, D. Talaga, J. Hunel, C. Cullin, S. Lecomte, Tip-enhanced Raman spectroscopy to distinguish toxic oligomers from Aβ1-42 fibrils at the nanometer scale, Angew. Chem. Int. Ed. 56 (2) (2017) 1771–1774.
[75] M. Paulite, C. Blum, T. Schmid, L. Opilik, K. Eyer, G.C. Walker, et al., Full spectroscopic tip-enhanced Raman imaging of single nanotapes formed from β-amyloid(1-40) peptide fragments, ACS Nano 7 (2) (2013) 911–920.
[76] C.C. van den Akker, T. Deckert-Gaudig, M. Schleeger, K.P. Velikov, V. Deckert, M. Bonn, et al., Morphology and persistence length of amyloid fibrils are correlated to peptide molecular structure, Small 11 (45) (2015) 4131–4139.
[77] D. Kurouski, T. Deckert-Gaudig, V. Deckert, I.K. Lednev, Structure and composition of insulin fibril surfaces probed by TERS, J. Am. Chem. Soc. 134 (32) (2012) 13323–13329.

[78] C.C. van den Akker, M. Scleeger, A.L. Bruinen, T. Deckert-Gaudig, K.P.R. Velikov, M.A. Heeren, et al., Multimodal spectroscopic study of amyloid fibril polymorphism, Phys. Chem. B 120 (34) (2016) 8809–8817.
[79] T. Deckert-Gaudig, D. Kurouski, M.A.B. Hedegaard, P. Singh, I.K. Lednev, V. Deckert, High resolution spectroscopy reveals fibrillation inhibition pathways of insulin, J. Phys. Chem. B 120 (2016) 8809–8817.
[80] C. Helbing, T. Deckert-Gaudig, I. Firkowska-Boden, G. Wei, V. Deckert, K.D. Jandt, Protein handshake on the nanoscale: how albumin and hemoglobin self-assemble into nanohybrid fibers, ACS Nano 12 (2) (2018) 1211–1219.
[81] T. Deckert-Gaudig, V. Deckert, High resolution spectroscopy reveals fibrillation inhibition pathways of insulin, Sci. Rep. 6 (2016) 39622.
[82] S. Kradolfer, E. Lipiec, C. Baldacchini, A.R. Bizzarri, S. Cannistraro, R. Zenobi, Vibrational changes induced by electron transfer in surface bound azurin metalloprotein studied by tip-enhanced Raman spectroscopy and scanning tunneling microscopy, ACSNano 11 (12) (2017) 12824–12831.
[83] C. Blum, T. Schmid, L. Opilik, S. Weidmann, S.R. Fagerer, R. Zenobi, Vibrational changes induced by electron transfer in surface bound azurin metalloprotein studied by tip-enhanced Raman spectroscopy and scanning tunneling microscopy, J. Raman Spectrosc. 43 (2012) 1895–1904.
[84] P. Hermann, H. Fabian, D. Naumann, A. Hermelink, Comparative study of far-field and near-field Raman spectra from silicon-based samples and biological nanostructures, J. Phys. Chem. 115 (50) (2011) 24512–24520.
[85] B.R. Wood, M. Asghari-Khiavi, E. Bailo, D. McNaughton, V. Deckert, Detection of nano-oxidation sites on the surface of hemoglobin crystals using tip-enhanced Raman scattering, Nano Lett. 12 (3) (2012) 1555–1560.
[86] M. Moretti, R.P. Zaccaria, E. Descrovi, G. Das, M. Leoncini, C. Liberale, et al., Plasmonics 8 (2013) 25–33.
[87] D. Nemecek, J. Stepanek, G.J. Thomas Jr, Raman spectroscopy of proteins and nucleoproteins, Curr. Protoc. Protein Sci. 71 (2013) 17.8.1–17.8.52.
[88] M.F. Rosario-Alomar, T. Quinones-Ruiz, D. Kurouski, V. Sereda, E.B. Ferreira, L. De Jesus-Kim, et al., Hydrogen sulfide inhibits amyloid formation, J. Phys. Chem. B 119 (2015) 1265–1274.

Further Reading

G. Gao, M. Zhang, P. Lu, G. Guo, D. Wang, T. Sun, Chirality-assisted ring-like aggregation of Aβ(1–40) at liquid–solid interfaces: a stereoselective two-step assembly process, Angew. Chem. Int. Ed. 127 (2015) 2273–2278.

CHAPTER 10

Vibrational spectroscopic analysis and quantification of proteins in human blood plasma and serum

James M. Cameron[1], Clément Bruno[2,3,4], Drishya Rajan Parachalil[5,6], Matthew J. Baker[1,7], Franck Bonnier[2], Holly J. Butler[1,7] and Hugh J. Byrne[5]

[1]WestCHEM, Department of Pure & Applied Chemistry, Technology and Innovation Centre, University of Strathclyde, Glasgow, United Kingdom, [2]Faculty of Pharmacy, EA 6295 NMNS, University of Tours, Tours, France, [3]CHRU Tours, Laboratory of Biochemistry and Molecular Biology, Tours, France, [4]UMR INSERM U 1253 - iBrain, University of Tours, Tours, France, [5]School of Physical and Clinical & Optometric Sciences, Technological University Dublin, Dublin, Ireland, [6]FOCAS Research Institute, Technological University Dublin, Dublin, Ireland, [7]ClinSpec Dx Ltd., Technology and Innovation Centre, Glasgow, United Kingdom

10.1 Introduction

Bodily fluids (e.g., plasma, serum, saliva, or urine) are emerging as an important source of samples for disease diagnostics and therapeutic monitoring, as their collection is largely non-invasive and relatively simple. Analyses are performed in specialized laboratories in hospitals or medical centers, and there is a considerable time lapse between collection of bodily fluids from patients and delivery of the results, while the samples are analyzed for disease diagnostic biomarkers. Bodily fluids are collected from a large number of patients in a hospital, further delaying the performance of the analysis and availability of results, which may in turn delay the therapeutic intervention, and prolong patient anxiety to the detriment of patient management. The accuracy of the test kits that enable point-of-care-testing is often poor and they are frequently avoided due to high cost [1,2]. Furthermore, some illnesses are still without biomarkers for diagnosis and disease stage monitoring. Recently, global analysis techniques based on "omics" have emerged using multicomponent concentrations as specific biomarkers. However, there is still a need for objective and cost-effective methods that are capable of accurate earlier diagnosis from bodily fluids, in a point-of-care clinical setting while also developing new tools for the identification of novel biomarkers. For both these purposes, vibrational spectroscopic techniques are appropriate,

Vibrational Spectroscopy in Protein Research.
DOI: https://doi.org/10.1016/B978-0-12-818610-7.00010-4

as they are non-destructive, label-free, rapid, cost-effective, easy to operate, and require minimal sample preparation.

Vibrational spectroscopic techniques, involving Raman scattering and infrared (IR) absorption, are valuable analytical techniques available to scientists, as they provide the unique opportunity to investigate the molecular composition of both organic and inorganic compounds. They are routine, standard techniques used for fingerprinting and identifying chemicals, as they can give molecularly specific chemical information without the use of extrinsic labels and without being extremely invasive or destructive to the system studied. Since both techniques are truly label-free, their potential for diagnostic applications has been well investigated and demonstrated, hence interest has turned toward their application to analysis of bodily fluids, particularly human blood plasma and/or serum for disease diagnostics and patient monitoring [3,4].

10.1.1 Analysis of biofluids

For many years, the field of medicine has used bodily fluids, or biofluids, for patient monitoring and screening for disease. Within the body there are two major fluid categories; intracellular fluid (ICF) and extracellular fluids (ECF). ICF refers to fluids contained inside the cell plasma membrane while EFC can be further differentiated into the interstitial fluid (IF) surrounding cells ($\sim$75%), the fluid component of the blood called plasma ($\sim$25%), and the transcellular fluids, a small specialized fraction of excreted biofluids (urine, saliva, gastric fluids, etc.). An adult human body is composed of 50%–60% water, about two thirds of which is found in the ICF and one third in the ECF, enabling the crucial physiological process of osmosis to take place. Described as the passage of water across a semipermeable membrane along a concentration gradient, osmosis is essential for cells to maintain homeostasis. Key to this transfer of biochemical constituents between the extra- and inter-cellular environment is the blood, and therefore its composition can evolve over time, reflecting changes in patient health.

However, within this complex matrix, only a limited number of molecules, or biomarkers, have well-established symptomatic deviations from normal ranges that can be used for diagnosing specific pathologies. As a consequence, biomarker discovery has become an active field of research, aiming to provide tools for early diagnosis associated with better prognosis for patients. Cerebrospinal, synovial, pleural, pericardial, peritoneal, lymphatic fluids, the aqueous humor, and mucosal secretions, such as urine, saliva, sputum, are thus considered as part of the ECF system. These fluids derive or interact with blood and can therefore similarly contain sentinel indicators of cellular and organ function or dysfunction. Thus, human biofluids are considered to be powerful sources of clinical biomarkers [5,6]. In terms of disease diagnosis and prognosis, bodily fluids are an interesting alternative to cells and tissues [7]. It is expected that modification of the overall composition of the

biochemical state of bodily fluids could deliver crucial information about patient health and disease states, enabling early disease diagnosis and administration of treatment [8]. Disease diagnosis from bodily fluids could potentially be developed into a dynamic diagnostic environment that will enable early disease diagnosis even before the disease becomes symptomatic.

These biofluids, which provide organ-specific information, are increasingly used for diagnostics; however, blood is considered the largest biomarker reservoir of the body. Since most of the clinical analytical instruments are accurate for both serum and plasma, these two terms are used interchangeably in most clinical tests [9]. Notably, many studies that have reported the use of serum, were in fact carried out in plasma [10–14].

10.1.2 Blood sample: preparation of plasma versus serum

A range of blood-based tests are routinely conducted in a clinical setting. Each analysis protocol has strict preanalytical requirements, according to which molecular or cell specific analysis assays are to be used, even starting with the choice of sampling tubes. For example, measurement of clotting factors level or complete blood cell count cannot be analyzed in dry tubes (without anticoagulant reagents) [15]. Thus, only citrate-treated tubes are compatible with hemostasis evaluation. While plasma and serum are both cell-free fluids obtained from blood samples by centrifugation, they differ on the basis of whether the sample has been allowed to clot or not (Fig. 10.1).

For plasma preparation, tubes treated with anticoagulant (such as citrate or heparin) are used for blood collection. Routinely, refrigerated centrifugation for 10 minutes at 2000 × g concentrates unwanted cells and platelets. For serum preparation, whole blood is allowed to clot at room temperature for about 15–30 minutes. The clot is removed by refrigerated

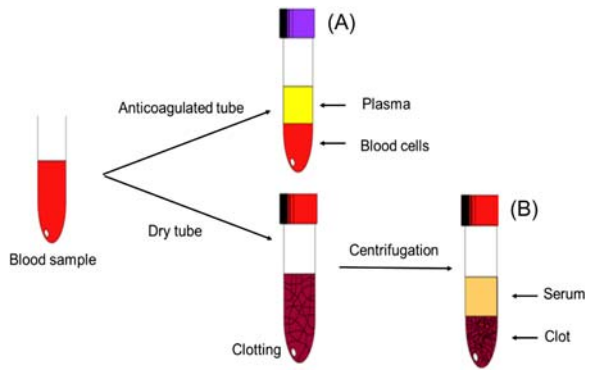


Figure 10.1
Obtaining (A) plasma and (B) serum from whole blood samples.

centrifugation at 1000−2000 × g for 10 minutes, often separated by a gel component to avoid contamination. It is important to immediately transfer the supernatant into a clean polypropylene tube and maintain samples at 2−8°C while handling. If the samples are not analyzed immediately, they should be stored at −20°C or preferably lower. It is also recommended to avoid freeze−thaw cycles, because this may have detrimental effects on many serum components.

10.1.3 Composition of plasma and serum

Blood serum and plasma are predominantly composed of water (∼90%), minerals, organic substances, and gas (oxygen and carbon dioxide). Proteins are the predominant molecular components of blood plasma, the remaining constituents being carbohydrates, lipids, and amino acids (Fig. 10.2). Serum albumin, globulins, fibrinogen, and a handful of other abundant proteins account for 99% of total serum proteins, while the remaining 1% is composed of low-abundance circulatory proteins. Additionally, plasma or serum contain more than 114,000 known metabolites at varying concentration level (<1 nmol/L to mmol/L) [16−18].

10.1.3.1 Nonprotein constituents

Electrolytes are a key constituent of blood plasma and serum that play key roles in the human body, including maintaining pH balance and cellular communication. While sodium, chloride, and bicarbonate (approximately 140, 100, and 25 mmol/L, respectively) are considered highly concentrated, others such as potassium, phosphate, calcium, and magnesium (4.5, 2.5, 2.5, and 1 mmol/L, respectively) are found with lower concentrations. Other minerals such as iron, copper, zinc, aluminum, and lead can be found in even lower concentrations (≤μmol/L).

Energy is provided to cells in the form of glucose, thus blood contains a large amount of glucose and its derivatives. Normal glucose levels are between 3.3 and 5.6 mmol/L, and it is strictly regulated by endocrine system. Diabetes mellitus is a disorder where this

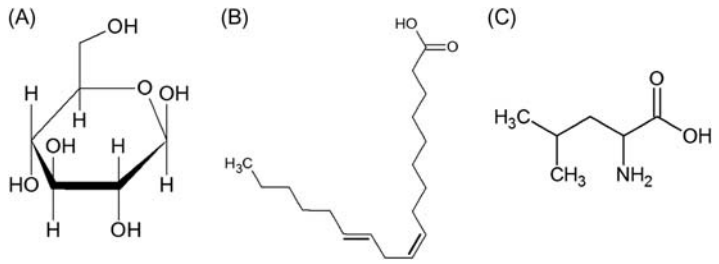


Figure 10.2
Molecular structural figures as examples of (A) sugar, (B) lipid, and (C) amino acid.

regulation has failed, either due to insulin resistance or nonproduction, and can be identified by the glucose concentration in the blood. Hyperglycemia is defined as above the normal range, whereas hypoglycemia is below.

Lipid concentrations also are indicators for pathological conditions and are therefore routinely monitored. A healthy range of total cholesterol is between 100 and 199 mg/dL in adults [16], while triglycerides are routinely below 150 mg/dL; both are linked to risk of heart and blood vessel disease and can be used to screen patients to prevent severe conditions leading to myocardial infarction (MI).

Amino acids, the building blocks of peptides and proteins, are found at levels of 2300−4000 μmol/L and are produced from protein catabolism. The most abundant amino acids, glutamine and alanine, are found at concentrations of 600 and 300 μmol/L, respectively [16]. Amino acids metabolism is complex and depends on a thin balance between protein anabolism and catabolism. Any dysfunction in a metabolic pathway leads to hyperammonemia (>400 μmol/L).

Nitrogen metabolites are also used routinely in clinics to witness body dysfunctions. Ammonia (6−35 μmol/L) is a highly toxic substance and thus is transported in blood mostly in the form of glutamine. An increased level of ammonia can indicate hepatic dysfunction or hypercatabolism. Urea is the end product of protein (or amino acid) degradation and is formed in the liver. Increased blood levels above 2.5−8.3 mmol/L generally indicate elimination dysfunction, and thus can highlight renal failure.

10.1.3.2 Proteins

In total, more than 20,000 human proteins are known, of which more than 1,500 are extracellular [19]. Blood-borne proteins form a heterogeneous group comprising more than 300 types of proteins. The total concentration of serum proteins (known as serum total protein) ranges from 60 to 80 g/L, and is dependent on many factors, such as the rate of synthesis and degradation, distribution in body fluids, hydration, and elimination. Physiologically higher concentrations can be observed in the standing position and in association with increased muscle activity. On the other hand, lower concentrations are physiological for children, pregnant women, or after prolonged fasting [20]. Some of the abundant proteins like fibrinogen and clotting factors are only found in plasma—about 8% of total proteins—while albumin and globulins can be measured in both plasma and serum. Albumin is the most abundant protein, accounting for about 55% of total proteins, which, combined with globulins, represents about 90% of the overall protein content of serum in healthy patients (Table 10.1). The dynamic range of concentrations in plasma and serum can be one of the greatest challenges in analyzing the proteome, requiring advanced and intensive methods such as mass spectrometry−based protocols. However, for a large

Table 10.1: Relative content of abundant proteins in plasma and serum.

Sample type	Protein type	Function	Abundance (%)
Plasma	Fibrinogen	Coagulation	7
	Clotting factors	Fibrin formation from fibrinogen	<1
Serum and *Plasma*	Albumin	Blood vessel integrity Carrier for insoluble molecules Tissue growth and healing	55–65
	α-1 Globulin	High-density lipoprotein (HDL)	1–5
	α-2 Globulin	Haptoglobulin binds hemoglobin (iron)	5–11
	β Globulin	Carrier and part of defense system against infection	7–13
	γ Globulin	Antibodies—immune system	10–18

number of clinically relevant protein biomarkers, routine analysis is required to enable early disease diagnosis.

Fibrinogen

The main difference between blood plasma and serum is the removal of clotting factors, predominantly fibrinogen, to produce the latter. Fibrinogen is a dimeric 340 kDa (0.4% in human plasma) plasma glycoprotein synthesized by the liver and plays a major role in blood coagulation. When blood clotting is activated, the circulating fibrinogen turns into fibrin and a stable clot is formed [21]. The normal concentration of fibrinogen in the human body is ~3 mg/mL, and any variation in this concentration can be an indicator of disease states. Many clinical studies have consistently shown elevated levels of fibrinogen in patients with cardiovascular disease and thrombosis [22–24]. A study conducted by Aleman et al. has shown that both elevated circulating fibrinogen (hyperfibrinogenemia) and abnormal fibrinogen levels are observable in plasma from patients with venous thrombosis [25]. However, the study undertaken by Klovaite et al. indicated that elevated plasma fibrinogen levels are associated with increased risk of pulmonary embolism rather than deep venous thrombosis [26]. Another study conducted by Toss et al. showed that increased fibrinogen levels are associated with persistent *Chlamydia pneumoniae* infection in unstable coronary artery disease [27]. A study conducted by the Emerging Risk Factors Collaboration, United Kingdom, also predicted a positive correlation between C-reactive protein (CRP), fibrinogen, and cardiovascular disease [28]. These studies and others suggest a strong correlation between increased levels of plasma fibrinogen and heart diseases that could be used as a diagnostic indicator [29,30].

Albumin

In humans, albumin is the most abundant plasma protein, normally constituting about 50% of human plasma protein, and has a molecular weight of 66 kDa. Albumin is a protein made by the liver and its main role is to maintain the osmotic pressure of the blood compartment,

provide nourishment of the tissues, and transport hormones, vitamins, drugs, and other substances such as calcium throughout the body [31]. The normal concentration of albumin in the human body is 30 g/L. Acute dehydration is the only clinical situation that is found to cause an increase in albumin concentration [8]. In the event of critical illness, the rates of synthesis and degradation of albumin are altered, leading to an abnormal distribution of albumin between the intravascular and extravascular compartments. The concentration of albumin decreases dramatically in critically ill patients and does not increase until the recovery phase of the illness [32]. Increased capillary leakage is the main reason for the altered distribution of albumin in critical illness. This is reported to occur in sepsis and after major surgical stress [33,34]. Cirrhotic patients are highly prone to suffer from septic shock. Several studies have demonstrated that the functions of albumin, such as ligand binding and transport of various molecules, can be applied in the treatment of cirrhotic patients and patients suffering from other end-stage liver diseases [35]. It is clear that closely monitoring the variation in albumin concentration could act as an indicator of liver diseases and other related pathologies.

Globulins

Globulins represent fractions of proteins identified in the plasma/serum known as alpha-1 globulins, alpha-2 globulins, beta globulins, and gamma globulins (Table 10.2). Each one of these fractions have various biological functions, including immune response.

Immunoglobulins

Immunoglobulins (Ig) are gamma globulin proteins present in blood plasma and mucosal secretions, and are delivered to sites of inflammation within tissue. They are antibodies produced by B lymphocytes—white blood cells —and are amongst the most abundant proteins in the blood, comprising 20% (by weight) of total plasma/serum proteins

Table 10.2: Globulin fractions with example of encompassed proteins.

Alpha-1 globulins	Alpha-2 globulins	Beta globulins	Gamma globulins
Range: 1–4 g/L	**Range: 3–9 g/L**	**Range: 7–15 g/L**	**Range: 8–16 g/L**
α-1 Antitrypsin	Very low-density lipoprotein (VLDL)	Transferrin	Immunoglobulins (Ig)
α-1 Acid glycoprotein (orosomucoid)	Haptoglobin	Low-density lipoprotein (LDL)	C-reactive protein (CRP)
α-1 Antichymotrypsin	Alpha-2 macroglobulin	Complement protein C3	Lysozyme
High-density lipoprotein (HDL)	Ceruloplasmin	Fibrinogen	Alpha fetoprotein
Thyroxine-binding globulin (TBG)	Glycoprotein	Hemopexin	
Prothrombin	Complement protein C4	Properdin	

(Table 10.1). Based on structure and protein composition, Ig are divided into five classes: IgG, IgA, IgM, IgD, and IgE (Table 10.3).

Ig classes vary by the number of units: one unit (IgG, IgD, and IgE), two units (IgA), or five units (IgM). The simplest antibody molecule can be represented as a "Y" shape (Fig. 10.3), composed of four polypeptide chains: two identical heavy chains (H) and two identical light chains (L) oriented parallel to each other and linked by disulfide bonds [36]. Both light and heavy chains are characterized by variable and constant regions. L chains

Table 10.3: Details of immunoglobulin (Ig) subtypes.

	IgM	IgG	IgA	IgE	IgD
Molecular weight	900 kDa	150 kDa	320 kDa	200 kDa	180 kDa
Normal concentrations	0.5–2 g/L	10–16 g/L	1–4 g/L	10–400 μg/L	0–0.4 g/L
Function	Secondary immune response: Naive response	Secondary immune response: immune memory	Local immune response on exposed mucosal surface	Immune response to parasites	Antibody production regulation
Abnormal situations		Myeloma ↑ Immunodeficiency ↓		Allergic reactions ↑	

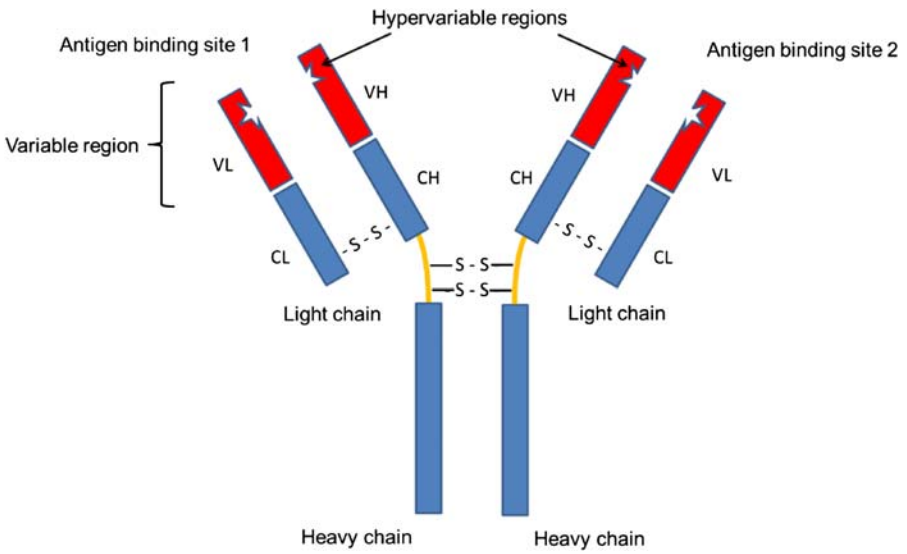


Figure 10.3
Generic structure of immunoglobulins (Ig).

only consist of one variable (VL = variable light) and one constant (CL = constant light) domain, while heavy chains have one VH (variable heavy) and three to four CH (constant heavy) domains, respectively, corresponding to IgG/IgA, and IgM/IgE. The terminal end of both L and H chains present extremely variable ("hypervariable") regions forming the antigen binding site. There are two types of L chains: λ (lambda) or κ (kappa), that can be found in all types of Ig, but never together. H chains are Ig specific and designated γ (IgG), α (IgA), μ (IgM), ε (IgE), and δ (IgD). Each heavy chain has about twice the number of amino acids and molecular weight (∼50 kDa) compared to light chains (∼25 kDa), resulting in a total Ig monomer molecular weight of approximately 150 kDa. The constant region confers Ig its biologic activity while the variable regions form a complex, conformational molecular arrangement for the attachment of each specific antigen.

10.1.4 Pathology of plasma proteins

10.1.4.1 Abundant proteins

Generally, routine serological evaluation of a patient includes analysis of a panel of serum electrolytes, total protein, and glucose concentrations. Estimating total protein, globulin, and albumin content is important to assess the global and nutritional status of patients [8,37]. The Biuret assay is the most common method used to quantify total protein levels in blood serum [38], which is the most compatible with routine application in terms of sensitivity and linearity of range, as well as in terms of reaction time. Apart from a few known physiological states, a low concentration is a sign of undernutrition, malabsorption, or hepatic disease and occurs in the case of severe protein loss—burn victims and renal failure for example. On the other hand, hyperproteinemia is found in conditions such as dehydration, myeloma, and systemic lupus erythematosus.

Although total serum protein estimation has limited diagnostic potential when compared to albumin or globulin due to lack of specificity, its relevance in the evaluation of patients with clinical conditions such as malnutrition, renal malfunction, liver diseases, and immune disorders cannot be ignored [37,39]. A normal level of total serum protein of 60−83 g/L indicates healthy nutritional status and normal liver function. Reduced serum total protein is predominantly found in patients with kidney disorders, HIV, and the elderly [40,41].

γ-Globulins, produced by lymphocytes and plasma cells in lymphoid tissue, are large protein molecules that consists of the Ig: IgM, IgA, and IgG [42]. An elevation in γ-globulins is a characteristic abnormality in the serum proteins in liver diseases and carcinoma of the gastrointestinal tract or breast [43,44]. Testing globulin levels in serum routinely provides key information that helps diagnose various conditions and diseases that affect the immune status. Liver diseases, chronic inflammatory diseases, hematological disorders, infections, and malignancies cause excess Ig levels [45], whereas humoral

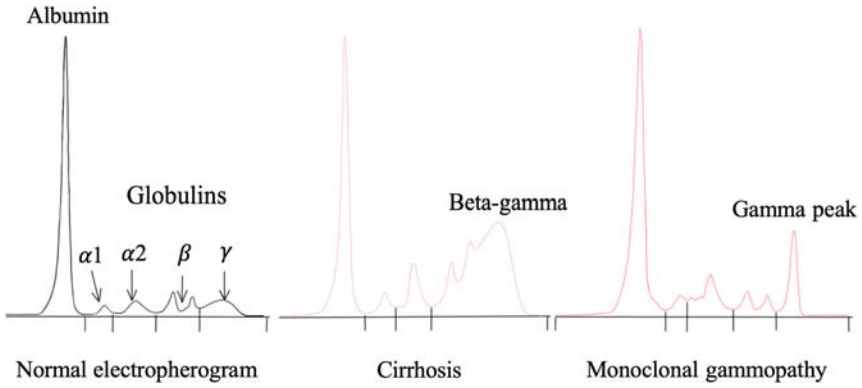


Figure 10.4
Disease patterns in serum protein electrophoresis (SPEP).

immunodeficiencies cause low Ig levels [46]. Radial immunodiffusion is the gold standard method for measuring globulins [47]. All conventional methods used for testing total protein content and globulin measurement make use of expensive disposables and are labor intensive. With escalating medical costs and budget constraints, a cost-effective alternative technology is desirable.

Serum protein electrophoresis (SPEP) is a cost-effective method for separation of proteins, based on their net charge (positive or negative), size, and shape. It enables visualization of the two major types of protein present in the serum: albumin and the globulin proteins. The largest peak, closest to the positive electrode, reflects the high concentration of albumin, while globulins are represented by multiple smaller features (Fig. 10.4). However, the pattern of five globulin categories (alpha-1, alpha-2, beta-1, beta-2, and gamma) contains the most relevant information for diagnosis. A dense narrow band that is composed of a single class of Ig is categorized as monoclonal gammopathy. It is the result of a malignant clone producing only one type of antibody that leads to a thin peak in protein electrophoresis. A broad-based band in the gamma region suggests a polyclonal increase in Ig (polyclonal gammopathy). When the β-globulins and γ-globulins do not separate, it can be specific to liver disease (cirrhosis), but is also common in autoimmune disease, chronic viral or bacterial infections. SPEP is a rapid technique to detect a number of conditions based on qualitative and quantitative patterns of the serum fractions. Table 10.4 provides examples of conditions with increased or decreased levels in protein fractions.

10.1.4.2 Low-abundance proteins

Currently, there are various proteins that are often used in the diagnosis and monitoring of different pathologies. For example, hepatic function is evaluated by aspartate amino transferase, alanine amino transferase, alkaline phosphatase, and gamma glutamyl

Table 10.4: Serum protein fractions and conditions associated with an increased or decreased level.

Serum protein		Decreased	Increased
Albumin		Malnutrition Cachexia (wasting syndrome) Liver disease Nephrotic syndrome Impaired liver function Protein-losing enteropathies Hemorrhage Severe burns	Indicator for dehydration
Globulins	Alpha-1	Alpha-1 antitrypsin deficiency Nephrotic syndrome Liver dysfunction	Pregnancy Inflammatory states
	Alpha-2	Hemolysis Liver disease	Nephritic syndrome Adrenal insufficiency Adenocorticosteroid therapy Advanced diabetes mellitus Hyperthyroidism
	Beta-1–2	Protein malnutrition	Biliary cirrhosis Hypothyroidism Nephrosis Polyarteritis nodosa Obstructive jaundice Cushing's disease Third-trimester pregnancy Iron-deficiency anemia
	Gamma	Agammaglobulinemia and hypogammaglobulinemia	Cirrhosis Multiple myeloma Hodgkin's disease Chronic lymphocytic Leukemia Amyloidosis Rheumatoid

transferase activities. For cardiac injuries, such as acute coronary syndrome, diagnostic and patient care is dependent on the troponin test. Recently, a number of studies have proven that imbalances in plasma protein levels can be linked to the presence of numerous disease states [48,49]. A method incorporating the ability of polyethylene glycol fractionation and immunoaffinity depletion to detect plasma biomarkers has been reported [50]. This method successfully identified 135 low-abundance proteins with concentration levels less than 100 ng/mL. A high-accuracy mass spectrometry–based proteomics method has been reported to characterize proteins in the plasma of patients with an acute bone fracture, leading to the discovery of several new proteins which were not previously reported in plasma [51]. Addonna et al. developed a pipeline by integrating the proteomic technologies

used from the various stages of discovery of plasma biomarker identification, to recognize early biomarkers of cardiac injury [52]. Patients were allowed to be their own controls by sampling blood directly from patient hearts before, during, and after controlled myocardial injury. Liquid chromatography mass spectroscopy detected 121 highly differentially expressed proteins and >100 novel candidate biomarkers for myocardial infarction (MI) [53]. Ray et al. identified 18 signaling proteins in blood plasma that can be used to differentiate Alzheimer's samples and control subjects with close to 90% accuracy and to identify patients who had mild cognitive impairment that progressed to Alzheimer's disease (AD) 2–6 years later. This molecular test for AD could lead to early diagnosis and better treatment [54].

10.1.4.3 Cytochrome c

Cytochrome *c* is a water soluble, ~12 kDa heme protein found loosely attached to the inner membrane of the mitochondrion. This heme protein normally resides in the mitochondrion and is released into the blood in the event of cell death, triggering inflammation [55]. It is essential in mitochondrial electron transport and also acts as an intermediate in apoptosis [56,57]. More recently, it was reported that cytochrome *c* can be used as an indicator of the apoptotic process in the cell [58]. This study demonstrated that it can act as an in vivo apoptosis indicator and prognostic marker during cancer therapy using a cytochrome *c* enzyme-linked immune sorbent assay (ELISA) kit that was modified to increase sensitivity. In another study, a sandwich ELISA method was used to measure serum cytochrome *c* levels to quantify the extent of apoptosis in systemic inflammatory response syndrome (SIRS) [59]. The prognostic significance of cytochrome *c* concentrations was investigated, and the ability of this method to assess the severity of organ dysfunction and to help to predict the prognosis of SIRS was demonstrated. Release of cytochrome *c* into circulation has been reported in patients with MI [60] and several liver diseases [61]. The mean cytochrome *c* level recorded in patients with liver diseases was found to be 187.1 ng/mL, whereas that of healthy controls was 39.8 ng/mL. A number of studies have been undertaken to show that cytochrome *c* can be used as a potential clinical marker of molecular and cellular damage [55]. Cytochrome *c* was identified in mitochondrial damage-associated molecular patterns, along with interleukin-6, as a marker of inflammation in hemodialysis patients [62]. An elevated level of cytochrome *c* in human plasma/serum can be indicative of the presence of various pathologies. Therefore, this mitochondrial protein has huge potential to be used as a clinical marker for these diseases at an early stage.

10.1.5 Vibrational spectroscopic analysis of bodily fluids

It is recognized that conventional test kits commonly employed in a hospital environment for plasma/serum analysis suffer from long time delays due to the need for specialized laboratories, which may in turn delay therapy, and prolong patient anxiety [1].

The development of optical methods for biomedical applications is an emerging field with huge potential [63], and has recently been explored through vibrational spectroscopic approaches [64,65]. The sensitivity to subtle changes in biochemical composition makes vibrational spectroscopy an ideal diagnostic tool. Considering the advancement in spectroscopic technologies and data analysis capabilities, coupled with filtration and fractionation techniques, bodily fluids can be analyzed rapidly and non-invasively to detect disease-related fluctuations in protein concentration [66–69].

10.1.6 Vibrational spectroscopy

Vibrational spectroscopy usually refers specifically to the optical techniques of IR absorption and Raman scattering spectroscopy, as well as inelastic neutron scattering. It is a subset of spectroscopy which analyzes vibrations within a molecule (or material). The vibrations are characteristic of the molecular structure and, in polyatomic molecules, give rise to a spectroscopic "fingerprint". Thus, the spectrum of vibrational energies can be employed to characterize a molecular structure, or alterations due to the local environment or external factors (e.g., radiation and chemical agents).

The number of vibrational modes for a given molecule will depend on its structure. A molecule with N number of atoms will have $3N$ degrees of freedom. Generally, non-linear molecules will exhibit $3N$-6 vibrational modes, the six nonvibrational degrees of freedom corresponding to three translational and three rotational modes around the x, y, and z axes. In contrast, as linear molecules are unable to rotate upon their axis, one of the rotational degrees of freedom is lost, and hence they can be described as having $3N$-5 vibrational modes [70]. Bond stretching and bending are the two fundamental types of molecular vibration; symmetric or asymmetric stretching alters the bond length, whilst bending vibrations consist of changing the bond angle, through twisting, rocking, wagging, and scissoring (Fig. 10.5).

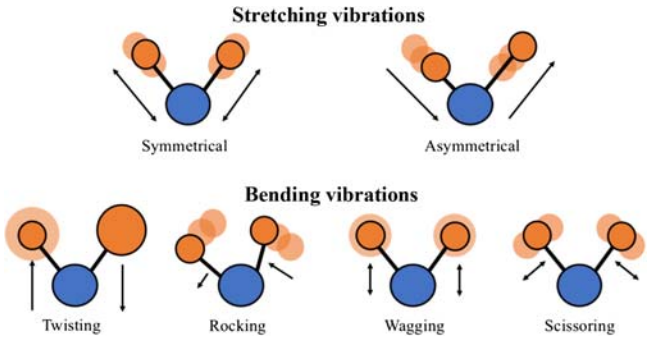


Figure 10.5
Common vibrational modes of chemical bonds.

Vibrational energies fall within the mid-infrared (mid-IR) region of the electromagnetic spectrum and are commonly probed through IR absorption spectroscopy. Following the discovery of IR radiation by Herschel in 1800 [71], initial applications of IR absorption spectroscopy were limited to astronomy and astrophysics [72]. In material sciences, significant advances were made by 1900 when Abney and Festing recorded spectra for 52 compounds, correlating absorption bands with molecular structures [73]. Coblentz helped establish IR spectroscopy as a routine analytical tool, cataloging the spectra of hundreds of substances, both organic and inorganic. Technological developments after the second world war aided considerably in establishing IR spectroscopy as a routine laboratory characterization technique, but none more so than the development of commercial Fourier-transform infrared (FTIR) spectrometers in the 1960s and 1970s [74,75], and FTIR microscopes in the late 1980s [76].

Raman spectroscopy is a complementary technique with its origin in the discovery of the Raman effect in 1928 [77], for which C.V. Raman was awarded the Nobel prize in physics in 1930. In 1998 the Raman effect was designated an ACS National Historic Chemical Landmark, in recognition of its importance in materials and process analysis. Raman spectroscopy remained largely a curiosity until the advent of the laser in the 1960s, and the revolution in charged coupled detector (CCD) arrays in the 1980s and 1990s added to the benefits of high laser source intensities. In addition, the development of narrow band laser line rejection filters meant that the huge losses in signal from traditional triple monochromator systems could be overcome with the combination of a filter set and a single spectroscopic grating. Furthermore, the significant reductions in acquisition time with multichannel signal detection enabled significant improvements in signal-to-noise ratios (SNRs) [78]. The combination of technological developments led to a new range of Raman spectroscopic microscopes in the 1990s, establishing Raman spectroscopic microscopy as a relatively inexpensive benchtop laboratory tool to complement conventional IR spectroscopy.

Both IR and Raman spectroscopy entail the coupling of incident radiation with molecular vibrations and the resultant spectrum is characteristic of the compound or material. However, whereas IR spectroscopy involves the absorption of radiation, inducing transitions between vibrational states, Raman spectroscopy is a scattering technique (Fig. 10.6), whereby the incident radiation couples with the vibrating polarization of the molecule and thus generates or annihilates a vibration. The differing underlying mechanisms give rise to a complementarity of the two techniques. For a vibration to be active in IR spectroscopy, a change in dipole is required, whereas to be Raman active, a change in polarizability is required. As a rule of thumb, vibrations of asymmetric polar bonds tend to be strong in IR spectra, whereas Raman is particularly suitable as a probe of symmetric, nonpolar groups. Notably, O−H vibrations of water are very strong

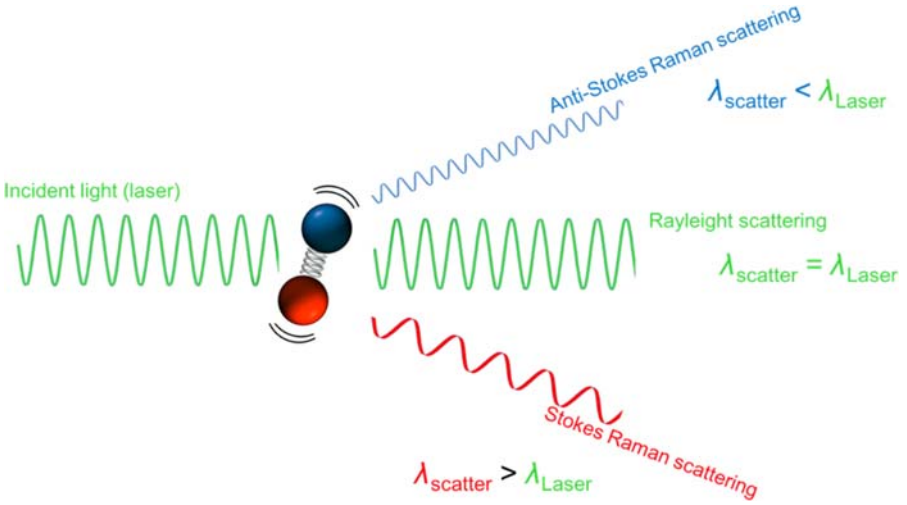


Figure 10.6
Depiction of light scattering by vibrating polarization.

in IR spectra, whereas they are extremely weak in Raman spectra, rendering Raman spectroscopy more suitable for some biomedical applications, particularly in vivo.

A further implication of the differing physical origins of the techniques is that, whereas IR directly monitors the absorption of IR radiation, Raman scattering can be employed in the ultraviolet, visible, or near-infrared (NIR) regions of the spectrum. Thus, Raman scattering offers intrinsically higher spatial resolution for mapping or profiling, the diffraction limit being determined by the wavelength (<1 μm for Raman, $\sim 5-10$ μm for IR). For many applications, however, near IR is favored as a source for Raman analysis, to minimize interference from scattering, fluorescence, or photodegradation of the sample [79].

10.1.7 Experimental approaches

10.1.7.1 Fourier-transform infrared spectroscopy

FTIR spectrometers have replaced traditional dispersive instruments, due to their superior speed and sensitivity [80]. They utilize a Michelson interferometer, which is a device that splits a single beam of IR light into two paths, and then recombines them after a variable path difference has been introduced (Fig. 10.7) [81,82]. The interferometers are composed of a fixed mirror, a movable mirror, and a beam splitter. The purpose of the beam splitter is to reflect some of the radiation toward the fixed mirror, meanwhile partially transmitting the rest to the adjustable mirror. When the waves return to the beam splitter, they interact and are then further reflected and transmitted. The split beams travel different path lengths as a result of the moving mirror, and hence produce waves of different of intensity when

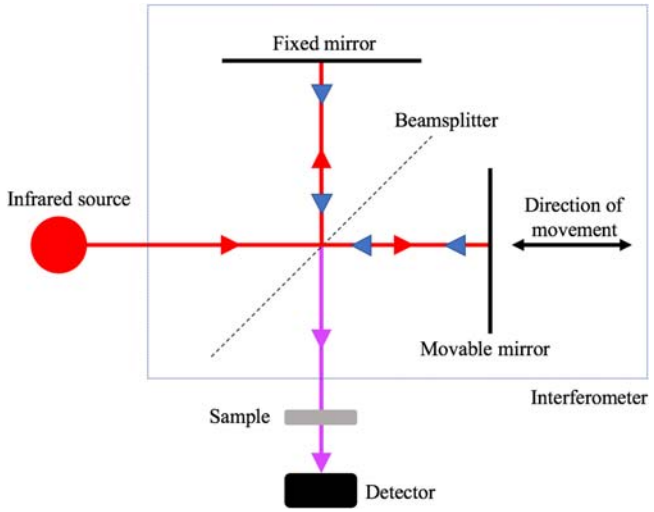


Figure 10.7
The Michelson interferometer found in FTIR instruments. Red: incident beam; blue: reflected; purple: combined. Source: *Adapted from P.R. Griffiths, J.A. De Haseth, Fourier Transform Infrared Spectrometry, second ed., Wiley-Interscience, Hoboken, NJ, 2007.*

recombined [83]. As a function of time, the light field varies spectrally, and can be converted to the frequency domain through Fourier transformation. The absorption, reflection or scattering of light by a sample can thus be recorded in the time domain by a single detector. FTIR absorption spectroscopy monitors the vibrational bending and stretching modes of molecules that are active within the IR region. The wavelengths at which they absorb the IR radiation are measured, and as every compound has a characteristic set of absorption bands, it results in a unique spectroscopic fingerprint [84].

There are three main sampling modes involved in FTIR spectroscopy: transmission, reflection, and attenuated total reflection. By default, most FTIR instruments use transmission mode, in which traditionally IR light is irradiated through a sample on an IR transparent window, such as calcium fluoride, and is collected by a detector on the other side [85]. The coupling of transmission mode to microscopy has allowed FTIR imaging to emerge in biomedical research [86,87]. There are, however, a number of flaws related to transmission mode. Sample preparation can be exhaustive, and short path lengths ($<10\,\mu m$) are required to prevent full absorption of IR radiation by the sample before reaching the detector. This limiting factor also affects aqueous samples, since water is highly IR active [88]. Furthermore, IR transparent substrates that are required for this technique are fragile and often rather expensive to replace [89].

In transmission/reflection, or so called transflection mode, the incident IR beam initially travels through a sample, and is then reflected back off an IR reflecting substrate, and again

passes through the sample toward the detector. It can be advantageous in comparison to transmission, in that the substrates are generally inexpensive low-emissivity (low e) slides, and the approximate sample thickness can usually be smaller than that required for transmission measurements (1−4 μm cf. 2−8 μm) which can be beneficial when sample quantity is limited. On the other hand, as the path length is effectively doubled there is also a maximum thickness limitation. Transflection mode may also be prone to standing wave artefacts that cause spectral variance, although the implications of this effect for diagnostic applications are still being assessed [85,90,91].

In assessing the suitability of the measurement mode for analysis of biological samples which are highly physically and chemically inhomogeneous, it is important to understand the physical processes involved. When a sample is measured in reflection or transflection mode, a proportion of the light registered derives from the top surface, the reflectance of which is governed by the real component of the refractive index of the material. The transmitted light, measured in transmission or transflectance, can be reduced by the intrinsic absorptions of the constituent molecules, giving rise to the desired fingerprint of the sample, but can also be reduced by "Mie-like" scattering from structures (cells and cell nuclei) which have dimensions similar to the wavelength of light employed (5−20 μm). This scattering is resonantly enhanced in the neighborhood of an absorption, and can give rise to spectral "artefacts" in reflection, transflection, and transmission modes [92,93]. These resonant Mie effects can be ameliorated by the application of specific preprocessing methods [94].

The development of attenuated total reflection−Fourier-transform infrared (ATR-FTIR) spectroscopy has attracted wide interest in the field in recent years. The technique is unique in that the incident IR beam does not actually travel through the sample, but is directed through a substrate with a high refractive index, such as diamond, germanium, or silicon, known as an internal reflection element (IRE). The sample must be placed in direct contact with the IRE, as when the incident radiation reflects off the internal surface of the IRE, an evanescent wave projects orthogonally into the sample, which then attenuates the IR beam before exiting the IRE to the detector [88]. The refractive index of the chosen IRE and the sample govern the basic attenuated total reflection (ATR) phenomenon, as shown in Eq. (10.1), where the critical angle θ_c, can be calculated from n_1 and n_2, which are the refractive indices of the IRE and sample, respectively.

$$\theta_c = sin^{-1}\left(\frac{n_2}{n_1}\right) \tag{10.1}$$

The IR radiation undergoes total internal reflection when the angle of incidence at the sample−crystal interface is greater than the critical angle, hence materials with a high refractive index are commonly chosen to minimize the critical angle [95]. An important

factor is the depth of penetration d_p of the evanescent wave into the sample, as it determines how much of the sample is actually analyzed—described in Eq. (10.2), where θ is the angle of incidence and λ is wavelength [81]. The penetration depth is dependent upon the angle of incidence, refractive indices, and the wavelength; at longer wavelengths, the evanescent wave will penetrate deeper into the sample.

$$d_p = \frac{\lambda}{2\pi n_1 \sqrt{sin^2\theta - (n_2/n_1)^2}} \tag{10.2}$$

One limitation with this approach is that the IRE must be kept clean to ensure there is no cross-contamination between samples, inhibiting throughput. Traditional IREs can also be expensive, such as a fixed diamond crystal. However, multi-IRE systems have been developed that will enable higher throughput analysis [96]. Scratches on the surface of the IRE are known to affect the sample-IRE contact, and loss of sensitivity is common due to the shallow penetration depths [82]. That said, ATR has become extremely popular in FTIR spectroscopy, as it has numerous advantages over the other IR techniques [81]. In contrast to transmission mode, for which the sample usually has to be pressed into a pellet or thin film, the ATR-FTIR mode negates the need for time-consuming preparation, as the sample can be examined directly on the IRE, in liquid or solid state. The shorter path length makes it more applicable for aqueous samples, as less IR radiation is lost through water absorbance, compared to transmission measurements [88]. Likewise, minimal scattering effects and relatively high SNR are valuable attributes [85]. Biological samples, such as human blood serum, are well suited to ATR analysis as only small volumes of biofluid drops are required to dry efficiently onto the IRE. The size of the crystal governs the volume of sample that is required, and to ensure intimate contact occurs the sample should cover the whole IRE, allowing effective penetration of the evanescent wave [97].

10.1.7.2 Instrumentation for Raman spectroscopy

A Raman spectrometer typically consists of three major components: an excitation source, a sampling apparatus, and a detector. While these three components have evolved in varying forms over the years, modern Raman instrumentation has developed around using a laser as an excitation source, a spectrometer for the detector, and either a microscope or a fiber optic probe for the sampling apparatus. Fig. 10.8 shows a schematic of a typical Raman setup. The laser sources provide a stable and intense beam of radiation. A wide range of lasers can be used as the light source, although, for biological applications, longer wavelength, near-IR sources are commonly employed to minimize photodamage, scattering, and/or fluorescence [98]. Band pass "interference" filters are employed to clean the laser spectrum and remove plasma lines. Dispersive instruments make use of a notch filter coupled with a high-quality grating monochromator. Double, or triple, grating

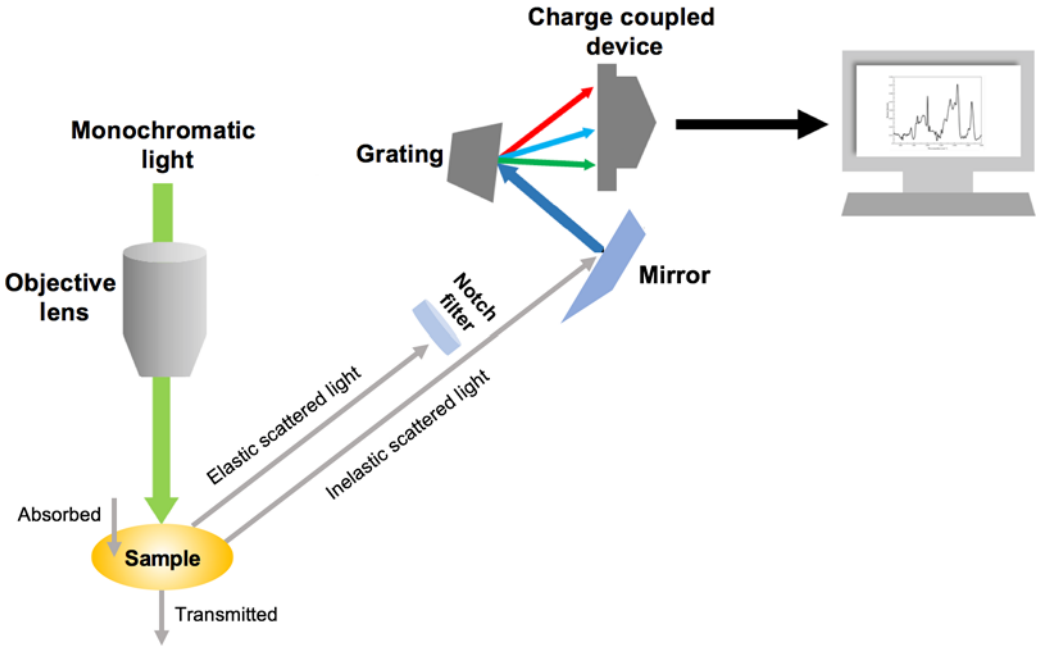


Figure 10.8
Typical instrumentation for Raman microspectroscopy.

monochromators, rejection filters, super notch filters, holographic notch or edge filters, and holographic filters are employed to separate relatively weak Raman lines from intense Rayleigh-scattered radiations [99]. Charge transfer devices (CTDs), such as charge-coupled devices and charge-injection devices, act as detectors and are used in the form of arrays. The role of the CTD arrays are to convert the incoming optical signal into charge which are then integrated and transferred to readout devices [100]. CTDs are commonly made of silicon so laser wavelengths of less than 1 μm can be detected, while laser wavelengths of greater than 1 μm use single element detectors based on a low bandgap semiconductor, such as germanium (Ge) or Indium–Gallium–Arsenic (InGaAs). The grating is used to disperse the light and the groove density determines the spectral resolution. Other factors that play a key role in determining spectral resolution include the wavelength, with shorter wavelengths having a higher spectral resolution, and the spectrometer length, which is the distance between grating and the detector, where longer distances provide higher resolution. The objective lens both delivers the incident light and collects the scattered light. The objective lens delivers the incident light to the sample, and the Raman scattered light can be observed at any angle. In the commonly employed backscattering geometry, the Raman is collected by the objective lens and delivered to the grating. The spectrally dispersed, detected Raman scattered light is displayed as a Raman shift from the source wavelength,

which is converted to units of wave number (cm^{-1}), such that Raman spectra can easily be compared and contrasted with equivalent FTIR spectra.

10.2 Biospectroscopy

FTIR spectroscopy has become an accepted tool in biophysics for analysis of the structure and interactions of biomolecular components [101–104]. Applications of FTIR spectroscopy for analyzing tissue samples for diagnostic purposes were first reported in the early 1990s, and since then a range of pathologies has been investigated [105–107]. The application of Raman spectroscopy to biomolecules and even tissues was first demonstrated as early as the 1960s, and by the mid-1970s biomedical applications were explored [108–110]. Whole cell and tissue studies have been carried out on a range of pathologies [111–113] and in vivo studies [114,115] have demonstrated the prospective for diagnostic applications. The potential of vibrational spectroscopy in conjunction with multivariate analysis techniques as a diagnostic tool has thus been well demonstrated and the concept of spectral cytopathology has been coined [116]. In this respect, Raman and infrared can be viewed as rival technologies, but to best advance the understanding of the potential of the techniques a combination of the two complementary techniques is recommended.

Lipids, proteins, nucleic acids, and carbohydrates are the four biomolecular groups characteristically found in a biological spectrum, as measured using either FTIR or Raman spectroscopy. Fig. 10.9 shows, for example, the (A) FTIR and (B) Raman microscopic spectra of human blood serum, where the protein-related bands are highlighted, while Tables 10.5 and 10.6 show typical band assignments across the full spectra. In the "high wavenumber region," $>2500\ cm^{-1}$, of the FTIR spectrum, the distinctive vibrations of

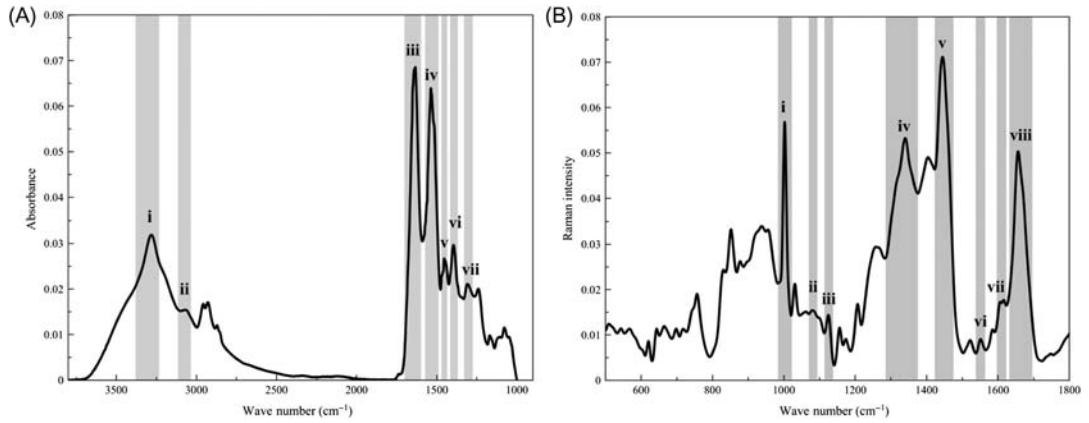


Figure 10.9
Typical (A) IR and (B) Raman spectrum of human blood serum.

Table 10.5: Tentative peak assignments for Fourier-transform infrared (FTIR) spectral data, (i)–(vii) corresponds to Fig. 10.9A [117–120].

Approximate Wave numbers (cm^{-1})	Vibration	Biochemical assignments
3300 **(i)**	ν(N–H)	Amide A of proteins/peptides
3100 **(ii)**	ν(N–H)	Amide B of proteins/peptides
2957	$\nu_{as}(CH_3)$	Lipids
2920	$\nu_{as}(CH_2)$	
2872	$\nu_{s}(CH_2)$	
2850	$\nu_{s}(CH_2)$	
1740	ν(C = O)	Phospholipid esters
1715–1680	ν(C = O)	Nucleic acids
1650 **(iii)**	>75% ν(C = O), ν(C–N), δ(N–H)	Amide I of proteins
1645	γ(HOH)	Water
1550 **(iv)**	~60% δ(N–H), ν(C–N), δ(C–O), and ν(C–C)	Amide II of proteins
1453	$\gamma(CH_2)$	CH_2 scissoring
1450 **(v)**	$\delta_{as}(CH_3)$	Lipid/proteins
1395 **(vi)**	$\delta_{s}(CH_3)$	lipid/proteins
1395	ν(C = O)	Carboxylate COO^-
1380	$\gamma_{s}(CH_3)$	Phospholipid/triglyceride
1350–1250 **(vii)**	δ(N–H), ν(C–N), γ(C = O), and ν(C–C)	Amide III–peptide/protein/collagen
1242	$\nu_{as}(PO^{2-})$	DNA/RNA/phospholipid
1170	ν_{as}(C–O)	Ester
1150	ν(C–O), γ(COH)	Carbohydrates
1090	$\nu_{s}(PO^{2-})$	DNA/RNA/phospholipid
1086	ν(C–O), ν(C–C), def(CHO)	Carbohydrates
1079	ν(C–C)	Glycogen
1065	ν(C–O)	DNA and RNA ribose
1050	ν(C–O)	Phosphate ester
1028	def(CHO)	Glycogen
965	$\nu(PO_3^{2-})$	DNA and RNA ribose
710–620	def(O = C–N)	Amide IV

ν, stretching; δ, bending; γ, wagging, twisting and rocking; *def*, deformation; *as*, asymmetric; *s*, symmetric.

N–H, C–H, and O–H of lipids and proteins can be found, whereas in the "fingerprint region," $<1800\ cm^{-1}$, the features are typically more complex combinations, including the amide I ($1650\ cm^{-1}$) and amide II ($1520\ cm^{-1}$) modes of proteins, nucleic acid phosphate stretching modes at 1070 and $1250\ cm^{-1}$, and lipid-derived features at 1310 and $1750\ cm^{-1}$. It should be noted that, although complementary techniques, the features in the respective spectra of FTIR and Raman have similar origin. Thus, the Raman spectrum of the nucleus exhibits similarly prominent signatures associated with proteins and lipids across the fingerprint region, as well as large peaks related to DNA and RNA at $785\ cm^{-1}$.

Table 10.6: Tentative peak assignments for Raman spectral data, (i)–(viii) corresponds to Fig. 10.9B [121,122].

Approximate Wave numbers (cm^{-1})	Vibration	Biochemical assignments
785–788	ν_s (PO^{2-})	Nucleic acid
1004 **(i)**	ν(ring breathing) Phenylalanine	Protein
1090 **(ii)**	ν_s (PO^{2-})	Nucleic acid
	ν(C–N)	Protein
1127 **(iii)**	ν(C–N)	Protein
	ν(C–C)	Lipid
1262	ν_s(PO^{2-})	Nucleic acid
	δ(N–H), ν(C–H)	Amide III of proteins
1319 **(iv)**	$\gamma(CH_2)$	Lipid/nucleic acid
	def(CHO)	Protein
1341 **(iv)**	ν_s (PO^{2-})	Nucleic acid
	def(C–H)	Protein/fatty acid
1451 **(v)**	$\delta(CH_2)$	Protein/lipid
1554 **(vi)**	δ(N–H), ν(C–N)	Amide II of proteins
1619 **(vii)**	ν(C = C)	Protein
1662 **(viii)**	ν(C = O)	Amide I of proteins
	ν(C = C)	Lipid

ν, Stretching; δ, bending, γ, twisting; *def*, deformation; *s*, symmetric.

A notable difference between the two spectra is the spectral range presented. FTIR covers the full spectral range in a single scan of the interferometer, which is recorded by a single detector. In contrast, Raman spectroscopy is currently predominantly performed by dispersive techniques, by which a spectral range is dispersed onto a CCD pixel array. The range covered depends on the dispersion element (grating) and use of higher resolution gratings means that multiple windows are required to cover the whole spectral range. Hence, Raman spectra are frequently presented in either the fingerprint or high wave number region. On the other hand, either Raman or FTIR can be limited in the lowerwavenumber range by instrument detectors, optical elements, and or substrate used. Typically, Raman spectra are recorded as low as 400 cm^{-1}, and FTIR only as low as 1000 or 600 cm^{-1}.

10.2.1 Vibrational spectroscopy of proteins

For the investigation of proteins, vibrational spectroscopy is particularly useful, as protein-related bands are dominant within biological spectra. Stretching vibrations are found in the higher-wave number region (3500–2500 cm^{-1}), such as C–H, N–H, and O–H stretches, whereas bending and carbon skeleton fingerprint vibrations tend to occur in the lower wave number regions. The most important spectral region in relation to biological materials is the

information-rich fingerprint region (1800−400 cm^{-1}), wherein the amide I and II peaks exist (1700−1500 cm^{-1}) [85].

The wealth of information that exists in a vibrational spectrum of a biological sample, detailed in Tables 10.5 and 10.6, renders the techniques as interesting tools for investigating molecular systems ranging from amino acids, peptides, and protein complexes [123−126]. Vibrational spectroscopy can enhance the understanding of protein function, as it is sensitive to changes to the protonation state of amino acid side chains and the strength of hydrogen bonding between amide bonds [127,128]. In both IR and Raman spectra, most characteristic bands are associated with the CONH group, referred to as amide A (NH stretching, ∼3300 cm^{-1}), amide B (NH stretching, ∼3100 cm^{-1}) and amide I to VII (I: 1600−1700 cm^{-1}, II: 1480−1580 cm^{-1}, III: 1230−1300 cm^{-1}, IV: 625−770 cm^{-1}, V: 640−800 cm^{-1}, VI: 540−600 cm^{-1}, VII: 200 cm^{-1}) (Fig. 10.10) [117,129,130]. The amide A band (∼3300 cm^{-1}) originates from the NH stretching vibration, which is often present as a resonance doublet with the weakly absorbing amide B (∼3170 cm^{-1}), arising from a

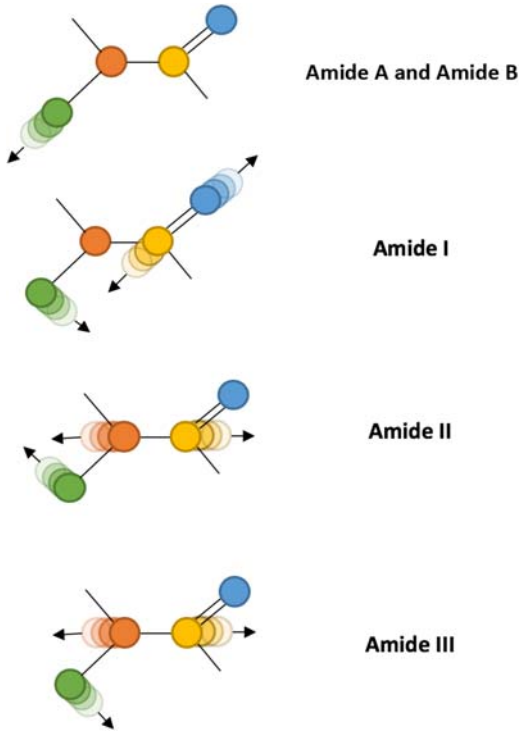


Figure 10.10
Molecular vibrations of the amide group—orange: hydrogen; red: nitrogen; purple: carbon; blue: oxygen. Source: *Adapted from K. Spalding, Developing Spectroscopic Biofluid Diagnostics: Monitoring and Therapeutic Profiling of Melanoma Patients, University of Strathclyde, 2018 [129].*

Fermi resonance between the first overtone of amide II [117]. The amide I, which absorbs near 1650 cm^{-1}, is primarily caused by the C = O stretching vibrations, with smaller contributions from CN stretching, deformation of CCN, and NH in-plane bending vibrations. The out-of-phase combination of the NH bending and the CN stretching vibrations, as well as minor contributions from the CO in-plane bend and the CC and NC stretching vibrations, give rise to the amide II band at $\sim$1550 cm^{-1} [131].

Similar analyses can be performed using IR and/or Raman spectroscopy, and even Raman optical activity, which is particularly sensitive to molecular chirality [132]. Although IR absorption and Raman scattering spectroscopy probe the same physical phenomenon of molecular vibrations, the spectral profile is discernibly richer in substructure in the case of Raman given similar instrumentational specifications in terms of spectral and spatial resolution [133]. IR absorption involves an electric dipole transition between two vibrational states, each of which has its own homogeneously and inhomogeneously broadened line width. The resultant spectral bandwidth is a convolution of these two individual line widths. Although often represented as a transition between a real vibrational level of the manifold of an electronic state and a virtual electronic level, Raman is a scattering process. In the representation of the transition to a virtual state, the bandwidth of that state is infinitesimally small, and so the scattering line width is intrinsically less than that of an equivalent IR absorption transition, giving rise to more distinct spectral features.

A characteristic feature of Raman spectra of many proteins, not observed in IR spectra, is the strong and often dominant feature at $\sim$1004 cm^{-1}, ascribed primarily to the ring breathing mode of the phenylalanine residue. Note, however, that its prominence in the spectrum does not reflect a similar relative prominence over other residues in the protein structure, but rather the large Raman scattering cross section of the highly polarizable, π-conjugated, ring structure. Similarly, the porphyrin moieties of cytochrome *c* contributes strongly to the Raman spectrum at $\sim$1585 cm^{-1}, and can be resonantly enhanced at source wavelengths of $<$550 nm (Fig. 10.11) [134].

Many studies have looked at the potential of vibrational spectroscopy to predict protein secondary structure, as the amide I band is highly sensitive to hydrogen bonding pattern, dipole–dipole interaction and the geometry of the polypeptide backbone [135]. A series of overlapping components that represent different structural elements, such as α-helices and β-sheets, are present in the broad amide I band [88,136].

Rygula et al. have reviewed the analysis of the secondary structure of proteins using Raman spectroscopy of 26 different proteins [130]. Raman spectroscopy for analysis of protein secondary structure has focused largely on the correlation of the position of the amide I and amide III vibrations with the crystallographically determined fraction of each secondary structural element present in the protein. Features related with α-helix (1662–1655 and 1272–1264 cm^{-1}) and β-sheet structures (1674–1672 and 1242–1227 cm^{-1}) were

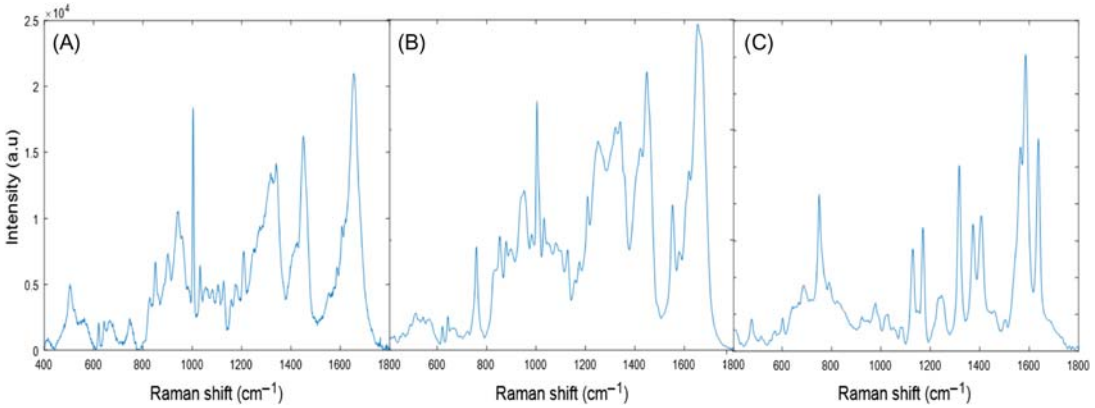


Figure 10.11
Raman spectra of (A) albumin, (B) fibrinogen and (C) cytochrome *c*.

Table 10.7: Assignment of amide I band positions of protein secondary structure in H_2O for both IR and Raman spectroscopies [127,130,135,136,138].

Secondary structure	Infrared (IR) band position (cm^{-1})	Raman band position (cm^{-1})
α-Helix	1648–1657	1650–1659
β-Sheets	1623–1641 and 1674–1695	1669–1674
Turns	1662–1686	1680–1690 and 1653–1656
Disordered/random coil	1642–1657	1640–1651

associated with the amide I and amide III modes. The review also classified structures as "mixed structures (α/β, α + β, s)" and "others" [137]. The bands relating to protein secondary structure of the amide I band are summarized in Table 10.7.

Particularly for the case of IR studies, protein secondary structures have been studied experimentally using both H_2O and D_2O. This is mainly due to the H_2O absorbance overlapping with the amide I band, but it is also thought to be easier to obtain spectra in D_2O as the bands occur at lower wave numbers than H_2O, meaning the region between 1400 and 1800 cm^{-1} exhibits relatively low absorbance (or scattering), providing an ideal window to observe the weaker bands of solubilized protein [135]. That said, using H_2O as a solvent is still preferable to D_2O when looking at protein structure, as D_2O can slightly alter the flexibility of proteins; for example, D_2O has been shown to increase the rigidity of most protein structures [139].

Fourier self-deconvolution (FSD) and derivative filtering are commonly employed methods for the investigation of protein secondary structure. FSD mathematically reduces bandwidths, so that the overlapping bands can be resolved [140]. This can also be achieved by differentiating the spectrum, commonly by calculating the second derivative, which exhibits a negative peak for every band or shoulder in the spectrum. For quantification,

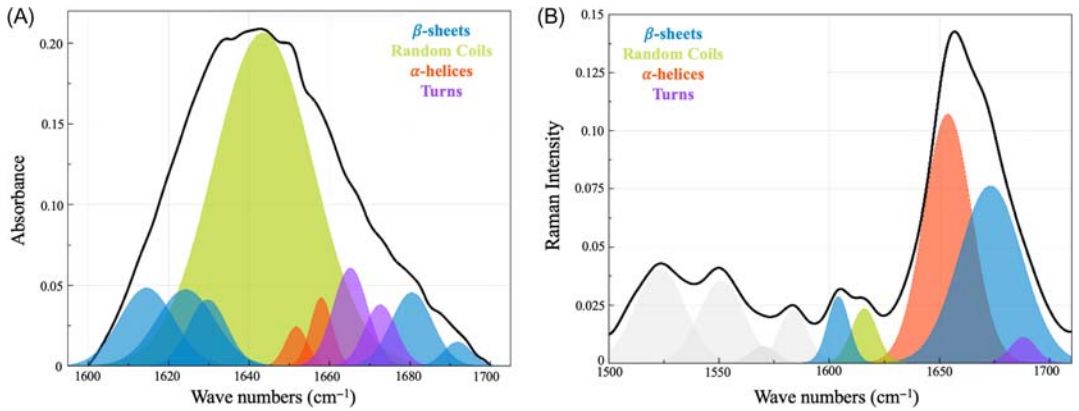


Figure 10.12
Curve fitting of the amide I band in serum, for (A) IR and (B) Raman spectra.

both FSD and second derivative spectra require curve fitting (Fig. 10.12), and the fractional areas of the fitted components correspond to the relative quantities of the different types of secondary structure [135]. Thus the band areas are directly proportional to the relative amount of secondary structure that is represented in that spectral region. The quantity of each component is expressed as a percentage, which provides a clear picture of overall protein structure. For example, second derivative analysis by Dong et al. predicted that immunoglobulin G (IgG) contains mostly β-sheets (64%) and turns (28%), with few random coils (5%) or α-helices (3%), a finding which was found to be closely correlated to the values obtained through X-ray crystallography [141].

Moreover, the estimation of amino acid side chain absorption must be considered in the analysis of protein spectra. Amino acid residues—arginine, asparagine, glutamine, etc.—also absorb in the amide I/II spectral region (Table 10.8). In the IR spectra of some globular proteins, the contribution of side chains can be as high as 10%−30% of the overall absorbance [136]. This can be a potential difficulty as the contribution from the amino acids will depend on their protonation state, which can be challenging to evaluate. The quantitative estimation of these groups can allow more refined predictions of the secondary structure of proteins and polypeptides by FTIR [135,142] (Table 10.8).

Rygula et al. also identify features of the Raman spectra that permit the description of the environment of numerous amino acid chains. These include amino acid side chain modes (e.g., tryptophan doublet 1360/1340 cm^{-1} and tyrosine doublet 860/833 cm^{-1}) or sulfur-containing residues in different physical states (C−S stretching with H at the *trans* position of the S atom: 640−680 cm^{-1} and C−S stretching with C at the *trans* position of the S atom; 740−760 cm^{-1}, S−S stretching: 508−512 cm^{-1} (GGG), 523−528 cm^{-1} (TGT), 540−545 cm^{-1} (TGT)) [137].

Table 10.8: Summary of main amino acid side chain absorptions found in infrared (IR) spectra in the 1400−1800 cm^{-1} region [127,135,143].

Side chain assignments	Approximate band position in IR (cm^{-1})
Asp, $\nu(C=O)$	1716
Glu, $\nu(C=O)$	1712
Asn, $\nu(C=O)$	1678
Arg, $\nu_{as}(CN_3H_5^+)$	1673
Gln, $\nu(C=O)$	1670
Arg, $\nu_s(CN_3H_5^+)$	1633
Lys, $\delta_{as}(NH_3^+)$	1629
Asn, $\delta(NH_2)$	1622
Gln, $\delta(NH_2)$	1610
Tyr, Ring-O^-	1602
Asp, $\nu_{as}(COO^-)$	1574
Glu, $\nu_{as}(COO^-)$	1560
Lys, $\delta_s(NH_3^+)$	1526
Tyr, Ring-OH	1518
Phe, Ring	1494

ν, Stretching; δ, bending; *as*, asymmetric; *s*, symmetric.

10.2.2 Spectroscopic signature of serum

Over the past decade, there has been a rapid increase of proof of concept publications for spectroscopic disease diagnostics, highlighting its potential for progression into the clinical environment [144]. The majority of the publications in the biomedical vibrational spectroscopy field have been based on the analysis of human tissue, with pilot studies showing it is possible to differentiate between healthy and cancerous tissue, as well as benign and malignant tumors [84]. Malignancies from various organs, such as breast, lung, colon, and prostate tissues, have previously been studied which has provided a platform of promising results [105,118,145−147]. Despite the high volume of published research, the technique has yet to make a successful transition into the clinic [69].

More recently there has been further interest in biofluid spectroscopy due to the ease of collection and handling, and minimal sample preparation required. Blood components such as serum and plasma are commonly analyzed for clinical reasons, carrying information regarding intra- and extra-cellular events. Biobanks exist as a valuable stock of both blood-derived biofluids, enabling the opportunity to repeat analysis or monitor treatment for disease progression [68]. More specifically, blood serum is the most complex biofluid, containing over 20,000 different proteins. As it perfuses all body organs, it gains proteomes from surrounding tissues and cells [148]. The lowmolecularweight fraction of serum (the peptidome) is information-rich, therefore the spectroscopic biosignature of serum is ideal for detecting disease states [149].

ATR-FTIR spectroscopy has proven to be a promising screening tool for detecting ovarian cancer from human blood, where both serum and plasma were used to discriminate ovarian cancer patients from healthy controls with a success rate of ~95% and ~97%, respectively [150]. Backhaus et al. used serum spectroscopy to differentiate between patients in good health and those with breast cancer, by applying unsupervised and supervised methods, reporting sensitivities and specificities of >92% for both [151]. Another pilot study found that the serum biosignature for cirrhotic patients, with and without hepatocellular carcinoma (HCC), could be successfully separated using support vector machine classification and leave-one-out cross-validation [152]. Furthermore, patients with extensive fibrosis in the liver have been separated from those without fibrosis, which is a common disorder in the early developmental stages of HCC, by using their FTIR serum spectra. Ollesch et al. introduced automated sampling for the first time, robotically spotting serum for high-throughput FTIR measurements, in their quest to identify and validate spectroscopic biomarker candidates for urinary bladder cancer [153]. By using as little as 1 μL of blood serum for ATR-FTIR analysis, Hands et al. were not only able to distinguish between serum of brain tumor patients and controls, but could effectively predict tumor grade by separating low-grade lesions from patients with glioblastoma (high-grade), highlighting the great potential of ATR-FTIR spectroscopic analysis of blood serum for determining the severity of brain tumors [154]. Paraskevaidi et al. demonstrated that ATR combined with chemometrics was capable of differentiating patients with various neurodegenerative diseases. Alzheimer's disease (AD) was identified with a sensitivity and specificity of 70%, and the AD patients were further segregated from those with dementia at 90% accuracy [155].

In order to maximize classification accuracy, feature extraction techniques can be utilized to pick out the most salient properties of the IR spectra. These methods isolate the features that are most highly correlated with a target set and rank them based on similarity, which in turn allows discrimination between classes and maximizes the intergroup differences [156]. In a further study, Hands et al. used feature extraction to select the most discriminatory spectral regions in their dataset [157]. Blood serum from a cohort of 433 patients, with and without a brain tumor, were collected for ATR-FTIR analysis. This substantial dataset consisted of control samples (non-cancer) and various brain tumor types; high- and low-grade glioma, meningioma, and metastatic tumors. The variable ranking function highlighted the wave number variables that were most salient between each spectral class. This technique proved to be effective in exposing the changes in the spectral signature between classes and tumor grade, and was capable of differentiating between cancer and noncancer; glioma and meningioma; metastatic and brain cancer; and high- and low-grade glioma, based on the most discriminatory spectral regions; amide I, amide II, C−O stretch of lipids/proteins, CH_2 of lipids/proteins, and PO_2^- of DNA/RNA. Furthermore, vibrations of C−O, C = O, and C−H of lipids/proteins, as well as the PO_2^- stretch from nucleic material, enabled discrimination between the three organs of origin of the metastatic cancer

samples (lung, melanoma, and breast). Previous studies have also highlighted these spectral regions when analyzing brain cancer and metastatic states via tissue spectroscopy [158,159]. To further explore the dataset gathered by Hands et al., machine learning techniques were employed in a computational study by Smith et al. [160]. The cancer versus non-cancer spectra were classified by random forest (RF), which utilized a Gini impurity metric to elucidate the most important wave number regions for classification. The carbohydrate region (997−1003 cm^{-1}) was found to be of highest RF importance, followed by the phosphate (1290−1294 cm^{-1}), lipid CH_2 (1462−1464 cm^{-1}), amide II (1527−1533 cm^{-1}), carbohydrate (1028−1034 cm^{-1}), and protein COO^- (1387−1390 cm^{-1}) regions. Two-dimensional correlation analysis was also performed alongside RF. The features highlighted by the 2D correlations were highly comparable to the RF results, further verifying the main spectral differences. The combination of these machine learning methods permitted successful discrimination of cancer and noncancer, with sensitivities and specificities of 92.8% and 91.5%, respectively, verifying the plausible use of orthogonal techniques to examine salient information for more accurate and rapid diagnostics [160].

Although ATR is preferred for biofluids analysis, the strong absorbance of water is still evident in the fingerprint region, which can obscure protein absorbance in liquid samples. Hence, the analysis of biofluids has been predominantly performed on air-dried samples, which can lead to chemical and physical inhomogeneity. The complex patterns that arise from dried biofluid drops have been of great interest over that past few decades, and various models have been published in an attempt to explain the complicated drying behavior [161]. When analyzing dried droplets using single-point transmission FTIR, Hughes et al. found the absorbance of the amide I/II region to be highly variable across a minute drop of blood serum. Spectra obtained from random locations across the dried serum spot, showed evidence of differences in sample thickness and heterogeneity. IR transmission imaging verified that there were biochemical differences across the drop [169]. Furthermore, the loss of light due to scattering, caused by the presence of cracks throughout the sample, along with varied drop thickness, led to the conclusion that samples need to be smooth and evenly spread for transmission measurements [162]. Deegan et al. proposed the coffee ring effect, whereby capillary flow forces the constituent biomolecules to move outwards to the edge of the drop, leaving behind a dense ring at the periphery [163]. This is common when drying biofluid drops, thus is a concern for spectroscopists, as the center of the drop may not be representative of the whole sample. Specifically in blood serum, the process is known as the Vroman effect, whereby a series of molecular displacements arise through protein exchange [164]. When a biological liquid is applied to a solid surface, lowmolecularweight proteins attach to the surface first, before being displaced by larger protein molecules over time. Therefore, the adsorption of proteins on to the substrate surface will be based on their differing affinities [165−167]. Gelation and cracking patterns have also been observed in dried biofluid drops, which are thought to be dependent on protein concentration [87,168].

These are the main limiting factors of using dried biofluids, as the surface inhomogeneity can cause peak shifts and alterations in band intensities [169]. Environmental—temperature and humidity—and experimental—volume and concentration—conditions have been shown to affect the drying patterns, hence it is vital the drying conditions are controlled and optimal protocols are developed in order to obtain a more homogenous deposition across the sample [161,170]. When measured in the ATR mode, however, the sample deposit can be completely contained within the area of the crystal, such that the evanescent wave measures the average of the entire drop, averaging out any inhomogeneities.

Raman spectroscopy is a complementary tool to IR spectroscopy and is compatible with aqueous samples. This technique allows the analysis to be carried out in the native state of bodily fluids, and therefore the additional drying step can be eliminated. A comprehensive proof of concept has been designed and conducted to detect HCC from patient serum using micro-Raman spectroscopy using dried and freeze-dried serum drops [171]. The aim of the study was to discriminate serum samples of patients with HCC and patients without HCC. Two groups of patients were classified with an overall accuracy of 84.5%−90.2% for dried serum drops and 86%−91.5% for freeze-dried serum. Although not specifically protein based, Mahmood et al. demonstrated the ability of Raman microspectroscopic analysis of dried blood serum to identify the presence of dengue infection, and to correlate the spectroscopic response with viral load [172]. Blood serum from patients with AD have been successfully separated from both healthy controls and those with neurodegenerative dementias, by applying Raman microspectroscopy and surface-enhanced Raman scattering (SERS) [173,174]. Another SERS-based immunoassay has been developed to monitor levels of the mucin protein MUC4 in patient serum, which will help in the early detection and diagnosis of pancreatic cancer [175].

Raman and SERS using silver nanoparticles, have been employed to identify signatures linked to ovarian cancer in blood plasma [176]. Both techniques provided satisfactory diagnostic accuracy for the detection of ovarian cancer, Raman achieving 94% sensitivity and 96% specificity, and SERS achieving 87% sensitivity and 89% specificity. For early ovarian cancer, Raman achieved sensitivity and specificity of 93% and 97%, respectively, while SERS had 80% sensitivity and 94% specificity.

As Raman microspectroscopy, by default, measures a very limited area of the sample, defined by the spot size of the objective focus, its suitability for analysis of chemically and physically inhomogeneously dried biofluid droplets is limited. However, an improved protocol for Raman spectroscopic analysis coupled with fractionation of serum using centrifugal filters to concentrate and separate low-molecular weight proteins has been demonstrated [162]. FTIR spectra were recorded in aqueous solutions of gelatin at concentrations as low as 100 mg/L; using Raman spectroscopy, high-quality spectra of gelatin solutions as low as 10 mg/L was achieved. Spectral features of human serum were

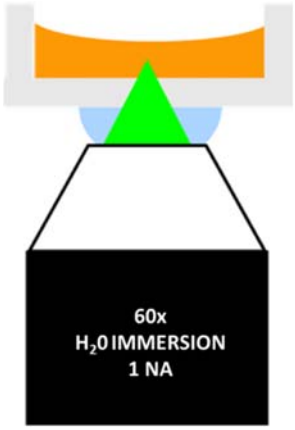


Figure 10.13
The inverted geometry used to analyze the serum focused by immersion objective.

found to be weak and partially obscured by water features. Dried deposits were shown to be physically and chemically inhomogeneous, resulting in unreliable results. Concentration of the serum using commercially available centrifugal filter devices resulted in enhanced spectral intensity and quality, and a 100% recovery of the analytes. Better analysis of serum using Raman spectroscopy was reported when the sample was analyzed in the inverted geometry using the water immersion objective with a 785 nm laser as source (Fig. 10.13). A drop of water is used to minimize the differences in the refractive indexes between sample, objective, and the substrate. However, the water drop does not contribute to the data collected, as it is outside the focus of the beam.

Using this analytical set up, Parachalil et al. recently reported on a systematic investigation of sample preparation considerations and data processing, for the analysis of blood plasma and serum [177]. In a solution of proteins, mixed in physiologically relevant concentrations, it was clearly seen that the poorly water-soluble fibrinogen presented significant challenges to measurement in the liquid phase, as it caused extensive Mie scattering of the source laser as well as the Raman scattering from itself, and the other protein constituents, preventing any quantitative measurement. This is a strong indication that the analysis of blood serum rather than plasma is favorable for optical-based techniques. In cases where the determination of fibrinogen levels is desired, the study showed that fibrinogen aggregates were broken down through mild sonication of the mixture, which significantly reduced the scatter.

10.2.3 Quantitative analysis

The ability of vibrational spectroscopy to quantify levels of biomolecules in blood serum has enabled various clinical studies over the past few decades [178]. In recent years, the

technique has been introduced for potential disease screening and monitoring for numerous health conditions, including arthritis, diabetes, heart disease, and a variety of cancers [179–181]. The development of a robust spectroscopic protocol could enable the replacement of the quantification methods currently used in medical practice.

An early study reported the quantitative analysis of dried serum films provided high accuracy for total protein, triglycerides, cholesterol, urea, and glucose, although the protocol was found to be less suitable for creatinine and uric acid [178]. Another demonstrated that both low-density lipoprotein (LDL) and high-density lipoprotein (HDL) cholesterol can be independently quantified using IR spectroscopy of dried serum films [182]. A later study showed patients with multiple myeloma are distinguishable from healthy patients through FTIR analysis of serum Ig, where the band intensity ratios in the spectral profile of the myeloma patients were higher than those with normal Ig levels [183].

Perez-Guaita et al. published an extensive quantitative study, using ATR-FTIR and partial least squares regression (PLSR) to determine the concentration of various proteins in human serum [184]. They established models for albumin and Ig, as well as total globulin and the albumin/globulin ratio, which are routinely determined in clinical practice as they can be indicative of diseased states [185]. Because the selected proteins exhibit different secondary structures, their analysis was focused on the behavior of the amide I/II bands. Their findings suggest that this technique could be useful in clinical practice as a routine assay for protein determination, as their prediction capability was extremely high for both albumin and total globulin, the root mean square error of prediction (RMSE) being lower than 5% for both. The RMSE for Ig and the albumin/globulin coefficient were slightly higher, between 7% and 14%, but the authors highlighted the ATR-FTIR method, coupled with PLSR, provides a promising quantification tool, especially at screening level [184].

Similar quantitative models using ATR-FTIR were constructed using PLSR in a recent study by Spalding et al., whereby human pooled serum was spiked with commercial human serum albumin (HSA) and IgG. The spiked serum was analyzed in both liquid and dried states to determine the optimal experimental protocol. In this particular study, in order to maintain reproducible spotting of the sample, the 10% air-dried sample preparation method was deemed to be the preferred approach. Using the same method, they analyzed serum gathered from 20 patients, to test the predictive power of the PLSR model when looking at more complex samples. The model was tested by two validation methods: leave one patient out cross-validation (LOPOCV) and k-fold cross-validation. For the prediction of total protein concentration, both models reported excellent results; the k-fold cross-validation produced an RMSE of 1.986 ± 0.778 mg/mL and an R^2 value of 0.934, whereas there was an RMSE of 1.534 ± 1.14 mg/mL and an R^2 value of 0.926 for the LOPOCV model. Both blind testing methods produced similar trends, with the prediction of the individual HSA and IgG concentrations not as effective as the total protein.

The prediction of IgG concentration was inferior to that of HSA, which Spalding et al. suggested could be due to the inability of FTIR to distinguish between the variable contributions of five major types of Ig that are present in human serum (IgA, IgD, IgE, IgG, and IgM) [186].

In comparison, Raman spectroscopy has been used for the quantitative analysis of blood to a lesser extent. Berger et al. used Raman microscopic analysis of liquid whole human blood and serum samples to quantify the content of six analytes, namely glucose, cholesterol, triglyceride, urea, total protein, and albumin [187]. Rohleder et al. undertook a comparison of mid-IR and Raman spectroscopy for the quantitative analysis of serum, based on spectra of the sera obtained from 247 blood donors [188]. PLSR analysis was used for the quantification of total protein, cholesterol, HDL and LDLs, triglycerides, glucose, urea, and uric acid. IR measurements were performed on dried samples, whereas Raman was performed on liquids. For all analytes, comparable root mean square error of cross-validation and root mean square error of validation were achieved.

Centrifugal filtration devices have been utilized to improve the sensitivity of quantitative analysis through Raman and IR spectroscopy, by separating the molecules according to their molecular weight. The proteins that are highly abundant in serum dominate the spectral profile, and through the removal of these proteins (albumin and globulins) the ability to monitor changes in the lowermolecular weight fraction (LMWF) is enhanced. Note, however, the importance of following a strict rinsing protocol to remove residual glycerine has been highlighted [66]. The centrifugal fractionation of human serum using both ATR-FTIR and Raman spectroscopy was evaluated by Bonnier et al. [189]. In this study, whole human serum was spiked with a wide range of known concentrations of glycine, between 0.5 and 50 mg/mL, in order to examine the capabilities of both ATR-FTIR and Raman spectroscopy in human serum monitoring. As glycine has a molecular weight of 75 Da, existing as the smallest amino acid, it can freely diffuse through the 10 kDa centrifugal filter membrane, making it an ideal target for quantitative analysis of the filtrate. Small aliquots (0.5 μL) of each concentration were measured in both liquid and dry state, after air drying for 10 minutes. Bonnier et al. used principal component analysis to explore and quantify the spectral variability caused by the adulteration of glycine to the human serum. Interestingly, the liquid samples produced a linear model ($R^2 = 0.9993$), whereas the results from the dried drops deviated from linearity above 10 mg/mL, and the relationship between the glycine concentration and spectral variations were expressed by a polynomial expansion ($R^2 = 0.9978$). In order to test the predictive power of both regression models, the lowest concentration, 0.5 mg/mL, was used as a blinded sample. The average predicted values for the liquid and dry models were 0.45 ± 0.16 mg/mL and 0.383 ± 0.007 mg/mL, respectively. Using the same protocol, more clinically relevant concentrations of glycine (0.01−2.5 mg/mL) were added to a centrifugal-filtered LMWF stock solution for further analysis. Similar to the model for whole serum in the dried state, the LMWF model

followed a polynomial fit achieving an R^2 value of 0.9981. In this case, the depletion of the abundant high molecular weight proteins greatly enhanced the sensitivity of the regression model for the detection of glycine, delivering a predictive value of 0.011 ± 0.006 mg/mL for the 0.01 mg/mL serum sample [189]. In the same study, it was demonstrated that the concentrate of the centrifugal process has been considerably concentrated (by factors up to $\times 10$ depending on the filter pore size), greatly enhancing the signal to (water) background levels for liquid phase Raman analysis. This indicates the potential for the prediction of other biomolecules that exist within the LMWF with this method, and with further research, such techniques could be translated into the clinical environment as a rapid tool for screening and monitoring.

10.3 Clinical translation

The high volume of research in the field of biomedical vibrational spectroscopy has indicated the potential utility of vibrational spectroscopy in a clinical environment. Numerous diagnostic and disease monitoring studies have reported extremely promising results, in many cases achieving sensitivity and specificity values greater than 90%. However, the techniques have not yet been successfully translated into the clinic [64]. The major hurdle to successful translation is arguably acceptance by health technology regulatory agencies, who determine which technologies are made available for public health. Criteria for successful acceptance require statistically verified clinical trials to prove clinical utility, but also clear understanding of the current clinical pathway in order to determine the economic and clinical impact of new technologies.

While promising proof-of-principle studies have supported clinical suitability, there have been few reported clinical trials employing either IR or Raman spectroscopy to confirm the utility in a prospective patient population. One of the examples closest to translation is the analysis of whole blood for the detection of malarial infection, currently being tested in a prospective cohort in Papua New Guinea [190]. This application of ATR-FTIR spectroscopy also looks to quantify levels of parasitemia in blood, providing clinically relevant information more rapidly than current methods [191]. The analysis of blood serum using ATR-FTIR spectroscopy for the early detection of brain cancer is also approaching clinical use, where a prospective patient cohort is currently being analyzed [96]. This application is also significant due to a clearly defined health economic study recently published, highlighting the clinical and economic benefits of implementing such a test into the current diagnostic pathway for brain cancers [192]. In summary, this study shows that a serum blood test at the primary care level could prioritize patients for neuroimaging, improving patient survival and quality of life while also providing cost savings by reducing unnecessary brain scans. Another benefit of this application is the potential of a seamless transition into the current clinical pathway. Blood tests at the primary care level are

commonly ordered, and the addition of a triage test into the clinical pathway would not significantly disrupt current clinic practices.

Current advances in technology may also facilitate the uptake of vibrational spectroscopy into standard clinical practice in the near future. Automated or high-throughput instrumentation would be best suited to clinical settings so as to minimize pressure on personnel resource. High-throughput technologies are available in IR transmission (or transflection) systems, largely attributed to the development of multichannel detectors and IR sources for discrete frequency spectroscopy, as well as the use of sample substrates which can be batch processed. This could have specific impacts on the translation of tissue imaging applications which have the potential to complement histopathology [106]. On the other hand, ATR-FTIR spectroscopy is inherently limited to a single point of analysis, the IRE, which restricts the overall sample throughput, particularly when considering cleaning the crystal between measurements as well as background subtraction. However, the development of low-cost IREs may provide a disposable substrate for ATR-FTIR spectroscopy, similarly enabling batch processing of samples and high-throughput spectral acquisition alongside the development of novel instrumentation [193]. High-throughput systems have similarly been explored for Raman analysis, and measurement in the liquid form avoids any further delay in the clinical work flow [194,195]. Importantly, the impact of reduced sampling times and consequent SNRs on classification and quantification accuracies have not yet been systematically explored.

References

[1] P.H. Lin, S.K. Yeh, W.C. Huang, H.Y. Chen, C.H. Chen, J.R. Sheu, et al., Research performance of biomarkers from biofluids in periodontal disease publications, J. Dental Sci. 10 (2015) 61–67. Available from: https://doi.org/10.1016/j.jds.2013.06.007.

[2] Z. Huang, A. McWilliams, H. Lui, D.I. McLean, S. Lam, H. Zeng, Near-infrared Raman spectroscopy for optical diagnosis of lung cancer, Int. J. Cancer 107 (2003) 1047–1052. Available from: https://doi.org/10.1002/ijc.11500.

[3] P. Crow, B. Barrass, C. Kendall, M. Hart-Prieto, M. Wright, R. Persad, et al., The use of Raman spectroscopy to differentiate between different prostatic adenocarcinoma cell lines, Br. J. Cancer 92 (2005) 2166–2170. Available from: https://doi.org/10.1038/sj.bjc.6602638.

[4] P.J. Caspers, G.W. Lucassen, E.A. Carter, H.A. Bruining, G.J. Puppels, In vivo confocal Raman microspectroscopy of the skin: noninvasive determination of molecular concentration profiles, J. Invest. Dermatol. 116 (2001) 434–442. Available from: https://doi.org/10.1046/j.1523-1747.2001.01258.x.

[5] T.D. Veenstra, T.P. Conrads, B.L. Hood, A.M. Avellino, R.G. Ellenbogen, R.S. Morrison, Biomarkers: mining the biofluid proteome, Mol. Cell. Proteom. 4 (2005) 409–418. Available from: https://doi.org/10.1074/mcp.M500006-MCP200.

[6] K. Kong, C. Kendall, N. Stone, I. Notingher, Raman spectroscopy for medical diagnostics—from in-vitro biofluid assays to in-vivo cancer detection, Adv. Drug. Deliv. Rev. 89 (2015) 121–134. Available from: https://doi.org/10.1016/j.addr.2015.03.009.

[7] S.-B. Su, T.C.W. Poon, V. Thongboonkerd, Human body fluid, BioMed. Res. Int. 2013 (2013) 1–2. Available from: https://doi.org/10.1155/2013/918793.

[8] J.T. Busher, Serum albumin and globulin, Clinical Methods: The History, Physical, Laboratory Examinations, Butterworths, Boston, MA, 1990, pp. 497–499.
[9] R.A. Shaw, S. Low-Ying, A. Man, K.-Z. Liu, C. Mansfield, C.B. Rileg, et al., Infrared spectroscopy of biofluids in clinical chemistry and medical diagnostics, Biomedical Vibrational Spectroscopy, John Wiley & Sons, Inc, 2007, pp. 79–103. Available from: https://doi.org/10.1002/9780470283172.ch4.
[10] L. Sheng, M. Luo, X. Sun, N. Lin, W. Mao, D. Su, Serum fibrinogen is an independent prognostic factor in operable nonsmall cell lung cancer, Int. J. Cancer 133 (2013) 2720–2725. Available from: https://doi.org/10.1002/ijc.28284.
[11] K.T. Nyuwi, C.H. Gyan Singh, S. Khumukcham, R. Rangaswamy, Y.S. Ezung, S.R. Chittvolu, et al., The role of serum fibrinogen level in the diagnosis of acute appendicitis, J. Clin. Diagn. Res. 11 (2017) PC13–PC15. Available from: https://doi.org/10.7860/JCDR/2017/21479.9319.
[12] S.-H. Yang, Y. Du, Y. Zhang, X.-L. Li, S. Li, R.-X. Xu, et al., Serum fibrinogen and cardiovascular events in Chinese patients with type 2 diabetes and stable coronary artery disease: a prospective observational study, BMJ Open. 7 (2017) e015041. Available from: https://doi.org/10.1136/bmjopen-2016-015041.
[13] M. Goicoechea, S.G. de Vinuesa, F. Gómez-Campderá, I. Aragoncillo, U. Verdalles, A. Mosse, et al., Serum fibrinogen levels are an independent predictor of mortality in patients with chronic kidney disease (CKD) stages 3 and 4, Kidney Int. Suppl. 68 (2008) S67–S70. Available from: https://doi.org/10.1038/ki.2008.519.
[14] R.L. Lundblad, Considerations for the use of blood plasma and serum for proteomic analysis, Internet J. Genomics Proteom. 1 (2005) 1–8.
[15] I.J. Mackie, S. Kitchen, S.J. Machin, G.D.O. Lowe, on behalf of the Haemostasis and Thrombosis Task Force of the British Committee for Standards in Haematology, Guidelines on fibrinogen assays, Br. J. Haematol. 121 (2003) 396–404. Available from: https://doi.org/10.1046/j.1365-2141.2003.04256.x.
[16] D.S. Wishart, T. Jewison, A.C. Guo, M. Wilson, C. Knox, Y. Liu, et al., HMDB 3.0—the human metabolome database in 2013, Nucleic Acids Res. 41 (2012) D801–D807. Available from: https://doi.org/10.1093/nar/gks1065.
[17] M.B. Bigler, S.B. Egli, C.M. Hysek, G. Hoenger, L. Schmied, F.S. Baldin, et al., Stress-induced in vivo recruitment of human cytotoxic natural killer cells favors subsets with distinct receptor profiles and associates with increased epinephrine levels, PLoS One 10 (2015) e0145635. Available from: https://doi.org/10.1371/journal.pone.0145635.
[18] S.K. Park, J.-H. Ryoo, C.-M. Oh, J.-M. Choi, Y.-J. Choi, K.O. Lee, et al., The risk of type 2 diabetes mellitus according to 2-h plasma glucose level: the Korean Genome and Epidemiology Study (KoGES), Diabetes Res. Clin. Pract. 146 (2018) 130–137. Available from: https://doi.org/10.1016/j.diabres.2017.08.002.
[19] Human Protein Reference Database, n.d. <http://hprd.org/browse/resultsBrowse?browse_type = localization&value = Extracellular&limit = 0>, (accessed 09.01.19).
[20] S.H. Rahman, Khairunnessa, Change of serum protein level during pregnancy and the impact of parity and diet on it, Bangladesh Med. Res. Counc. Bull. 4 (1978) 16–20.
[21] R.M. Cappelletti, Fibrinogen and Fibrin: Structure and Functional Aspects, Thrombin Funct. Pathophysiol (2012) 263–291.
[22] A. Chitsaz, S.A. Mousavi, Y. Yousef, V. Mostafa, Comparison of changes in serum fibrinogen level in primary intracranial hemorrhage (ICH) and ischemic stroke, ARYA Atherosclerosis 7 (2012) 142–145.
[23] J.G. van der Bom, M.P. de Maat, M.L. Bots, F. Haverkate, P.T. de Jong, A. Hofman, et al., Elevated plasma fibrinogen: cause or consequence of cardiovascular disease? Arterioscler. Thromb. Vasc. Biol. 18 (1998) 621–625.
[24] J.J. Stec, H. Silbershatz, G.H. Tofler, T.H. Matheney, P. Sutherland, I. Lipinska, et al., Association of fibrinogen with cardiovascular risk factors and cardiovascular disease in the Framingham Offspring Population, Circulation 102 (2000) 1634–1638.

[25] R.A.S. Ariëns, Elevated fibrinogen causes thrombosis, Blood 117 (2011) 4687. Available from: https://doi.org/10.1182/blood-2011-03-340422.
[26] J. Klovaite, B.G. Nordestgaard, A. Tybjærg-Hansen, M. Benn, Elevated fibrinogen levels are associated with risk of pulmonary embolism, but not with deep venous thrombosis, Am. J. Respir. Crit. Care Med. 187 (2013) 286–293. Available from: https://doi.org/10.1164/rccm.201207-1232OC.
[27] H. Toss, J. Gnarpe, H. Gnarpe, A. Siegbahn, B. Lindahl, L. Wallentin, Increased fibrinogen levels are associated with persistent *Chlamydia pneumoniae* infection in unstable coronary artery disease, Eur. Heart J. 19 (1998) 570–577.
[28] C. Salmon-Gandonnière, I. Benz-de Bretagne, E. Mercier, A. Joret, J.-M. Halimi, S. Ehrmann, et al., Iohexol clearance in unstable critically ill patients: a tool to assess glomerular filtration rate, Clin. Chem. Lab. Med. (CCLM) 54 (2016) 1777–1786. Available from: https://doi.org/10.1515/cclm-2015-1202.
[29] L.M. Lima, M. das G. Carvalho, M. de O. Sousa, Plasminogen and fibrinogen plasma levels in coronary artery disease, Rev. Bras. Hematol. Hemoter. 34 (2012) 298–301. Available from: https://doi.org/10.5581/1516-8484.20120075.
[30] I.O. Tekin, B. Pocan, A. Borazan, E. Ucar, G. Kuvandik, S. Ilikhan, et al., Positive correlation of CRP and fibrinogen levels as cardiovascular risk factors in early stage of continuous ambulatory peritoneal dialysis patients, Ren. Fail. 30 (2008) 219–225. Available from: https://doi.org/10.1080/08860220701813350.
[31] J.P. Nicholson, M.R. Wolmarans, G.R. Park, The role of albumin in critical illness, Br. J. Anaesth. 85 (2000) 599–610. Available from: https://doi.org/10.1093/bja/85.4.599.
[32] A. Fleck, F. Hawker, P.I. Wallace, G. Raines, J. Trotter, I.M. Ledingham, et al., Increased vascular permeability: a major cause of hypoalbuminaemia in disease and injury, Lancet 325 (2017) 781–784. Available from: https://doi.org/10.1016/S0140-6736(85)91447-3.
[33] X. Sun, M. Iles, C. Weissman, Physiologic variables and fluid resuscitation in the postoperative intensive care unit patient, Crit. Care Med. 21 (1993) 555–561.
[34] M.-L. Hu, S. Louie, C.E. Cross, P. Motchnik, B. Halliwell, Antioxidant protection against hypochlorous acid in human plasma, J. Lab. Clin. Med. 121 (2017) 257–262. Available from: https://doi.org/10.5555/uri:pii:002221439390154Q.
[35] J.S. Lee, Albumin for end-stage liver disease, Korean J. Intern. Med. 27 (2012) 13–19. Available from: https://doi.org/10.3904/kjim.2012.27.1.13.
[36] C.J. Janeway, P. Travers, M. Walport, The structure of a typical antibody molecule, in: Immunobiology: The Immune System in Health and Disease, fifth ed., Garland Science, New York, 2001. Available from: https://www.ncbi.nlm.nih.gov/books/NBK27144/.
[37] M.S.I. Chowdhury, N. Akhter, M. Haque, R. Aziz, N. Nahar, Serum total protein and albumin levels in different grades of protein energy malnutrition, J. Bangladesh Soc. Physiol. 3 (1970) 58–60. Available from: https://doi.org/10.3329/jbsp.v3i0.1799.
[38] B. Okutucu, Ö. Habib, Z. Figen, Comparison five methods for determination of total plasma protein concentration, J. Biochem. Biophys. Methods 70 (2007) 709–711. Available from: https://doi.org/10.1016/j.jbbm.2007.05.009.
[39] K. Hayden, C. van Heyningen, Measurement of total protein is a useful inclusion in liver function test profiles, Clin. Chem. 47 (2001) 793–794.
[40] B.G. Gazzard, HIV disease and the gastroenterologist, Gut 29 (1988) 1497–1505.
[41] C. Tian, L. Qian, X. Shen, J. Li, J. Wen, Distribution of serum total protein in elder Chinese, PLoS One 9 (2014) 1–5. Available from: https://doi.org/10.1371/journal.pone.0101242.
[42] E. Merler, F.S. Rosen, The gamma globulins, N. Engl. J. Med. 275 (1966) 536–542. Available from: https://doi.org/10.1056/NEJM196609082751006.
[43] T.B. Tomasi, W.A. Tisdale, Serum gamma-globulins in acute and chronic liver diseases, Nature 201 (1964) 834–835. Available from: https://doi.org/10.1038/201834a0.
[44] W. Gross, R.S. Snell, The serum gamma-globulin-level in malignant disease, Nature 178 (1956) 855. Available from: https://doi.org/10.1038/178855a0.

[45] A. Dispenzieri, M.A. Gertz, T.M. Therneau, R.A. Kyle, Retrospective cohort study of 148 patients with polyclonal gammopathy, Mayo Clin. Proc. 76 (2001) 476–487. Available from: https://doi.org/10.4065/76.5.476.
[46] R.H. Buckley, Humoral immunodeficiency, Clin. Immunol. Immunopathol. 40 (1986) 13–24.
[47] J.T. Whicher, C. Warren, R.E. Chambers, Immunochemical assays for immunoglobulins, Ann. Clin. Biochem. Int. J. Biochem. Lab. Med. 21 (1984) 78–91. Available from: https://doi.org/10.1177/000456328402100202.
[48] L. Thadikkaran, M.A. Siegenthaler, D. Crettaz, P.A. Queloz, P. Schneider, J.D. Tissot, Recent advances in blood-related proteomics, Proteomics 5 (2005) 3019–3034. Available from: https://doi.org/10.1002/pmic.200402053.
[49] N.L. Anderson, N.G. Anderson, The human plasma proteome, Mol. Cell. Proteom. 1 (2002) 845–867. Available from: https://doi.org/10.1074/mcp.R200007-MCP200.
[50] Z. Liu, S. Fan, H. Liu, J. Yu, R. Qiao, M. Zhou, et al., Enhanced detection of low-abundance human plasma proteins by integrating polyethylene glycol fractionation and immunoaffinity depletion, PLoS One 11 (2016) 1–17. Available from: https://doi.org/10.1371/journal.pone.0166306.
[51] L. Grgurevic, B. Macek, D. Durdevic, S. Vukicevic, Detection of bone and cartilage-related proteins in plasma of patients with a bone fracture using liquid chromatography-mass spectrometry, Int. Orthop. 31 (2007) 743–751. Available from: https://doi.org/10.1007/s00264-007-0404-z.
[52] M. Walker, J.G. Kublin, J.R. Zunt, Parasitic central nervous system infections in immunocompromised hosts: malaria, microsporidiosis, leishmaniasis, and african trypanosomiasis, Clin. Infect. Dis. 42 (2009) 115–125. Available from: https://doi.org/10.1086/498510.
[53] T.L. Daines, K.W. Morse, Determination of glucose in blood serum, J. Chem. Educ. 53 (1976) 126–127.
[54] S. Ray, M. Britschgi, C. Herbert, Y. Takeda-Uchimura, A. Boxer, K. Blennow, et al., Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins, Nat. Med. 13 (2007) 1359–1362.
[55] T. Eleftheriadis, G. Pissas, V. Liakopoulos, I. Stefanidis, Cytochrome c as a potentially clinical useful marker of mitochondrial and cellular damage, Front. Immunol. 7 (2016) 1–5. Available from: https://doi.org/10.3389/fimmu.2016.00279.
[56] M. Hüttemann, P. Pecina, M. Rainbolt, T.H. Sanderson, E. Valerian, L. Samavati, et al., The multiple functions of cyctochrome c and their regulation in life and death decisons of the mammalian cell: from respiration to apoptosis, Mitochondrion 11 (2012) 369–381. Available from: https://doi.org/10.1016/j.mito.2011.01.010.The.
[57] K. Matsuura, K. Canfield, W. Feng, M. Kurokawa, Chapter Two—Metabolic regulation of apoptosis in cancer, in: K.W. Jeon, L. Galluzzi (Eds.), International Review of Cell and Molecular Biology, Academic Press, 2016, pp. 43–87. Available from: https://doi.org/10.1016/bs.ircmb.2016.06.006.
[58] K. Barczyk, M. Kreuter, J. Pryjma, E.P. Booy, S. Maddika, S. Ghavami, et al., Serum cytochrome c indicates in vivo apoptosis and can serve as a prognostic marker during cancer therapy, Int. J. Cancer 116 (2005) 167–173. Available from: https://doi.org/10.1002/ijc.21037.
[59] N. Adachi, M. Hirota, M. Hamaguchi, K. Okamoto, K. Watanabe, F. Endo, Serum cytochrome c level as a prognostic indicator in patients with systemic inflammatory response syndrome, Clin. Chim. Acta 342 (2004) 127–136. Available from: https://doi.org/10.1016/j.cccn.2003.12.011.
[60] J. Radhakrishnan, S. Wang, I.M. Ayoub, J.D. Kolarova, R.F. Levine, R.J. Gazmuri, Circulating levels of cytochrome c after resuscitation from cardiac arrest: a marker of mitochondrial injury and predictor of survival, Am. J. Physiol. Heart Circ. Physiol. 292 (2007) H767–75. Available from: https://doi.org/10.1152/ajpheart.00468.2006.
[61] Z. Ben-Ari, H. Schmilovotz-Weiss, A. Belinki, O. Pappo, J. Sulkes, M.G. Neuman, et al., Circulating soluble cytochrome c in liver disease as a marker of apoptosis, J. Intern. Med. 254 (2003) 168–175. Available from: https://doi.org/10.1046/j.1365-2796.2003.01171.x.

[62] T. Eleftheriadis, G. Pissas, G. Antoniadi, V. Liakopoulos, I. Stefanidis, Damage-associated molecular patterns derived from mitochondria may contribute to the hemodialysis-associated inflammation, Int. Urol. Nephrol. 46 (2014) 107–112. Available from: https://doi.org/10.1007/s11255-013-0417-z.
[63] S.P. Singh, O. Ibrahim, H.J. Byrne, J.W. Mikkonen, A.P. Koistinen, A.M. Kullaa, et al., Recent advances in optical diagnosis of oral cancers: review and future perspectives: optical diagnosis of oral cancers, Head Neck 38 (2016) E2403–E2411. Available from: https://doi.org/10.1002/hed.24293.
[64] H.J. Byrne, M. Baranska, G.J. Puppels, N. Stone, B. Wood, K.M. Gough, et al., Spectropathology for the next generation: Quo vadis? Analyst 140 (2015) 2066–2073. Available from: https://doi.org/10.1039/C4AN02036G.
[65] M.J. Baker, H.J. Byrne, J. Chalmers, P. Gardner, R. Goodacre, A. Henderson, et al., Clinical applications of infrared and Raman spectroscopy: state of play and future challenges, Analyst 143 (2018) 1735–1757. Available from: https://doi.org/10.1039/C7AN01871A.
[66] F. Bonnier, M.J. Baker, H.J. Byrne, Vibrational spectroscopic analysis of body fluids: avoiding molecular contamination using centrifugal filtration, Anal. Methods 6 (2014) 5155. Available from: https://doi.org/10.1039/c4ay00891j.
[67] A.A. Bunaciu, Ş. Fleschin, V.D. Hoang, H.Y. Aboul-Enein, Vibrational spectroscopy in body fluids analysis, Crit. Rev. Anal. Chem. 47 (2017) 67–75. Available from: https://doi.org/10.1080/10408347.2016.1209104.
[68] M.J. Baker, C.S. Hughes, K.A. Hollywood, Biophotonics: Vibrational Spectroscopic Diagnostics, Morgan & Claypool Publishers, 2016. Available from: https://dx.doi.org/10.1088/978-1-6817-4071-3.
[69] A.L. Mitchell, K.B. Gajjar, G. Theophilou, F.L. Martin, P.L. Martin-Hirsch, Vibrational spectroscopy of biofluids for disease screening or diagnosis: translation from the laboratory to a clinical setting, J. Biophotonics 7 (2014) 153–165. Available from: https://doi.org/10.1002/jbio.201400018.
[70] N. Sheppard, The historical development of experimental techniques in vibrational spectroscopy, in: J.M. Chalmers, P.R. Griffiths (Eds.), Handbook of Vibrational Spectroscopy, John Wiley & Sons, Ltd, Chichester, UK, 2006. Available from: https://doi.org/10.1002/0470027320.s0101.
[71] Herschel Discovers Infrared Light, Cool Cosmos, 2019. http://coolcosmos.ipac.caltech.edu/cosmic_classroom/classroom_activities/herschel_bio.html.
[72] G.H. Rieke, History of infrared telescopes and astronomy, Exp. Astron. 25 (2009) 125–141. Available from: https://doi.org/10.1007/s10686-009-9148-7.
[73] A.L. Smith, Applied Infrared Spectroscopy: Fundamentals, Techniques, and Analytical Problem-Solving, Wiley, New York, 1979.
[74] L. Mertz, Astronomical infrared spectrometer, J. Astron 70 (1965) 548–551.
[75] P.R. Griffiths, R. Curbelo, C.T. Foskett, S.T. Dunn, Analytical Instrumentation, 8, Institute Society of America, 1970, pp. 11–14.
[76] R. Messerchmidt, M. Harthcock, Infrared Microscopy, Theory and Applications, Marcel Dekker Inc, New York, 1988.
[77] C.V. Raman, K.S. Krishnan, A new type of secondary radiation, Nature. 121 (1928) 501–502. Available from: https://doi.org/10.1038/121501c0.
[78] C. Adjouri, A. Elliasmine, Y. Le Duff, Low-light-level Raman spectroscopy using a charge-coupled device detector, Spectroscopy 44 (1996) 46.
[79] G. Puppels, Laser irradiation and Raman spectroscopy of single living cells and chromosomes: sample degradation occurs with 514.5 nm but not with 660 nm laser light*1, Exp. Cell Res. 195 (1991) 361–367. Available from: https://doi.org/10.1016/0014-4827(91)90385-8.
[80] T. Vo-Dinh, G. Gauglitz (Eds.), Handbook of Spectroscopy, Wiley-VCH, Weinheim; [Cambridge], 2003.
[81] P.R. Griffiths, J.A. De Haseth, Fourier Transform Infrared Spectrometry, second ed., Wiley-Interscience, Hoboken, NJ, 2007.
[82] B.C. Smith, Fundamentals of Fourier Transform Infrared Spectroscopy, CRC Press, Boca Raton, FL, 2011. Available from: http://www.crcnetbase.com/isbn/9781420069303 (accessed 31.01.18).
[83] A.A. Michelson, E.W. Morley, On the relative motion of the Earth and the luminiferous ether, Am. J. Sci. s3-34 (1887) 333–345. Available from: https://doi.org/10.2475/ajs.s3-34.203.333.

[84] Z. Movasaghi, S. Rehman, I. ur Rehman Dr., Fourier transform infrared (FTIR) spectroscopy of biological tissues, Appl. Spectrosc. Rev. 43 (2008) 134–179. Available from: https://doi.org/10.1080/05704920701829043.
[85] M.J. Baker, J. Trevisan, P. Bassan, R. Bhargava, H.J. Butler, K.M. Dorling, et al., Using Fourier transform IR spectroscopy to analyze biological materials, Nat. Protoc. 9 (2014) 1771–1791. Available from: https://doi.org/10.1038/nprot.2014.110.
[86] M.J. Pilling, P. Bassan, P. Gardner, Comparison of transmission and transflectance mode FTIR imaging of biological tissue, Analyst 140 (2015) 2383–2392. Available from: https://doi.org/10.1039/C4AN01975J.
[87] L. Lovergne, G. Clemens, V. Untereiner, R.A. Lukaszweski, G.D. Sockalingum, M.J. Baker, Investigating optimum sample preparation for infrared spectroscopic serum diagnostics, Anal. Methods 7 (2015) 7140–7149. Available from: https://doi.org/10.1039/C5AY00502G.
[88] S.E. Glassford, B. Byrne, S.G. Kazarian, Recent applications of ATR FTIR spectroscopy and imaging to proteins, Biochim. Biophys. Acta (BBA) Proteins Proteom. 1834 (2013) 2849–2858. Available from: https://doi.org/10.1016/j.bbapap.2013.07.015.
[89] K. Dorling, M.J. Baker, Highlighting attenuated total reflection Fourier transform infrared spectroscopy for rapid serum analysis, Trends Biotechnol. 31 (2013) 325–327. Available from: https://doi.org/10.1016/j.tibtech.2013.03.009.
[90] J. Filik, M.D. Frogley, J.K. Pijanka, K. Wehbe, G. Cinque, Electric field standing wave artefacts in FTIR micro-spectroscopy of biological materials, Analyst 137 (2012) 853. Available from: https://doi.org/10.1039/c2an15995c.
[91] J. Lee, E. Gazi, J. Dwyer, M.D. Brown, N.W. Clarke, J.M. Nicholson, et al., Optical artefacts in transflection mode FTIR microspectroscopic images of single cells on a biological support: the effect of back-scattering into collection optics, Analyst 132 (2007) 750. Available from: https://doi.org/10.1039/b702064c.
[92] P. Bassan, H.J. Byrne, J. Lee, F. Bonnier, C. Clarke, P. Dumas, et al., Reflection contributions to the dispersion artefact in FTIR spectra of single biological cells, Analyst 134 (2009) 1171. Available from: https://doi.org/10.1039/b821349f.
[93] P. Bassan, H.J. Byrne, F. Bonnier, J. Lee, P. Dumas, P. Gardner, Resonant Mie scattering in infrared spectroscopy of biological materials—understanding the 'dispersion artefact', Analyst 134 (2009) 1586. Available from: https://doi.org/10.1039/b904808a.
[94] P. Bassan, A. Kohler, H. Martens, J. Lee, H.J. Byrne, P. Dumas, et al., Resonant Mie Scattering (RMieS) correction of infrared spectra from highly scattering biological samples, Analyst 135 (2010) 268–277. Available from: https://doi.org/10.1039/B921056C.
[95] B.H. Stuart, Infrared Spectroscopy: Fundamentals and Applications, John Wiley & Sons, Ltd., Hoboken, NJ, 2005. Available from: http://www.SLQ.eblib.com.au/patron/FullRecord.aspx?p = 194354. accessed 20.02.18.
[96] H.J. Butler, P.M. Brennan, J.M. Cameron, et al., Development of high-throughput ATR-FTIR technology for rapid triage of brain cancer, Nat Commun 10 (2019). Available from: https://doi.org/10.1038/s41467-019-12527-5.
[97] R.A. Shaw, S. Low-Ying, A. Man, K.-Z. Liu, C. Mansfield, C.B. Rileg, et al., Infrared spectroscopy of biofluids in clinical chemistry and medical diagnostics, Biomedical Vibrational Spectroscopy, John Wiley and Sons, Inc, Hoboken, NJ, 2008, pp. 79–103.
[98] G.S. Bumbrah, R.M. Sharma, Raman spectroscopy—basic principle, instrumentation and selected applications for the characterization of drugs of abuse, Egypt. J. Forensic Sci. 6 (2016) 209–215. Available from: https://doi.org/10.1016/j.ejfs.2015.06.001.
[99] E. Smith, G. Dent, The Raman experiment—Raman instrumentation, sample presentation, data handling and practical aspects of interpretation, in: Modern Raman Spectroscopy—A Practical Approach, John Wiley & Sons, Ltd, Chichester, UK, 2005, pp. 23–70. Available from: https://doi.org/10.1002/0470011831.ch2.

[100] M.J. Bertrand, Handbook of Instrumental Techniques for Analytical Chemistry Edited by Frank A. Settle. Prentice Hall: Upper Saddle River. 1997. xxi + 995 pp. ISBN 0-13-177338-0, J. Am. Chem. Soc. 120 (1998). Available from: https://doi.org/10.1021/ja975671k. 6633.

[101] M. Jackson, H.H. Mantsch, The use and misuse of FTIR spectroscopy in the determination of protein structure, Crit. Rev. Biochem. Mol. Biol. 30 (1995) 95–120. Available from: https://doi.org/10.3109/10409239509085140.

[102] H.L. Casal, H.H. Mantsch, Polymorphic phase behaviour of phospholipid membranes studied by infrared spectroscopy, Biochim. Biophys. Acta 779 (1984) 381–401.

[103] M. Mathlouthi, J.L. Koenig, Vibrational spectra of carbohydrates, Adv. Carbohydr. Chem. Biochem. 44 (1986) 7–89.

[104] E. Taillandier, J. Liquier, J.A. Taboury, Advances in Spectroscopy: advances in infrared and raman spectroscopy, in: R.J.H. Clark, R.E. Hester (Eds.), Advances in Spectroscopy, Wiley, New York, 1985, p. 65.

[105] E. Gazi, M. Baker, J. Dwyer, N.P. Lockyer, P. Gardner, J.H. Shanks, et al., A correlation of FTIR spectra derived from prostate cancer biopsies with Gleason grade and tumour stage, Eur. Urol. 50 (2006) 750–761. Available from: https://doi.org/10.1016/j.eururo.2006.03.031.

[106] D.C. Fernandez, R. Bhargava, S.M. Hewitt, I.W. Levin, Infrared spectroscopic imaging for histopathologic recognition, Nat. Biotechnol. 23 (2005) 469–474. Available from: https://doi.org/10.1038/nbt1080.

[107] R. Dukor, Vibrational spectroscopy in the detection of cancer, in: J.M. Chalmers, P.R. Griffths (Eds.), Handbook of Vibrational Spectroscopy, Wiley, Chichester, 2002.

[108] R.C. Lord, N.T. Yu, Laser-excited Raman spectroscopy of biomolecules. I. Native lysozyme and its constituent amino acids, J. Mol. Biol. 50 (1970) 509–524.

[109] M.C. Tobin, Raman spectra of crystalline lysozyme, pepsin, and alpha chymotrypsin, Science 161 (1968) 68–69. Available from: https://doi.org/10.1126/science.161.3836.68.

[110] A.G. Walton, M.J. Deveney, J.L. Koenig, Raman spectroscopy of calcified tissue, Calcif. Tissue Res. 6 (1970) 162–167. Available from: https://doi.org/10.1007/BF02196195.

[111] G.J. Puppels, J. Breve, Biomedical applications of spectroscopy, in: R. H.H. Clark, R.E. Hester (Eds.), Advances in Apectroscopy, John Wiley & Sons, New York, 1996.

[112] M. Gniadecka, H.C. Wulf, O.F. Neilsen, D.H. Christensen, J. Hercogova, Distinctive molecular abnormalities in benign and malignant skin-lesions—studies by Raman-spectroscopy, Photochem. Photobiol. 66 (1997) 418–423.

[113] J. Smith, C. Kendall, A. Sammon, J. Christie-Brown, N. Stone, Raman spectral mapping in the assessment of axillary lymph nodes in breast cancer, Technol. Cancer Res. Treat. 2 (2003) 327–331. Available from: https://doi.org/10.1177/153303460300200407.

[114] E.B. Hanlon, R. Manoharan, T.W. Koo, K.E. Shafer, J.T. Motz, M. Fitzmaurice, et al., Prospects for in vivo Raman spectroscopy, Phys. Med. Biol. 45 (2000) R1–R59.

[115] P.J. Caspers, G.W. Lucassen, R. Wolthuis, H.A. Bruining, G.J. Puppels, In vitro and in vivo Raman spectroscopy of human skin, Biospectroscopy 4 (1998) S31–39. Available from: https://dx.doi.org/10.1002/(SICI)1520-6343(1998)4:5 + < S31::AID-BSPY4 > 3.0.CO;2-M.

[116] M. Miljković, B. Bird, K. Lenau, A.I. Mazur, M. Diem, Spectral cytopathology: new aspects of data collection, manipulation and confounding effects, Analyst 138 (2013) 3975. Available from: https://doi.org/10.1039/c3an00185g.

[117] P. Garidel, H. Schott, Fourier-transform midinfrared spectroscopy for analysis and screening of liquid protein formulations, BioProcess Int 4(6) (2006) 48–55.

[118] M. Baker, E. Gazi, M. Brown, et al., FTIR-based spectroscopic analysis in the identification of clinically aggressive prostate cancer, Br J Cancer 99 (2008) 1859–1866. Available from: https://doi.org/10.1038/sj.bjc.6604753.

[119] G. Bellisola, C. Sorio, Infrared spectroscopy and microscopy in cancer research and diagnosis, Am. J. Cancer Res. 2 (2012) 1.

[120] D. Naumann, FT-infrared and FT-Raman spectroscopy in biomedical research, Appl. Spectrosc. Rev. 36 (2001) 239–298. Available from: https://doi.org/10.1081/ASR-100106157.
[121] Y. Chen, J. Dai, X. Zhou, Y. Liu, W. Zhang, G. Peng, Raman spectroscopy analysis of the biochemical characteristics of molecules associated with the malignant transformation of gastric mucosa, PLoS One 9 (2014) e93906. Available from: https://doi.org/10.1371/journal.pone.0093906.
[122] C. Molony, J. McIntyre, A. Maguire, R. Hakimjavadi, D. Burtenshaw, G. Casey, et al., Label-free discrimination analysis of de-differentiated vascular smooth muscle cells, mesenchymal stem cells and their vascular and osteogenic progeny using vibrational spectroscopy, Biochim. Biophys. Acta Mol. Cell Res. 2018 (1865) 343–353. Available from: https://doi.org/10.1016/j.bbamcr.2017.11.006.
[123] J.L.R. Arrondo, A. Muga, J. Castresana, F.M. Goñi, Quantitative studies of the structure of proteins in solution by fourier-transform infrared spectroscopy, Prog. Biophys. Mol. Biol. 59 (1993) 23–56. Available from: https://doi.org/10.1016/0079-6107(93)90006-6.
[124] J.L. Arrondo, F.M. Goñi, Structure and dynamics of membrane proteins as studied by infrared spectroscopy, Prog. Biophys. Mol. Biol. 72 (1999) 367–405.
[125] P.I. Haris, D. Chapman, Does Fourier-transform infrared spectroscopy provide useful information on protein structures? Trends Biochem. Sci. 17 (1992) 328–333.
[126] H. Fabian, W. Mantele, Infrared spectroscopy of proteins, in: J.M. Chalmers, P.R. Griffiths (Eds.), Handbook of Vibrational Spectroscopy, John Wiley & Sons, Ltd, Chichester, UK, 2006. Available from: https://doi.org/10.1002/0470027320.s8201.
[127] A. Barth, Infrared spectroscopy of proteins, Biochim. Biophys. Acta (BBA) Bioenerg. 1767 (2007) 1073–1101. Available from: https://doi.org/10.1016/j.bbabio.2007.06.004.
[128] W. Gallagher, FTIR analysis of protein structure, Course Man. Chem. 455 (2009) 1–8.
[129] K. Spalding, Developing Spectroscopic Biofluid Diagnostics: Monitoring and Therapeutic Profiling of Melanoma Patients, University of Strathclyde, 2018.
[130] A. Rygula, K. Majzner, K.M. Marzec, A. Kaczor, M. Pilarczyk, M. Baranska, Raman spectroscopy of proteins: a review: Raman spectroscopy of proteins, J. Raman Spectrosc. 44 (2013) 1061–1076. Available from: https://doi.org/10.1002/jrs.4335.
[131] A. Barth, C. Zscherp, What vibrations tell us about proteins, Q. Rev. Biophys. 35 (2002) 369–430. Available from: https://doi.org/10.1017/S0033583502003815.
[132] M. Kinalwa, E.W. Blanch, A.J. Doig, Determination of protein fold class from Raman or Raman optical activity spectra using random forests, Protein Sci. 20 (2011) 1668–1674. Available from: https://doi.org/10.1002/pro.695.
[133] S.M. Ali, F. Bonnier, H. Lambkin, K. Flynn, V. McDonagh, C. Healy, et al., A comparison of Raman, FTIR and ATR-FTIR micro spectroscopy for imaging human skin tissue sections, Anal. Methods. 5 (2013) 2281. Available from: https://doi.org/10.1039/c3ay40185e.
[134] S. Hu, I.K. Morris, J.P. Singh, K.M. Smith, T.G. Spiro, Complete assignment of cytochrome c resonance Raman spectra via enzymic reconstitution with isotopically labeled hemes, J. Am. Chem. Soc. 115 (1993) 12446–12458. Available from: https://doi.org/10.1021/ja00079a028.
[135] H. Fabian, W. Mantele, Infrared spectroscopy of proteins, Biomedical Applications, John Wiley & Sons, 2002, p. 27.
[136] J. Kong, S. Yu, Fourier transform infrared spectroscopic analysis of protein secondary structures, Acta Biochim. Biophys. Sin. 39 (2007) 549–559. Available from: https://doi.org/10.1111/j.1745-7270.2007.00320.x.
[137] T. Kitagawa, S. Hirota, Raman spectroscopy of proteins, in: J.M. Chalmers, P.R. Griffiths (Eds.), Handbook of Vibrational Spectroscopy, John Wiley & Sons Ltd, Chichester, 2002, pp. 3426–3446.
[138] S.U. Sane, S.M. Cramer, T.M. Przybycien, A. Holistic, Approach to protein secondary structure characterization using amide I band raman spectroscopy, Anal. Biochem. 269 (1999) 255–272. Available from: https://doi.org/10.1006/abio.1999.4034.
[139] P. Cioni, G.B. Strambini, Effect of heavy water on protein flexibility, Biophys. J. 82 (2002) 3246–3253. Available from: https://doi.org/10.1016/S0006-3495(02)75666-X.

[140] P.B. Tooke, Fourier self-deconvolution in IR spectroscopy, Trends Anal. Chem. (TrAC) 7 (1988) 130–136. Available from: https://doi.org/10.1016/0165-9936(88)87010-9.
[141] A. Dong, P. Huang, W.S. Caughey, Protein secondary structures in water from second-derivative amide I infrared spectra, Biochemistry 29 (1990) 3303–3308. Available from: https://doi.org/10.1021/bi00465a022.
[142] S.Y. Venyaminov, N.N. Kalnin, Quantitative IR spectrophotometry of peptide compounds in water (H2O) solutions. I. Spectral parameters of amino acid residue absorption bands, Biopolymers 30 (1990) 1243–1257. Available from: https://doi.org/10.1002/bip.360301309.
[143] L.K. Tamm, S.A. Tatulian, Infrared spectroscopy of proteins and peptides in lipid bilayers, Q. Rev. Biophys. 30 (1997) 365–429. Available from: https://doi.org/10.1017/S0033583597003375.
[144] N.M. Ralbovsky, I.K. Lednev, Raman spectroscopy and chemometrics: a potential universal method for diagnosing cancer, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 219 (2019) 463–487. Available from: https://doi.org/10.1016/j.saa.2019.04.067.
[145] M.J. Walsh, S.E. Holton, A. Kajdacsy-Balla, R. Bhargava, Attenuated total reflectance Fourier-transform infrared spectroscopic imaging for breast histopathology, Vib. Spectrosc. 60 (2012) 23–28. Available from: https://doi.org/10.1016/j.vibspec.2012.01.010.
[146] B. Bird, S. Remiszewski, A. Akalin, M. Kon, M. Diem, et al., Infrared spectral histopathology (SHP): a novel diagnostic tool for the accurate classification of lung cancer, Lab. Invest. 92 (2012) 1358.
[147] P. Lasch, W. Haensch, D. Naumann, M. Diem, Imaging of colorectal adenocarcinoma using FT-IR microspectroscopy and cluster analysis, Biochim. Biophys. Acta (BBA) Mol. Basis Dis. 1688 (2004) 176–186. Available from: https://doi.org/10.1016/j.bbadis.2003.12.006.
[148] R.S. Tirumalai, K.C. Chan, D.A. Prieto, H.J. Issaq, T.P. Conrads, T.D. Veenstra, Characterization of the low molecular weight human serum proteome, Mol. Cell. Proteom. 2 (2003) 1096–1103. Available from: https://doi.org/10.1074/mcp.M300031-MCP200.
[149] E.F. Petricoin, C. Belluco, R.P. Araujo, L.A. Liotta, The blood peptidome: a higher dimension of information content for cancer biomarker discovery, Nat. Rev. Cancer. 6 (2006) 961–967. Available from: https://doi.org/10.1038/nrc2011.
[150] K. Gajjar, J. Trevisan, G. Owens, P.J. Keating, N.J. Wood, H.F. Stringfellow, et al., Fourier-transform infrared spectroscopy coupled with a classification machine for the analysis of blood plasma or serum: a novel diagnostic approach for ovarian cancer, Analyst 138 (2013) 3917–3926. Available from: https://doi.org/10.1039/C3AN36654E.
[151] J. Backhaus, R. Mueller, N. Formanski, N. Szlama, H.-G. Meerpohl, M. Eidt, et al., Diagnosis of breast cancer with infrared spectroscopy from serum samples, Vib. Spectrosc. 52 (2010) 173–177. Available from: https://doi.org/10.1016/j.vibspec.2010.01.013.
[152] X. Zhang, G. Thiéfin, C. Gobinet, V. Untereiner, I. Taleb, B. Bernard-Chabert, et al., Profiling serologic biomarkers in cirrhotic patients via high-throughput Fourier transform infrared spectroscopy: toward a new diagnostic tool of hepatocellular carcinoma, Transl. Res. 162 (2013) 279–286. Available from: https://doi.org/10.1016/j.trsl.2013.07.007.
[153] J. Ollesch, M. Heinze, H.M. Heise, T. Behrens, T. Brüning, K. Gerwert, It's in your blood: spectral biomarker candidates for urinary bladder cancer from automated FTIR spectroscopy: spectral cancer biomarkers from high-throughput FTIR spectroscopy, J. Biophotonics 7 (2014) 210–221. Available from: https://doi.org/10.1002/jbio.201300163.
[154] J.R. Hands, K.M. Dorling, P. Abel, K.M. Ashton, A. Brodbelt, C. Davis, et al., Attenuated total reflection Fourier transform infrared (ATR-FTIR) spectral discrimination of brain tumour severity from serum samples: serum spectroscopy gliomas, J. Biophotonics 7 (2014) 189–199. Available from: https://doi.org/10.1002/jbio.201300149.
[155] M. Paraskevaidi, C.L.M. Morais, K.M.G. Lima, J.S. Snowden, J.A. Saxon, A.M.T. Richardson, et al., Differential diagnosis of Alzheimer's disease using spectrochemical analysis of blood, Proc. Natl Acad. Sci. U.S.A. 114 (2017) E7929–E7938. Available from: https://doi.org/10.1073/pnas.1701517114.
[156] D. Vicinanza, R. Stables, G. Clemens, M. Baker, Assisted differentiated stem cell classification in infrared spectroscopy using auditory feedback, in: 20th International Conference on Auditory Display

(ICAD−2014), New York University's Steinhardt School of Culture, Education and Human Development, New York, 2014, p. 6.
[157] J.R. Hands, G. Clemens, R. Stables, K. Ashton, A. Brodbelt, C. Davis, et al., Brain tumour differentiation: rapid stratified serum diagnostics via attenuated total reflection Fourier-transform infrared spectroscopy, J. Neuro-Oncol. 127 (2016) 463−472. Available from: https://doi.org/10.1007/s11060-016-2060-x.
[158] K. Gajjar, L.D. Heppenstall, W. Pang, K.M. Ashton, J. Trevisan, I.I. Patel, et al., Diagnostic segregation of human brain tumours using Fourier-transform infrared and/or Raman spectroscopy coupled with discriminant analysis, Anal. Meth 5 (2013) 89−102. Available from: https://doi.org/10.1039/C2AY25544H.
[159] C. Krafft, L. Shapoval, S.B. Sobottka, G. Schackert, R. Salzer, Identification of primary tumors of brain metastases by infrared spectroscopic imaging and linear discriminant analysis, Technol. Cancer Res. Treat. 5 (2006) 291−298. Available from: https://doi.org/10.1177/153303460600500311.
[160] B.R. Smith, K.M. Ashton, A. Brodbelt, T. Dawson, M.D. Jenkinson, N.T. Hunt, et al., Combining random forest and 2D correlation analysis to identify serum spectral signatures for neuro-oncology, Analyst 141 (2016) 3668−3678.
[161] J.M. Cameron, H.J. Butler, D.S. Palmer, M.J. Baker, Biofluid spectroscopic disease diagnostics: a review on the processes and spectral impact of drying, J. Biophotonics 11 (2018) e201700299. Available from: https://doi.org/10.1002/jbio.201700299.
[162] C. Hughes, M. Brown, G. Clemens, A. Henderson, G. Monjardez, N.W. Clarke, et al., Assessing the challenges of FTIR spectroscopic analysis of blood serum, J. Biophotonics 7 (2014) 180−188. Available from: https://doi.org/10.1002/jbio.201300167.
[163] R.D. Deegan, Pattern formation in drying drops, Phys. Rev. E 61 (2000) 475.
[164] S.L. Hirsh, D.R. McKenzie, N.J. Nosworthy, J.A. Denman, O.U. Sezerman, M.M.M. Bilek, The Vroman effect: competitive protein exchange with dynamic multilayer protein aggregates, Colloids Surf. B Biointerfaces 103 (2013) 395−404. Available from: https://doi.org/10.1016/j.colsurfb.2012.10.039.
[165] L. Vroman, A.L. Adams, G.C. Fischer, P.C. Munoz, Interaction of high molecular weight kininogen, factor XII, and fibrinogen in plasma at interfaces, Blood 55 (1980) 156−159.
[166] A.H. Schmaier, L. Silver, L. Ann, G.C. Adams, P.C. Fischer, L. Munoz, et al., Colman, The effect of high molecular weight kininogen on surface-adsorbed fibrinogen, Thrombosis Res. 33 (2017) 51−67. Available from: https://doi.org/10.1016/0049-3848(84)90154-3.
[167] A.L. Adams, G.C. Fischer, P.C. Munoz, L. Vroman, Convex-lens-on-slide: a simple system for the study of human plasma and blood in narrow spaces, J. Biomed. Mater. Res. 18 (1984) 643−654. Available from: https://doi.org/10.1002/jbm.820180606.
[168] C.C. Annarelli, J. Fornazero, J. Bert, J. Colombani, Crack patterns in drying protein solution drops, Eur. Phys. J. E Soft Matter Biol. Phys. 5 (2001) 599−603.
[169] F. Bonnier, F. Petitjean, M.J. Baker, H.J. Byrne, Improved protocols for vibrational spectroscopic analysis of body fluids: improved protocols for vibrational spectroscopic analysis of body fluids, J. Biophotonics 7 (2014) 167−179. Available from: https://doi.org/10.1002/jbio.201300130.
[170] L. Lovergne, P. Bouzy, V. Untereiner, R. Garnotel, M.J. Baker, G. Thiéfin, et al., Biofluid infrared spectro-diagnostics: pre-analytical considerations for clinical applications, Faraday Discuss. 187 (2016) 521−537. Available from: https://doi.org/10.1039/C5FD00184F.
[171] I. Taleb, G. Thiefin, C. Gobinet, V. Untereiner, B. Bernard-Chabert, A. Heurgue, et al., Diagnosis of hepatocellular carcinoma in cirrhotic patients: a proof-of-concept study using serum micro-Raman spectroscopy, Analyst 138 (2013) 4006−4014. Available from: https://doi.org/10.1039/C3AN00245D.
[172] T. Mahmood, H. Nawaz, A. Ditta, M.I. Majeed, M.A. Hanif, N. Rashid, et al., Raman spectral analysis for rapid screening of dengue infection, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 200 (2018) 136−142. Available from: https://doi.org/10.1016/j.saa.2018.04.018.
[173] E. Ryzhikova, O. Kazakov, L. Halamkova, D. Celmins, P. Malone, E. Molho, et al., Raman spectroscopy of blood serum for Alzheimer's disease diagnostics: specificity relative to other types of dementia, J. Biophotonics 8 (2015) 584−596. Available from: https://doi.org/10.1002/jbio.201400060.

[174] E. Ryzhikova, N.M. Ralbovsky, L. Halámková, D. Celmins, P. Malone, E. Molho, et al., Multivariate statistical analysis of surface enhanced Raman spectra of human serum for Alzheimer's disease diagnosis, Appl. Sci. 9 (2019) 3256. Available from: https://doi.org/10.3390/app9163256.

[175] G. Wang, R.J. Lipert, M. Jain, S. Kaur, S. Chakraboty, M.P. Torres, et al., Detection of the potential pancreatic cancer marker MUC4 in serum using surface-enhanced Raman scattering, Anal. Chem. 83 (2012) 2554–2561. Available from: https://doi.org/10.1021/ac102829b.Detection.

[176] M. Paraskevaidi, K.M. Ashton, H.F. Stringfellow, N.J. Wood, P.J. Keating, A.W. Rowbottom, et al., Raman spectroscopic techniques to detect ovarian cancer biomarkers in blood plasma, Talanta 189 (2018) 281–288. Available from: https://doi.org/10.1016/j.talanta.2018.06.084.

[177] D.R. Parachalil, B. Brankin, J. McIntyre, H.J. Byrne, Raman spectroscopic analysis of high molecular weight proteins in solution—considerations for sample analysis and data pre-processing, Analyst 143 (2018) 5987–5998. Available from: https://doi.org/10.1039/C8AN01701H.

[178] R.A. Shaw, S. Kotowich, M. Leroux, H.H. Mantsch, Multianalyte serum analysis using mid-infrared spectroscopy, Ann. Clin. Biochem. 35 (1998) 624–632. Available from: https://doi.org/10.1177/000456329803500505.

[179] S.L. Haas, R. Müller, A. Fernandes, K. Dzeyk-Boycheva, S. Würl, J. Hohmann, et al., Spectroscopic diagnosis of myocardial infarction and heart failure by Fourier transform infrared spectroscopy in serum samples, Appl. Spectrosc. 64 (2010) 262–267. Available from: https://doi.org/10.1366/000370210790918508.

[180] L. Lechowicz, M. Chrapek, J. Gaweda, M. Urbaniak, I. Konieczna, Use of Fourier-transform infrared spectroscopy in the diagnosis of rheumatoid arthritis: a pilot study, Mol. Biol. Rep. 43 (2016) 1321–1326. Available from: https://doi.org/10.1007/s11033-016-4079-7.

[181] D.A. Scott, D.E. Renaud, S. Krishnasamy, P. Meriç, N. Buduneli, S. Cetinkalp, et al., Diabetes-related molecular signatures in infrared spectra of human saliva, Diabetol. Metab. Syndr. 2 (2010) 48. Available from: https://doi.org/10.1186/1758-5996-2-48.

[182] K.Z. Liu, R.A. Shaw, A. Man, T.C. Dembinski, H.H. Mantsh, Reagent-free, simultaneous determination of serum cholesterol in HDL and LDL by infrared spectroscopy, Clin. Chem. 48 (2002) 499–506.

[183] G. Sankari, E. Krishnamoorthy, S. Jayakumaran, S. Gunasekaran, V.V. Priya, S. Subramaniam, et al., Analysis of serum immunoglobulins using Fourier transform infrared spectral measurements, Biol. Med. 2 (2010) 1321–1326.

[184] D. Perez-Guaita, J. Ventura-Gayete, C. Pérez-Rambla, M. Sancho-Andreu, S. Garrigues, M. de la Guardia, Protein determination in serum and whole blood by attenuated total reflectance infrared spectroscopy, Anal. Bioanal. Chem. 404 (2012) 649–656. Available from: https://doi.org/10.1007/s00216-012-6030-7.

[185] B. Suh, S. Park, D.W. Shin, J.M. Yun, B. Keam, H.-K. Yang, et al., Low albumin-to-globulin ratio associated with cancer incidence and mortality in generally healthy adults, Ann. Oncol. 25 (2014) 2260–2266. Available from: https://doi.org/10.1093/annonc/mdu274.

[186] K. Spalding, F. Bonnier, C. Bruno, H. Blasco, R. Board, I. Benz-de Bretagne, et al., Enabling quantification of protein concentration in human serum biopsies using attenuated total reflectance-Fourier transform infrared (ATR-FTIR) spectroscopy, Vib. Spectrosc. 99 (2018) 50–58. Available from: https://doi.org/10.1016/j.vibspec.2018.08.019.

[187] A.J. Berger, T.-W. Koo, I. Itzkan, G. Horowitz, M.S. Feld, Multicomponent blood analysis by near-infrared Raman spectroscopy, Appl. Opt. 38 (1999) 2916–2926. Available from: https://doi.org/10.1364/AO.38.002916.

[188] D. Rohleder, G. Kocherscheidt, K. Gerber, W. Kiefer, W. Köhler, J. Möcks, et al., Comparison of mid-infrared and Raman spectroscopy in the quantitative analysis of serum, J. Biomed. Opt. 10 (2005) 031108. Available from: https://doi.org/10.1117/1.1911847.

[189] F. Bonnier, G. Brachet, R. Duong, T. Sojinrin, R. Respaud, N. Aubrey, et al., Screening the low molecular weight fraction of human serum using ATR-IR spectroscopy, J. Biophotonics 9 (2016) 1085–1097. Available from: https://doi.org/10.1002/jbio.201600015.

[190] D. Perez-Guaita, K.M. Marzec, A. Hudson, C. Evans, T. Chernenko, C. Matthäus, et al., Parasites under the spotlight: applications of vibrational spectroscopy to malaria research, Chem. Rev. 118 (2018) 5330–5358. Available from: https://doi.org/10.1021/acs.chemrev.7b00661.

[191] M. Martin, D. Perez-Guaita, D.W. Andrew, J.S. Richards, B.R. Wood, P. Heraud, Detection and quantification of *Plasmodium falciparum* in aqueous red blood cells by attenuated total reflection infrared spectroscopy and multivariate data analysis, J. Vis. Exp. 141 (2018) e56797. Available from: https://doi.org/10.3791/56797.

[192] E. Gray, H.J. Butler, R. Board, P.M. Brennan, A.J. Chalmers, T. Dawson, et al., Health economic evaluation of a serum-based blood test for brain tumour diagnosis: exploration of two clinical scenarios, BMJ Open 8 (2018) e017593. Available from: https://doi.org/10.1136/bmjopen-2017-017593.

[193] M. Koç, E. Karabudak, History of spectroscopy and modern micromachined disposable Si ATR-IR spectroscopy, Appl. Spectrosc. Rev. 53 (2018) 420–438. Available from: https://doi.org/10.1080/05704928.2017.1366341.

[194] D.K.R. Medipally, A. Maguire, J. Bryant, J. Armstrong, M. Dunne, M. Finn, et al., Development of a high throughput (HT) Raman spectroscopy method for rapid screening of liquid blood plasma from prostate cancer patients, Analyst 142 (2017) 1216–1226. Available from: https://doi.org/10.1039/C6AN02100J.

[195] C.A. Jenkins, R.A. Jenkins, M.M. Pryse, K.A. Welsby, M. Jitsumura, C.A. Thornton, et al., A high-throughput serum Raman spectroscopy platform and methodology for colorectal cancer diagnostics, Analyst 143 (2018) 6014–6024. Available from: https://doi.org/10.1039/C8AN01323C.

CHAPTER 11

Vibrational spectroscopy in protein research toward virus identification: challenges, new research, and future perspectives

Marfran C.D. Santos[1], Camilo L.M. Morais[2], Kássio M.G. Lima[1] and Francis L. Martin[2,3]

[1]Biological Chemistry and Chemometrics, Institute of Chemistry, Federal University of Rio Grande do Norte, Natal, Brazil, [2]School of Pharmacy and Biomedical Sciences, University of Central Lancashire, Preston, United Kingdom, [3]Biocel Ltd, Hull, United Kingdom

11.1 Introduction

One of the major public health problems is the large number of different types of viruses that circulate in our environment. From the point of view of clinical diagnosis, most of the time, there are no exemplar sensitive or specific techniques for each type of virus. This is because the more sensitive and specific the techniques are, the more expensive, complex, and time-consuming they become; thus inexpensive and less specific techniques become more prevalent often because of an ability to provide quick results [1]. One of the main current examples in this sense is related to cases of arboviruses, such as dengue, Zika, and chikungunya. These viruses are transmitted by the same mosquito and usually circulate in the same environments. In addition, they cause very similar clinical manifestations in the patient [2–6]. Another problem is that some viruses, such as Zika and dengue, belong to the same family, resulting in them having very similar surface proteins, which often confuse the immune system that produces antibodies to a wrong virus, aggravating the disease and making antibody-based diagnoses generate false positive results [7,8]. For this reason, specific techniques for these infectious agents need to be available.

Currently, there are no techniques that are, at the same time, specific, fast, cheap, and provide a diagnosis that differentiates dengue, Zika, and chikungunya infections in a single

Vibrational Spectroscopy in Protein Research.
DOI: https://doi.org/10.1016/B978-0-12-818610-7.00011-6

analysis. In Fig. 11.1, a comparison can be made of the merits of the main diagnostic methods. According to the World Health Organization (WHO), an optimal diagnosis for these viruses would be through three polymerase chain reactions (PCRs). That is, a patient who arrives in a clinic or hospital with suspected arbovirus infection would have their blood collected, and this sample would undergo a PCR for dengue, when in case of a negative result, a PCR for Zika would be performed, and then, in case of a second negative result, a PCR for chikungunya would be made [9,10]. However, this makes the diagnosis time-consuming and expensive.

For this reason, many efforts are made in the attempt to develop techniques that are capable, with a single analysis, of providing a reliable diagnosis that differentiates these viruses. In this chapter we suggest the use of vibrational spectroscopy techniques, such as infrared (IR) spectroscopy and Raman spectroscopy, as well as the study of spectral regions related to proteins, to detect the presence of certain viruses. Additionally, we will also discuss chemometric techniques that can be used in conjunction with spectroscopy and the difficulties that can be found in this field of research, as well as bring recent studies that have been made in this area and discuss the perspectives for the future in this field of research.

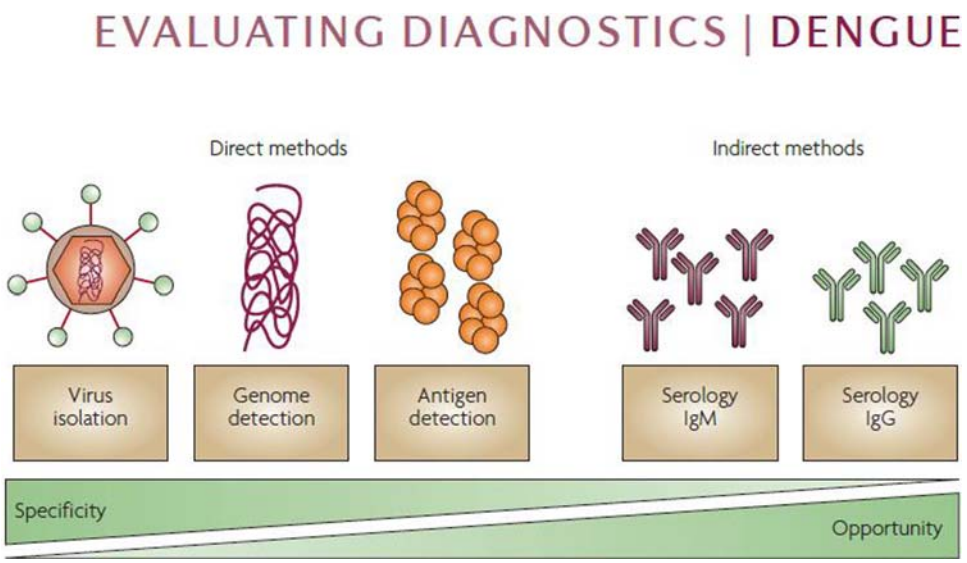


Figure 11.1

Comparison of the merits of direct and indirect laboratory methods for viral diagnosis. Opportunity refers to the feasibility of finding the method in the diagnostic routine. Source: *Taken from R.W. Peeling, H. Artsob, J.L. Pelegrino, P. Buchy, M.J. Cardosa, S. Devi, et al., Evaluation of diagnostic tests: dengue, Nat. Rev. Microbiol. 8 (2010) S30–S37 [1]. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature Publishing AG. Copyright © 2010.*

11.2 General structure of viruses

Today we know different types of viruses, from different families and genera, where each one has its own structural peculiarity. However, some similarities can be observed between them. One is that they are mostly made up of proteins. A general structure of a virus can be described as an outer lipid membrane, called a viral envelope (in the case of enveloped viruses); another inner membrane composed of a protein called capsid; and, finally, the genetic material (RNA or DNA) [11]. Since there are different types of viruses, in this topic we will focus on the Flaviviridae family viruses to talk about their structure. In this family, we will find viruses like dengue, Zika, and yellow fever virus.

Flaviviruses are small, encapsulated viruses that have three structural proteins (the glycoproteins of capsid (C), membrane (M), and envelope (E)) and seven nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5). The structure of a flavivirus is shown in Fig. 11.2. In its outer part, there are the proteins E and M, which are spicules encrusted in the viral envelope. The viral envelope is a lipid bilayer. Internally, we can find the nucleocapsid that is formed by protein and protects the viral genome [12–15].

When a person is infected by one of these viruses, the immune system kicks in. Antibodies are produced by the immune system based on the viral surface proteins (proteins E and M in Fig. 11.2, for example) [11]. Viral proteins are specific for each type of virus. For viruses

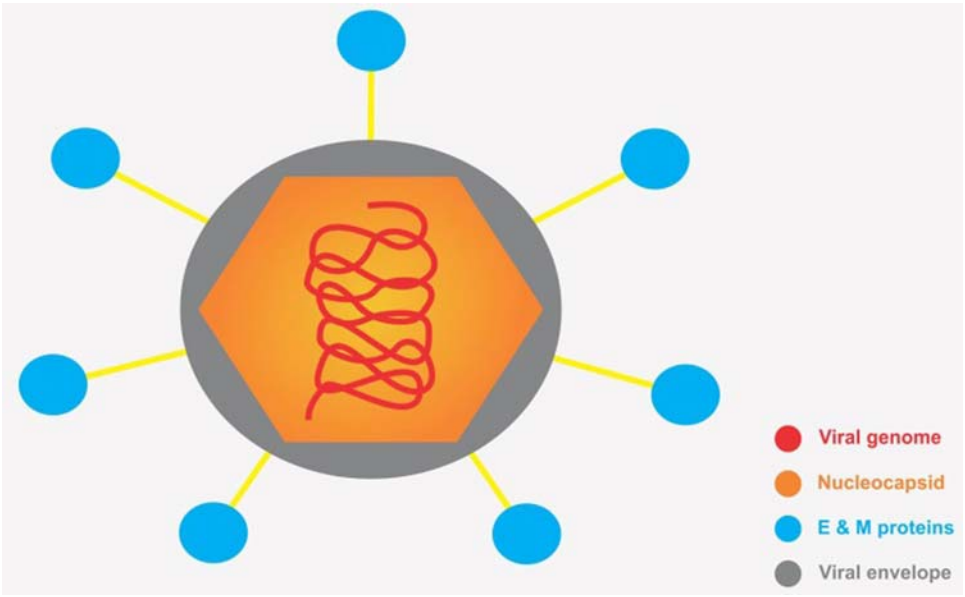


Figure 11.2
General structure of a flavivirus with its parts identified by colors.

from the same family, these surface proteins are very similar, but they still have small differences. The high degree of similarity between the surface proteins often confuses the immune system, which produces the wrong antibodies; for example, the patient has Zika, but the antibodies produced are for dengue. This aggravates the disease and causes false positive results for the serological diagnoses that are widely used in diagnostic routines.

Vibrational spectroscopy techniques are very susceptible to small variations in the chemical composition of samples and may arise as a new methodology in the virologic field for the identification and/or typing of viruses based on their respective proteins. We will now discuss three techniques of vibrational spectroscopy that can be used for this purpose.

11.3 A brief overview of vibrational biospectroscopy

Studies involving the use of spectroscopy in areas related to biology have been growing in the last decades. These approaches have given scope to the field of biospectroscopy. The spectroscopic techniques that will be discussed in this chapter, as already mentioned, will be IR spectroscopy [near-infrared (NIR) and mid-infrared (MIR)], and Raman spectroscopy. These techniques are the main ones in the field of vibrational spectroscopy and have been widely used in several studies in virology, which we will see in the course of this chapter.

11.3.1 Infrared spectroscopy

The frequency of IR radiation is of the same order as the frequency with which the molecules vibrate. For this reason, it is possible to obtain chemical information when certain molecules interact with radiation of this energy. The active biomolecules in IR are those that have an electric dipole moment that can change during the vibration. These biomolecules will vibrate when they interact and absorb IR radiation, making it possible to obtain a spectrum of the samples in which these molecules are present [16].

11.3.1.1 Mid-infrared

The MIR comprises electromagnetic radiation in the range of 4000−400 cm^{-1}. Spectra in this range can be obtained through transmission or reflectance modes, such as using attenuated total reflection Fourier-transform infrared (ATR-FTIR) spectroscopy. In ATR-FTIR spectroscopy, the sample (e.g., a drop of 10 μL of sample) is placed on the ATR crystal (e.g., a diamond of 250 × 250 μm), where the radiation is fully reflected, generating an evanescent wave capable of penetrating the sample by a few micrometers (see Fig. 11.3). The sample absorbs part of this radiation, and this absorption is detected by the equipment, which generates an absorbance spectrum of that sample [16,17].

There is a wide range of interest in MIR spectroscopy. The so-called fingerprint region for biological samples comprises the range between 1800 and 900 cm^{-1}. In this region it is

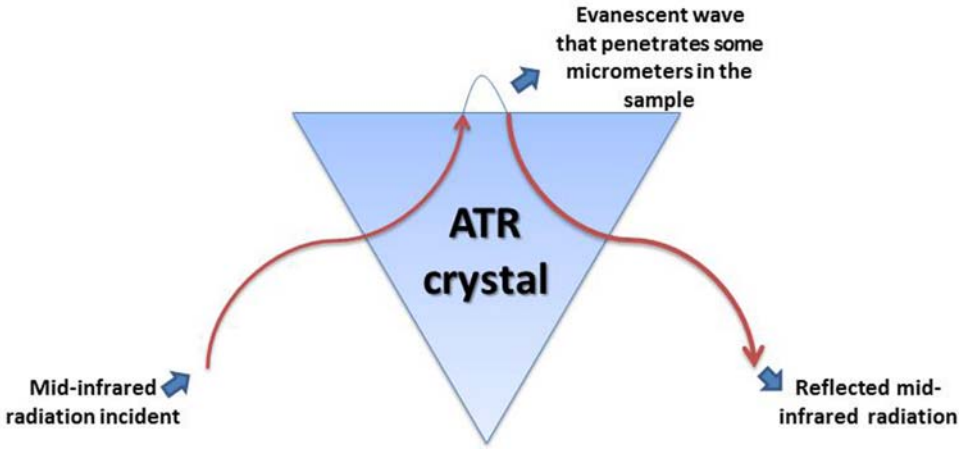


Figure 11.3
Incident radiation being reflected, generating an evanescent wave that penetrates the sample being analyzed.

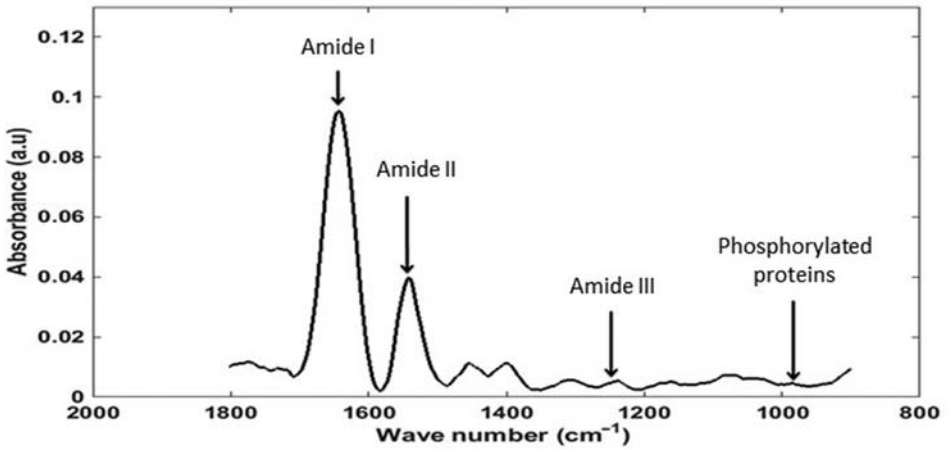


Figure 11.4
Mid-infrared (MIR) spectrum of the blood of a patient diagnosed with Zika virus. Bands related to protein structures are shown in the spectrum.

possible to find spectral bands related to biomolecules of interest when investigating viruses. Viruses are composed primarily of protein (structural and nonstructural proteins), lipids, and genetic material. In the fingerprint region of the MIR range, the bands related to proteins are 1650 cm^{-1} for amide I (C = C stretching), 1550 cm^{-1} for amide II (N−H bending coupled to C−N stretching), 1260 cm^{-1} for amide III (C−N stretching in proteins), and 960 cm^{-1} for symmetric stretching of phosphorylated proteins [16–19]. In Fig. 11.4 a

spectrum of the blood of a patient diagnosed with Zika is shown. It is possible to observe the peaks related to protein structures. Among the advantages of this technique, we can mention the small amount of sample required for analysis, rapid response, and nondestructive nature.

11.3.1.2 Near-infrared

NIR spectroscopy is a spectroscopic technique that uses NIR radiation, which covers the range of 700–2500 nm. It is a fast and multicomponent technique that does not require sample preparation or reagents. The absorption involved when biomolecules interact with NIR radiation is mainly related to overtones and combinations of stretching–bending vibrations of bonds containing hydrogen atoms, such as CH, OH, and NH. Overtones and stretching–bending combinations are types of vibrations that relate to the fundamental vibrations found in the MIR region: overtones will be observed in the second or third order of the fundamental vibration frequency, that is, about two or three times this frequency. By increasing the overtones frequency, the absorption intensity decreases. Stretching–bending combination bands are found at energies lower than overtones, and represent the sum of several fundamental absorptions of different types of vibrations [20].

Within the NIR region, we find the region called the "optical window," which comprises the range 650–1110 nm. Below 650 nm and above 1110 nm, there are overwhelming absorptions of water and hemoglobin. Several molecules present in biological samples are in sufficient quantities to have a NIR absorption signal, hence, making it possible to analyze biochemical changes in samples of different classes [20,21]. Among the advantages of this technique, we can cite the small amount of sample required for analysis, rapid response, low-cost instrumentation, and high availability of portable equipment.

11.3.2 Raman spectroscopy

Raman spectroscopy is a technique that has been also widely used in biospectroscopic studies. Normally, when a beam of light is incident on a substance, the molecules of that substance that are able to absorb in the wavelengths of that radiation are excited to a higher vibrational level, and when they return to their fundamental state, they emit radiation with the same frequency as that of the incident radiation (elastic effect), with no change in the net energy. However, in approximately 1 of 100,000 cases, Raman scattering occurs. In this phenomenon, molecules that absorb energy of a certain wavelength will emit radiation at a different wavelength range. This emitted radiation may be of higher or lower energy than that which has been absorbed, although in most cases the emitted energy will be less than that absorbed (longer wavelength). These events are related to the Stokes shift. The relationship between these high and low energy shifts can be detected and translated into what we call the Raman signal. This signal is proportional to the chemical structure and

concentration; thus it is possible to carry out quantitative and qualitative analyses using Raman spectroscopy [17,22].

Raman spectroscopy is a sensitive and noninvasive technique with a high capacity to provide information about the sample composition. In Fig. 11.5 a Raman spectrum is depicted with the peaks related to protein structures [17]. This information is useful for classification of virus-containing samples, using appropriate mathematical tools. Raman spectra can be used for virus investigation, whether for screening, determination of viral concentration, or diagnosis.

These vibrational spectroscopy techniques provide spectra with important information on the chemical composition of the samples being analyzed. These spectra have several variables (various absorbance or transmittance values, associated with certain wave number or wavelength values). In order to analyze these multivariate data, it is necessary to use appropriate mathematical tools. Preprocessing, as well as spectral cuts to use only regions of interest, are performed to improve the signal-to-noise ratio and remove undesired signals or interferences. After preprocessing, multivariate analysis and classification techniques are extremely important in the studies that involve the perspectives discussed in this chapter. In the next topic, we will present a brief introduction of the computational analysis methods that are used in studies containing virologic-related spectral data, from preprocessing techniques to multivariate analysis, so that the reader has an idea of how these tools work, and a brief theoretical support that will aid interpretation for future applications of biospectroscopy and chemometrics in the identification of viruses from protein-related spectral regions.

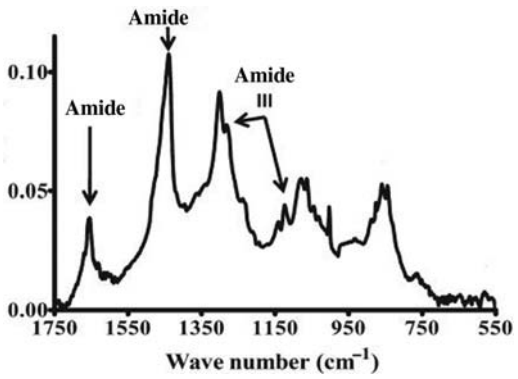


Figure 11.5

Assignment of protein-related peaks in a Raman spectrum. Source*: Adapted with permission from J.G. Kelly, J. Trevisan, A.D. Scott, P.L. Carmichael, H.M. Pollock, P.L. Martin- Hirsch, et al., Biospectroscopy to metabolically profile biomolecular structure: a multistage approach linking computational analysis with biomarkers, J. Proteome Res. 10 (2011) 1437–1448 [17]. Copyright © 2011 American Chemical Society.*

11.4 Computational analysis

After spectral acquisition, spectra are imported to some suitable software for further data processing. MATLAB (MathWorks, Inc., United States) is ones of the most popular software for spectral processing and multivariate analysis. Other options include The Unscrambler (Camo Analytics, Norway), Pirouette (InfoMetrix, Inc., United States), and free software such as GNU Octave (https://www.gnu.org/software/octave/) and R (https://www.r-project.org/). In this chapter, we will deal with the entire computational procedure, from spectral preprocessing to the classification techniques that can be used to investigate information that relates to proteins in viruses.

11.4.1 Preprocessing

A step of extreme importance after spectral acquisition is the preprocessing. Preprocessing is required to improve the signal-to-noise ratio of the data, for example, removing random errors, correcting baseline absorption, and correcting different sample thicknesses. These phenomena are related to the physical nature of the sample, instrument, or to the environment. Since they are not related to the chemical analyte of interest, they can be removed to filter only chemically relevant information for model construction. The first procedure to be used is the *spectral truncation*.

Spectral truncation is very common. The truncation is performed so that only spectral-relevant bands of the substances of interest are analyzed. If we analyze Figs. 11.4 and 11.5, for example, we will see that we can perform an even larger cut in the spectra, so that only the regions related to the proteins are considered. For the mid-IR spectrum (Fig. 11.4), we could make a cut in the range of 1800−1200 cm^{-1}; and for the Raman spectrum (Fig. 11.5), we could make a cut between 1750 and 1000 cm^{-1}. The user can still go beyond, and use only smaller spectral bands, for example, by cutting the spectrum of Fig. 11.4 in the region of 1700−1590 cm^{-1} to study only the band related to amide I. With this, as we shall see later, we will be speeding computational analysis and improving multivariate analysis outcomes, by increasing the chances of detecting differences in viral proteins in a virus-infected sample. However, note that by performing these cuts, we will be reducing the amount of information that will be analyzed later, thus this procedure must be performed carefully in order to not remove chemical-relevant information for the application of interest.

Another widely used preprocessing technique is *baseline correction.* Baseline correction is used for the removal of physical interferences. It is important to open a space to cite what causes these phenomena: it is believed that these effects of baseline mismatches occur when the wavelength of incident radiation is smaller or similar to some biochemical structures through which the light passes. This causes the radiation to spread, and this scattering is called Mie scattering. Consequently, underlying oscillations occur which result in absorption

spectra that do not truly represent the concentrations of the species in the samples [23,24]. Other causes of baseline distortion are reflection, temperature variation, or anomalies in the instrument, though whatever the cause, these effects may be partially or totally removed by baseline correction techniques. Several preprocessing techniques have been used to correct baseline abnormalities in MIR, NIR, and Raman spectra, such as rubberband baseline correction and automatic weighted least squares baseline correction [17,25–27].

Alternatively, to eliminate baseline problems, *differentiation* can be also used. Differentiation can be used both to correct baseline and to solve overlapping bands. Differentiation works to improve the differences between the spectral bands of interest and the existing baseline. Second-order differentiation provides symmetric characteristics consistent with what is expected of the absorption process. Two problems with this preprocessing are (1) after differentiation the absorbance scale is lost, thus the resultant signal has no physical–chemical meaning; and (2) that as the order of differentiation gets higher, the spectral noise increases by an order of magnitude [17,18,25–27].

Both baseline correction and differentiation are preprocessing techniques used to remove physical interference. Spectra can also be preprocessed in an attempt to remove random noise. A well-used preprocessing technique for smoothing is the *Savitzky–Golay (SG) smoothing* [28]. However, it is important to keep in mind that this technique can introduce spectral distortions and/or smooth out important spectral information. Therefore it is necessary to well adjust the smoothing parameters so that only the noise problem is solved without distorting the peaks. It is common to find techniques that smooth and calculate derivatives, such as the Norris–Williams derivative and the SG polynomial derivative [25,27].

Preprocessing can be also used to perform a *normalization* of a spectra set. Spectral variations between samples are usually related to differences in thickness and/or concentration. Often, this masks the biochemical differences of interest (e.g., structural differences related to viral proteins). These effects can be minimized when the spectral data set is scaled to a single specific criterion. Commonly, normalization can be performed by the highest peak, for example, for a set of MIR spectra with the same profile as that spectrum in Fig. 11.4, we could normalize by the peak of the amide I; and for a set of Raman spectra with the same profile of that of Fig. 11.5, we could normalize by the peak of the amide II. It is important to note that normalization by a particular peak should only be applied when such a peak is present across all spectra in the data set [17,18], and is not a discriminant feature. Several techniques have been employed in the normalization of MIR, NIR, and Raman spectra [17,18,25–27].

Table 11.1 summarizes these preprocessing techniques. It is important to note that to use spectroscopy in virologic analyses, if we want to analyze only proteins groups, the choice of correct preprocessing is extremely important, since if chosen in the wrong way, one can lose the information that is sought.

Table 11.1: List some preprocessing techniques for vibrational spectroscopy data (NIR, MIR, and Raman).

Purpose	Technique	References
Removal of baseline signal	Rubberband, automatic weighted least squares, or polynomial baseline correction 2nd differentiation	[16–27]
Normalization	To the amide I peak To the amide II peak Vector normalization	
Noise removal	Savitzky–Golay smoothing	

11.4.2 Multivariate analysis techniques

If we analyze a set of spectra of a certain type of sample (e.g., several blood spectra of healthy individuals and patients diagnosed with dengue), we would observe that they are extremely similar, making it impossible to visually distinguish between the healthy or disease spectra. For this, multivariate analysis techniques are generally used. Several techniques of multivariate analysis have been used, and in this section we will briefly present a few of these: principal component analysis (PCA), successive projections algorithm (SPA), genetic algorithm (GA), linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), support vector machines (SVM), and partial least squares discriminant analysis (PLS-DA).

PCA is a well-known exploratory analysis technique. PCA performs a linear transformation of the original data, generating new orthogonal variables called principal components (PCs). Each PC explains a portion of the original data variance, so that the first PC has the largest explained variance, the second PC has the second largest, and so on. In this process, PCA decomposes the preprocessed spectral matrix into three data matrices: scores, loadings, and residuals. The scores represent the variance on sample direction, being used to assess similarities/dissimilarities between the samples; and the loadings contain weights indicating the most important spectral variables in the construction of each component, being used to identify "biological markers" [29–31].

Another data reduction technique is SPA. In SPA the spectral variables are reduced to a few selected variables. For this, SPA treats each spectral variable as a vector. It starts with a variable (vector) and interacts with the respective projections of the other variables in the orthogonal space of that initial variable. This is performed in an iterative process until a certain number of selected variables is reached. In this way, the selected variables will be those that obtain the largest projections. As an advantage, SPA eliminates multicollinearity problems by removing redundant information from the data, while keeping the same original spectral space for the selected variables [32,33].

GA is another technique of variable selection. GA selects variables by a process that mimics Darwin's theory of evolution. In this technique, an initial population is randomly created, consisting of subsets of spectral variables called chromosomes. In a iterative process, these chromosomes are cross-combined, their offspring are randomly mutated, and the outcome is evaluated by a fitness function, where the selection criterion can be, for instance, a linear discriminant-based misclassification rate. Based on this, the chromosomes can be eliminated or copied until the number of generations ends. At the end, the set of variables (chromosome) with the best fitness performance is selected [34].

As mentioned previously, PCA, SPA, and GA are techniques that promote a reduction in data complexity, either by transforming the original spectral data to a small number of orthogonal factors (PCs), or by reducing them to some variables (SPA and GA) with great dissimilarity potential. It has been said that the loadings obtained by PCA can be treated as biological markers, where the most important chemical bonds for the construction of each component are shown. The variables selected by SPA and GA can also be treated as biological markers, since they are the main discriminant features within the data set, and they have the same spectral scale. From this point of view, we can imagine that, in a classification model where it is desired to distinguish samples between infected and noninfected, or between samples infected by different viruses and uninfected samples, we can obtain in the loadings of each PC, or in the variables selected by SPA and GA, information regarding viral proteins.

In order to classify samples into categories, we need a discriminant analysis (DA) technique. For this, LDA can be used. This technique is very robust and has been widely used in chemometrics. As the name suggests, LDA assigns samples into their respective classes based on a linear discriminant function calculated using a Mahalanobis distance between the classes. LDA works based on a pooled covariance matrix between the classes, thus assuming classes have similar variance structures. On the other hand, QDA, although also based on Mahalanobis distance calculation between the classes, considers classes having different variance structures, thus using a separated variance−covariance matrix for each class individually. LDA works better for data sets having a smaller number of samples and/or when the classes have similar variance distributions; while QDA works better for larger data sets and/or when the classes have different variance distributions [35,36].

LDA and QDA are limited to a respective linear or quadratic discriminant function applied in the input data space (e.g., PCA scores, SPA or GA selected variables); however, for complex data, a nonlinear classification may provide a better performance by using a SVM classifier. SVM transform the input data space, which can be the result of a data reduction technique or the whole preprocessed spectra, into a feature space where a linear classification is performed between the samples, according to a function fitted between the samples close to the classes boundaries, called support vectors. The key step of SVM is the

transformation of the input space into the feature space, which is performed by a Kernel function. There are many different types of Kernel functions, such as linear, quadratic, polynomial, and radial basis function (RBF), which are decisive in the SVM classification performance. RBF is the most indicated Kernel function because it is able to fit a transformation function according to the data distribution, maximizing the differences between the classes [37,38].

Another common technique for multivariate analysis of spectral data is PLS-DA [39]. PLS-DA combines both feature extraction and classification into a single technique. It is based on PLS decomposition of the preprocessed spectral data followed by a linear discriminant classifier; that is, it makes use of PLS to find a straight line that divides the data space into two regions, where each region is related to the space of each class [39]. In addition to biomarker detection, feature extraction or selection is employed together with classification techniques to solve problems of ill-conditioned data; since spectral data are ill-conditioned by nature, which highly affects the computation of the inverse matrix operation in LDA and QDA, for example. Therefore the spectral data must be reduced to a small number of variables ($\leq$number of observations), preferably orthogonal. Therefore these multivariate analysis techniques described herein have been often used as feature extraction/selection and classification techniques combined.

11.4.3 Performance evaluation

The quality of a classification model must be measured. Generally, the following figures of merit are used: sensitivity (SENS), representing the proportion of positives that are correctly classified; specificity (SPEC), representing the proportion of negatives that are correctly classified; positive predictive value (PPV), representing how many of the positives tests are true positives; negative predictive value (NPV), representing how many negative tests are true negatives; Youden's index (YOU), that evaluates the ability of a model to avoid errors; positive likelihood ratio (LR +), that represents the ratio between the probability of predicting a sample as positive when it is truly positive, and the probability of predicting a sample as positive when it is actually negative; and negative likelihood ratio (LR−), representing the ratio between the probability of predicting a sample as negative when it is actually positive, and the probability of predicting a sample as negative when it is truly negative [40]. The formulas for calculating these parameters are depicted in Table 11.2.

11.5 Applications

Several studies have been published using vibrational spectroscopy techniques in conjunction with chemometrics in virologic analysis. Applications using NIR spectroscopy are lesser in comparison with FTIR and Raman, where most of the work done using NIR in

Table 11.2: Classification quality parameters, where TP stands for true positives; TN for true negatives; FP for false positives; and FN for false negatives.

Parameter	Equation
Sensitivity (SENS)	$\text{SENS}(\%) = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100$
Specificity (SPEC)	$\text{SPEC}(\%) = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100$
Positive predictive value (PPV)	$\text{PPV}(\%) = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100$
Negative predictive value (NPV)	$\text{NPV}(\%) = \frac{\text{TN}}{\text{TN} + \text{FN}} \times 100$
Youden's index (YOU)	$\text{YOU} = \text{SENS} - (1 - \text{SPEC})$
Positive likelihood ratio (LR +)	$\text{LR} + = \frac{\text{SENS}}{1 - \text{SPEC}}$
Negative likelihood ratio (LR−)	$\text{LR} - = \frac{\text{SPEC}}{1 - \text{SENS}}$

the medical field is based on oxygenated or deoxygenated hemoglobin [20]. However, the advantages of NIR spectroscopy, such as its low-cost instrumentation and nondestructive nature, make this technique interesting for future clinical applications. A good example where NIR has been applied for virus diagnostics was published by Sakudo et al. [41], where this technique was used to identify human immunodeficiency virus type-1 in plasma. Sakudo et al. [42] also reported a study for the diagnosis of HIV serotype 1 (HIV-1) infection using molecular clones of various HIV-1 subtypes through NIR spectroscopy and PCA. In one of the steps of this study, the authors performed an ultracentrifugation of supernatant containing HIV-1 virus to analyze the NIR spectra of the supernatant before and after centrifugation (knowing that centrifugation will decrease the viral content). The main procedure in this stage was the use of 293 cells that were transinfected with pNL4-3 (pNL4-3 comprises pUC18 with an insert of HIV-1 subtype B gene), using cell cultures. The amount of HIV-I in the supernatant was determined by HIV-1 p24 antigen-capture enzyme-linked immunosorbent assay, confirming the reduction of viral content by 10%. The authors analyzed the changes in the NIR spectra before and after ultracentrifugation for HIV-1 removal. For three different culture supernatants, the acquisition of three NIR spectra was performed, totaling nine spectra. For NIR spectra acquisition, aliquots of 2 mL of each sample were placed in a polystyrene cuvette and the measurement was performed in a sealed system. The ambient temperature during spectral acquisition was 37°C, and the spectral resolution was 2 nm in the region of 600−1100 nm. The spectra were preprocessed by standard normal variate and smoothing using a window of 19 data points. PCA was then applied, where the scores and loadings were calculated. Based on the PCA loadings, the authors focused on differences in factors such as the interaction of viruses with the water in the medium. In fact, of the three techniques tested by the authors, NIR was the least selective and the hardest to determine with certainty what type of molecule the band is associated with, unlike Raman and MIR, for example.

Raman spectroscopy has been also widely used to investigate protein-related changes in viruses. This technique can act as structural probes of native viruses, precursor viral assemblies, and their constituent proteins and nucleic acids, being a powerful technique to observe macromolecular structural transformations that define the viral morphogenetic pathway, such as conformational rearrangements of viral proteins and nucleic acids and site-specific interactions of important molecular groups [43]. For example, Blanch et al. [44] reported Raman optical activity spectra to investigate secondary and tertiary structures of the polypeptide backbone, hydration, side chain conformation, and structural elements present in denatured protein states; and sugar ring conformation, base stacking arrangement, and the mutation orientation of the sugar and base rings around the C−N glycosidic linkage in nucleic acid. This is achieved by measuring small differences in vibrational Raman scattering from chiral molecules in right- and left-circularity polarized incident light. Salman et al. [45] evaluated the potential of Raman spectroscopy as a sensitive, rapid, and reliable method for the detection and identification of Herpes Simplex Virus (HSV-1) infections in cell culture. The Raman spectra were acquired after the whole process of cell culture, viral infection, and sample preparation. In a 5-day period, 22 uninfected and 21 HSV-1 infected cells were analyzed. The spectra were collected from three regions: cell center, membrane-rich sites, and sites within the cell border. In the spectral data obtained, PCA and LDA were used. The plot of PCA scores on PC1 versus PC2 for Vero cell spectral data with HSV-1 and Vero without HSV showed a significant discriminant pattern with excellent class segregation, proving that when LDA is applied, perfect results are obtained. In this study, the authors made several truncations in the spectra, so that they applied PCA-LDA in the uncut spectra and in seven truncated regions: 1195−1726 cm^{-1} (region that relates to lipid proteins and nucleic acid); 600−1995 cm^{-1} (region that is related to carbohydrates); 1195−1380 cm^{-1} (relates to amide III); 1380−1520 cm^{-1} (relates to nucleic acids); 1520−1726 cm^{-1} (relates to amide I and II); 1520−1626 cm^{-1} (relates to amide II); and 1626−1726 cm^{-1} (relates to amide I). This study showed that the best results, that is, the region with best classification among infected or uninfected HSV-1 samples, was the region between 1195−1380 cm^{-1}, which is associated with amide III in proteins. This finding demonstrates that the technique was able to find spectral difference caused by the presence of viral proteins, achieving 100% correct classification [45].

Santos et al. [46] used MIR spectroscopy and multivariate analysis techniques to classify healthy and infected blood samples with different concentrations of Dengue 3 serotype (Denv-3). In this study, ATR-FTIR spectroscopy was used for the acquisition of the MIR spectra, and multivariate classification was performed by means of PCA-LDA, SPA-LDA, and GA-LDA algorithms. The sample set consisted of 40 samples, 10 samples for each Denv-3 concentration in blood. The spectra were collected in the region of 4000−400 cm^{-1}, and then truncated in the region 1800−900 cm^{-1}. The spectra were separated into training (60%), validation (20%), and test (20%) sets using the

Kennard–Stone sample selection algorithm. Preprocessing was carried out by SG smoothing and baseline correction. When applying the PCA, the authors selected five PCs that explained 95% of the total data variance. The scores on these PCs were used for LDA classification. SPA selected 19 wave numbers, while GA selected 11 (notice the significant reduction of data by PCA, SPA, and GA). Among the wave numbers selected by SPA there are carbohydrate ($\approx 1151\ cm^{-1}$), protein structures (amide I, $\approx 1653\ cm^{-1}$, amide III, $\approx 1317\ cm^{-1}$), and RNA ($\approx 1227\ cm^{-1}$, $\approx 1076\ cm^{-1}$) vibrations. Among the wave numbers selected by GA there are RNA ($\approx 1238\ cm^{-1}$), protein structures (amide III, $\approx 1329\ cm^{-1}$, amide I, $\approx 1661\ cm^{-1}$), and lipid ($\approx 1743\ cm^{-1}$) vibrations. Both variable selection techniques selected spectral regions related to protein structures. Based on these selected variables, LDA was then applied. The best results obtained in the classification were for SPA-LDA and GA-LDA. In Fig. 11.6 we can see the two-dimensional plots of the discriminant functions obtained by SPA-LDA and GA-LDA, where an excellent class separation is observed.

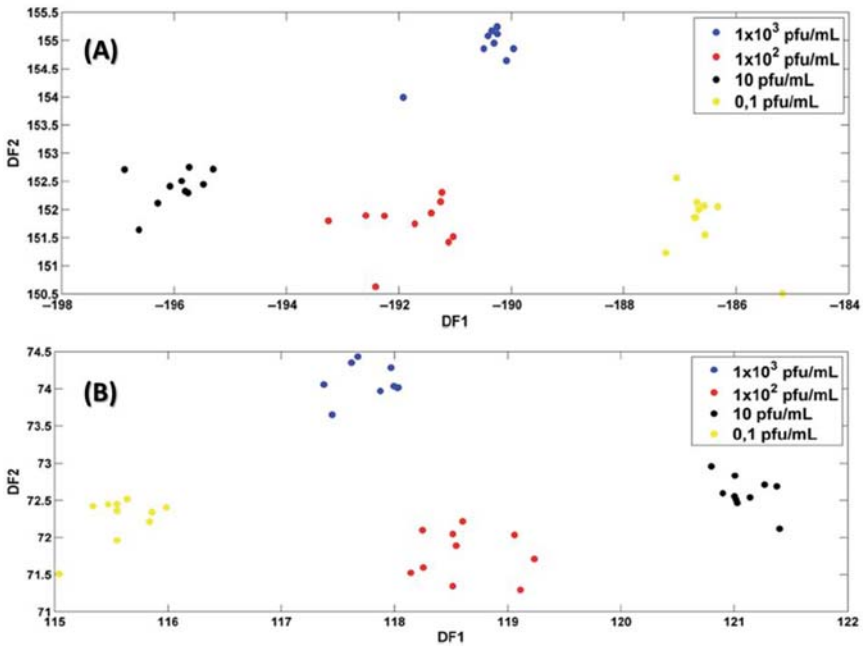


Figure 11.6

(A) DF1 × DF2 discriminant function values calculated with the variables selected by SPA-LDA for DENV-3 in blood (•1×10^3, •1×10^2, •10, and •0.1 PFU/mL). (B) DF1 × DF2 discriminant function values calculated with the variables selected by GA-LDA for DENV-3 in blood (•1×10^3, •1×10^2, •10, and •0.1 PFU/mL). Source: *Adapted from M.C.D. Santos, Y.M. Nascimento, J.M.G. Araújo, K.M.G. Lima, ATR-FTIR spectroscopy coupled with multivariate analysis techniques for the identification of DENV-3 in different concentrations in blood and serum: a new approach, RSC Adv. 7 (2017) 25640–25649 [46]. Published by the Royal Society of Chemistry.*

The figures of merit demonstrated 100% classification efficiency. That is, based on the variables selected by SPA and GA, LDA was able to predict with 100% accuracy which concentrations of Denv-3 were present in the blood samples. Knowing that many of the selected variables were related to protein structures, we can suggest that the technique was able to identify the differences in the relative amounts of the viral proteins and, based on this, to classify the samples correctly.

Santos et al. [47] recently published another study involving ATR-FTIR spectroscopy with multivariate classification techniques in virus investigation. In the study, they used MIR spectroscopy in conjunction with PCA-LDA, SPA-LDA, and GA-LDA to classify clinical samples as Dengue (blood samples from patients diagnosed with dengue), Chikungunya (blood samples from patients diagnosed with chikungunya), Zika (blood samples from patients diagnosed with Zika), and Healthy (blood samples from healthy volunteers). This study used 45 blood samples from patients diagnosed with dengue, 30 blood samples from patients diagnosed with Zika, 10 blood samples of patients diagnosed with Zika, and 45 blood samples from healthy individuals (not infected with any of these viruses). A total of 130 samples were used, where from each sample, two spectra were collected, totaling 260 spectra. The infections were confirmed by reverse transcriptase PCR assay and samples from healthy volunteers were also analyzed. The spectra were cut between 1800 and 900 cm^{-1}, followed by baseline correction and smoothing. The two highest bands in the preprocessed spectra were found to be related to proteins. The samples were divided into 60% for training (156 samples), 20% for validation (52 samples), and 20% for test (52 samples) using the Kennard–Stone algorithm. PCA was applied using three PCs accounting for 95% of the data explained variance. Based on the scores of these components, LDA was applied, generating 100% sensitivity and specificity results for the Healthy, Dengue, and Chikungunya groups; and 92% for the Zika group. The few errors related to blood samples from patients with Zika are most likely related to the similarity of the surface proteins of Zika and dengue viruses, since viruses of the same family have similar surface proteins. SPA selected only four variables, where we can highlight the wave number 1641 cm^{-1} which is related to the amide II protein. This again indicates that the technique was able to identify the small spectral variations caused by the presence of different viruses in blood (or absence of these viruses in the case of healthy controls). When applying LDA to the variables selected by SPA, 100% sensitivity and specificity were also obtained for the Healthy, Dengue, and Chikungunya groups; and 92% for the Zika group. Since in the four variables selected by SPA there is a protein-related variable, it is believed that the errors related to the prediction of Zika samples are related to the similarity of Zika and dengue proteins. GA selected 56 variables, where two protein-related variables were observed (amide I at 1627 cm^{-1}, and amide II at 1551 cm^{-1}). Based on the variables selected by GA, LDA was applied in the discrimination of the four classes, where 100% sensitivity

and specificity for Healthy, Dengue and Chikungunya classes were obtained; and 92% of sensitivity and 86% of specificity for the Zika class. Likewise, the errors related to Zika are most likely related to the fact that Zika and dengue belong to the same family.

Other applications where IR spectroscopy has been successfully used to investigate viral proteins include the assessment of modifications in viral proteins' profiles induced by a cationic porphyrinic photosensitizer, where toxic reactive oxygen species produced during this process were found to affect phage proteins [48]; monitoring structural changes in hepatitis C virus core protein by interactions of protein and nucleic acid components [49]; and structural analysis and mapping of individual protein complexes in tobacco mosaic virus [50].

In many cases, the results of sensitivity and specificity obtained in these studies are superior to some techniques used as standard in clinical settings, demonstrating that the approach of spectroscopy in conjunction with chemometric techniques has great potential to be used in virus detection routines, where protein-related bands are of great importance.

11.6 Challenges

The results obtained in studies using vibrational spectroscopy in conjunction with chemometric techniques in virus identification have demonstrated enormous potential for the future application of these techniques in virologic routine analysis. However, of course, much still needs to be developed until these applications can be translated to the clinical theater. Below are some of the difficulties that can be found in this field of research, and we come up with some answers as to what path one can take to solve these problems.

- One of the difficulties that can be found in this field of research is the fact that there is still little information about spectral measurements of viral proteins. For example, how can one identify the chemical structures of these proteins based on spectral data? It may be necessary to isolate the proteins from each virus so that a spectral database of these isolated proteins can be assembled. Thus with PCA, for example, it is possible to evaluate whether the loadings on some PC applied to clinical samples with dengue have the same profile of the protein spectrum isolated from dengue. This would be an important finding to ascertain whether the PCA-LDA classification, in this case, is based on these viral protein structures, and not on any other protein structure of the blood. This could be a way to solve one of the major problems of this area, which is to confirm that those peaks, or selected variables, are related to virus proteins, not to other proteins present in the medium. Another difficulty that could be solved by this procedure would be in cases of coinfections, since it is possible that one patient is infected at the same time by more than one virus. In this case, more than one analytical

technique and/or biochemical method should be analyzed together to obtain correlation between results and narrow down the results to a more precise diagnosis.
- In the case of variable selection techniques, a procedure that could be interesting, knowing the profile of the isolated viral protein spectra, would be performing the multivariate analysis only in the regions with the highest absorbance intensities of protein-related bands. That is, specifying the region where it is known that this chemical species has a considerable absorbance.
- Another problem is the variability inherent in each individual. Different people, with different lifestyles, have biochemical differences in their blood, hence, resulting in spectral differences. How to make the classification based only on viral information and not on person-to-person variation? One way to solve this is sampling. A large database is required so that numerous samples enter the training set and all of these variations are included in the control group. In this way, the supervised training model would be more robust and selective to a certain type of virus.

The fact is that the results generally found in this field of study are encouraging, however, many studies need to be developed until we can apply vibrational spectroscopy techniques for the purpose described in this chapter. It is necessary to performed massive inter- and intralaboratorial tests to reach a conclusion for practical use. However, one thing is sure, the more we investigate this area, the closer we will be to an actual application of these techniques in routine virologic analysis.

11.7 Future perspectives

As described in this chapter, the diagnostic, detection, or virus typing techniques used today have some advantages and disadvantages. Direct methods are very sensitive and specific; however, they are expensive, complex, and time-consuming. On the other hand, indirect methods are usually faster and cheaper, but are less specific and can give wrong results. With this in mind, vibrational spectroscopy techniques, such as MIR, NIR, or Raman, appear to have the potential to address these shortcomings in standard techniques. Vibrational spectroscopic techniques are known for their rapid response and the ability to provide reliable information about sample composition, being used in many biological applications with near-future prospects in the field of virology.

As we know, spectroscopic techniques analyze the response of certain samples when they interact with a specific radiation frequency. With this in mind, we can imagine that, in the near future, with new studies and advances in techniques and technology, we can count on spectroscopic tools in clinics, hospitals, and health centers functioning as a reliable tool to aid virus screening or even diagnosis. For this, the patient suspected of having a virus infection would have a minimal amount of biofluid collected (remember that only a few microliters is needed), this biofluid would undergo a spectral acquisition on an instrument

connected to a computer where this spectral information would automatically pass through software that would perform the multivariate analysis, providing an almost real-time result. Another interesting point to keep in mind is that, in the case of NIR instruments, there is the possibility of manufacturing portable tools, which would facilitate *in loco* analysis, for example, in surveillance studies.

Looking not only from the standpoint of diagnostic routines, but also of routines involving centers for virologic studies, we can also imagine a great application of vibrational spectroscopy. The techniques used in research centers are, for the most part, techniques that are within the so-called direct methods. They are very specific and sensitive tools, but very expensive and relatively time-consuming. For example, PCR is generally used for diagnosis, virus typing, or viral load determination. However, for each PCR made, primers are needed for each type of virus or serotypes of the same virus. This makes this technique rather expensive and laborious, especially when compared to the spectroscopic techniques described in this chapter. The rationale is pretty much the same. Since the spectroscopic techniques are faster and easier to use, they do not require wet-chemistry reactant kits, thus enhancing and optimizing daily routines. Finally, in view of the perspectives discussed here, we can affirm that vibrational spectroscopy has a great potential to be used in virologic studies, both for its characteristics of providing reliable information about the chemical composition of samples, as well as for the low-cost instrumentation, nondestructive nature, and fast analytical frequency.

References

[1] R.W. Peeling, H. Artsob, J.L. Pelegrino, P. Buchy, M.J. Cardosa, S. Devi, et al., Evaluation of diagnostic tests: dengue, Nat. Rev. Microbiol. 8 (2010) S30–S37.

[2] M.G. Guzman, S.B. Halstead, H. Artsob, P. Buchy, J. Farrar, D.J. Gubler, et al., Dengue: a continuing global threat, Nat. Rev. Microbiol. 8 (2010) S7–S16.

[3] S. Bhatt, P.W. Gething, O.J. Brady, J.P. Messina, A.W. Farlow, C.L. Moyes, et al., The global distribution and burden of dengue, Nature 496 (2013) 504–507.

[4] A.R. Plourde, E.M. Bloch, A literature review of Zika virus, Emerg. Infect. Dis. 22 (2016) 1185–1192.

[5] D. Musso, D.J. Gubler, Zika virus, Clin. Microbiol. Rev. 29 (2016) 487–524.

[6] A.L. Presti, A. Lai, E. Cella, G. Zehender, M. Ciccozzi, Chikungunya virus, epidemiology, clinics and phylogenesis: a review, Asian Pac. J. Trop. Med. 7 (2014) 925–932.

[7] W. Dejnirattisai, P. Supasa, W. Wongwiwat, A. Rouvinski, G. Barba-Spaeth, T. Duangchinda, et al., Dengue virus sero-cross-reactivity drives antibody-dependent enhancement of infection with Zika virus, Nat. Immunol. 17 (2016) 1002–1008.

[8] A.B. Kawiecki, R.C. Christofferson, Zika-induced antibody response enhances dengue serotype 2 replication *in vitro*, J. Infect. Dis. 214 (2016) 1357–1360.

[9] Pan American Health Organization, Tool for the Diagnosis and Care of Patients With Suspected Arboviral Diseases, World Health Organization, 2017.

[10] N. Boonham, J. Kreuze, S. Winter, R.V. Vlugt, J. Bergervoet, J. Tomlinson, et al., Methods in virus diagnostics: from ELISA to next generation sequencing, Virus Res. 186 (2014) 20–31.

[11] A.J. Cann, Principles of Molecular Virology, fourth ed., Elsevier Academic Press, Burlington, 2005.

[12] Y. Modis, S. Ogata, D. Clements, S.C. Harrison, Structure of the dengue virus envelope protein after membrane fusion, Nature 427 (2004) 313–319.
[13] S. Mukhopadhyay, R.J. Kuhn, M.G. Rossmann, A structural perspective of the flavivirus life cycle, Nat. Rev. Microbiol. 3 (2005) 13–22.
[14] F.A. Rey, K. Stiasny, F.X. Heinz, Flavivirus structural heterogeneity: implications for cell entry, Curr. Opin. Virol. 24 (2017) 132–139.
[15] M. Bollati, K. Alvarez, R. Assenberg, C. Baronti, B. Canard, S. Cook, et al., Structure and functionality in flavivirus NS-proteins: perspectives for drug design, Antivir. Res. 87 (2010) 125–148.
[16] F.L. Martin, J.G. Kelly, V. Llabjani, P.L. Martin-Hirsch, I.I. Patel, J. Trevisan, et al., Distinguishing cell types or populations based on the computational analysis of their infrared spectra, Nat. Protoc. 5 (2010) 1748–1760.
[17] J.G. Kelly, J. Trevisan, A.D. Scott, P.L. Carmichael, H.M. Pollock, P.L. Martin- Hirsch, et al., Biospectroscopy to metabolically profile biomolecular structure: a multistage approach linking computational analysis with biomarkers, J. Proteome Res. 10 (2011) 1437–1448.
[18] J. Trevisan, P.P. Angelov, P.L. Carmichael, A.D. Scott, F.L. Martin, Extracting biological information with computational analysis of Fourier-transform infrared (FTIR) biospectroscopy datasets: current practices to future perspectives, Analyst 137 (2012) 3202–3215.
[19] M.J. Baker, J. Trevisan, P. Bassan, R. Bhargava, H.J. Butler, K.M. Dorling, et al., Using Fourier transform IR spectroscopy to analyze biological materials, Nat. Protoc. 9 (2014) 1771–1791.
[20] A. Sakudo, Y. Suganuma, T. Kobayashi, T. Onodera, K. Ikuta, Near-infrared spectroscopy: promising diagnostic tool for viral infections, Biochem. Biophys. Res. Commun. 341 (2006) 279–284.
[21] A. Sakudo, Near-infrared spectroscopy for medical applications: current status and future perspectives, Clin. Chim. Acta 455 (2016) 181–188.
[22] P.J. Lambert, A.G. Whitman, O.F. Dyson, S.M. Akula, Raman spectroscopy: the gateway into tomorrow's virology, Virol. J. 3 (2006) 51.
[23] P. Bassan, H.J. Byrne, F. Bonnier, J. Lee, P. Dumas, P. Gardner, Resonant Mie scattering in infrared spectroscopy of biological materials – understanding the 'dispersion artefact', Analyst 134 (2009) 1586–1593.
[24] P. Bassan, A. Kohler, H. Martens, J. Lee, E. Jackson, N. Lockyer, et al., RMieS-EMSC correction for infrared spectra of biological cells: extension using full Mie theory and GPU computing, J. Biophotonics 3 (2010) 609–620.
[25] P. Lasch, Spectral pre-processing for biomedical vibrational spectroscopy and microspectroscopic imaging, Chemometr. Intell. Lab. Syst. 117 (2012) 100–114.
[26] N.K. Afseth, V.H. Segtnan, J.P. Wold, Raman spectra of biological samples: a study of preprocessing methods, Appl. Spectrosc. 60 (2006) 1358–1367.
[27] A. Rinnan, F. van den Berg, S.B. Engelsen, Review of the most common pre-processing techniques for near-infrared spectra, Trends Analyt. Chem. 28 (2009) 1201–1222.
[28] A. Savitzky, M.J.E. Golay, Smoothing and differentiation of data by simplified least squares procedures, Anal. Chem. 36 (1964) 1627–1639.
[29] S. Wold, Principal component analysis, Chemometr. Intell. Lab. Syst. 2 (1987) 37–52.
[30] H. Abdi, L.J. Williams, Principal component analysis, WIREs Comput. Stat. 2 (2010) 433–459.
[31] R. Bro, A.K. Smilde, Principal component analysis, Anal. Meth. 6 (2014) 2812–2831.
[32] M.C.U. Araújo, T.C.B. Saldanha, R.K.H. Galvão, T. Yoneyama, H.C. Chame, V. Visani, The successive projections algorithm for variable selection in spectroscopic multicomponent analysis, Chemometr. Intell. Lab. Syst. 57 (2001) 65–73.
[33] M.J.C. Pontes, R.K.H. Galvão, M.C.U. Araújo, P.N.T. Moreira, O.D.P. Neto, G.E. José, et al., The successive projections algorithm for spectral variable selection in classification problems, Chemometr. Intell. Lab. Syst. 78 (2005) 11–18.
[34] D. Broadhurst, R. Goodacre, A. Jones, J.J. Rowland, D.B. Kell, Genetic algorithms as a method for variable selection in multiple linear regression and partial least squares regression, with applications to pyrolysis mass spectrometry, Anal. Chim. Acta 348 (1997) 71–86.

[35] C.L.M. Morais, K.M.G. Lima, Principal component analysis with linear and quadratic discriminant analysis for identification of cancer samples based on mass spectrometry, J. Braz. Chem. Soc. 29 (2018) 472–481.
[36] W. Wu, Y. Mallet, B. Walczak, W. Penninckx, D.L. Massart, S. Heuerding, et al., Comparison of regularized discriminant analysis, linear discriminant analysis and quadratic discriminant analysis, applied to NIR data, Anal. Chim. Acta 329 (1996) 257–265.
[37] C. Cortes, V. Vapnik, Support-vector networks, Mach. Learn. 20 (1995) 273.
[38] C.L.M. Morais, F.S.L. Costa, K.M.G. Lima, Variable selection with a support vector machine for discriminating *Cryptococcus* fungal species based on ATR-FTIR spectroscopy, Anal. Meth. 9 (2017) 2964–2970.
[39] R.G. Brereton, G.R. Lloyd, Partial least squares discriminant analysis: taking the magic way, J. Chemometrics 28 (2014) 213–225.
[40] M.C.D. Santos, C.L.M. Morais, Y.M. Nascimento, J.M.G. Araujo, K.M.G. Lima, Spectroscopy with computational analysis in virological studies: a decade (2006-2016), Trends Analyt. Chem. 97 (2017) 244–256.
[41] A. Sakudo, R. Tsenova, T. Onozuka, K. Morita, S. Li, J. Warachit, et al., A novel diagnostic method for human immunodeficiency virus type-1 in plasma by near-infrared spectroscopy, Microbiol. Immunol. 49 (2005) 695–701.
[42] A. Sakudo, Y. Suganuma, R. Sakima, K. Ikuta, Diagnosis of HIV-1 infection by near-infrared spectroscopy: analysis using molecular clones of various HIV-1 subtypes, Clin. Chim. Acta 413 (2012) 467–472.
[43] D. Němeček, G.J. Thomas Jr., Chapter 16 – Raman spectroscopy of viruses and viral proteins, in: J. Laane (Ed.), Frontiers of Molecular Spectroscopy, Elsevier, 2009, pp. 553–595.
[44] E.W. Blanch, L. Hecht, L.D. Barron, Vibrational Raman optical activity of proteins, nucleic acids, and viruses, Methods 29 (2003) 196–209.
[45] A. Salman, E. Shufan, L. Zeiri, M. Huleihel, Characterization and detection of Vero cells infected with Herpes Simplex virus type 1 using Raman spectroscopy and advanced statistical methods, Methods 68 (2014) 364–370.
[46] M.C.D. Santos, Y.M. Nascimento, J.M.G. Araújo, K.M.G. Lima, ATR-FTIR spectroscopy coupled with multivariate analysis techniques for the identification of DENV-3 in different concentrations in blood and serum: a new approach, RSC Adv. 7 (2017) 25640–25649.
[47] M.C.D. Santos, Y.M. Nascimento, J.D. Monteiro, B.E.B. Alves, M.F. Melo, A.A.P. Paiva, et al., ATR-FTIR spectroscopy with chemometric algorithms of multivariate classification in the discrimination between healthy vs. dengue vs. chikungunya vs. zika clinical samples, Anal. Meth. 10 (2018) 1280–1285.
[48] L. Costa, A.C. Esteves, A. Correia, C. Moreirinha, I. Delgadillo, A. Cunha, et al., SDS-PAGE and IR spectroscopy to evaluate modifications in the viralprotein profile induced by a cationic porphyrinic photosensitizer, J. Virol. Meth. 209 (2014) 103–109.
[49] P. Carmona, M. Molina, Interactions of protein and nucleic acid components of hepatitis C virus as revealed by Fourier transform infrared spectroscopy, Biochemistry 49 (2010) 4724–4731.
[50] I. Amenabar, S. Poly, W. Nuansing, E.H. Hubrich, A.A. Govyadinov, F. Huth, et al., Structural analysis and mapping of individual protein complexes by infrared nanospectroscopy, Nat. Commun. 4 (2013) 2890.

This page intentionally left blank

CHAPTER 12

Two-dimensional correlation spectroscopy of proteins

Yeonju Park[1], Isao Noda[2] and Young Mee Jung[1]

[1]Department of Chemistry, and Institute for Molecular Science and Fusion Technology, Kangwon National University, Chunchon, Korea, [2]Department of Materials Science and Engineering, University of Delaware, Newark, DE, United States

12.1 Introduction

Two-dimensional correlation spectroscopy (2D-COS) is a versatile technique applicable to the in-depth analysis of spectral data obtained under the influence of external perturbations in various experiments [1−36]. It has been well-accepted as a powerful analytical technique in many fields of spectroscopic studies, which can provide new insights at the molecular level into the system understanding [1−36]. Among many applications, proteins are one of the most popularly studied samples in 2D-COS [30−151].

In this chapter, we firstly introduce the background of generalized 2D-COS method and then review its specific applications in protein studies. Various analytical techniques, such as infrared (IR), Raman, fluorescence, Raman optical activity (ROA), and vibrational circular dichroism (VCD) spectroscopies, have been successfully applied in conjunction with 2D-COS in protein studies. Advanced methods, such as hetero two-dimensional (2D) correlation analysis [68,89,102,104−107,116,128,152] and 2D-COS combined with principal component analysis (PCA) [45,81,113−122], have also been performed for better understanding of proteins.

12.2 Background

12.2.1 Generalized two-dimensional correlation spectroscopy

The original concept of 2D-COS was firstly introduced by Noda in 1986 for investigating dynamic IR linear dichroism spectra of polymer films under the influence of a small-amplitude sinusoidal strain [2]. Noda then expanded this initial 2D-COS concept to a more general form of spectroscopic approach in 1993 that has become a much more useful

Vibrational Spectroscopy in Protein Research.
DOI: https://doi.org/10.1016/B978-0-12-818610-7.00012-8

analytical technique applicable to various fields of spectroscopic studies [1]. The generalized 2D-COS is a versatile technique to sort out important information in spectral variations observed with a various probe under a variety of applied external perturbations. It can often provide very interesting results, which are sometimes hardly detected in one-dimensional (1D) spectral analysis.

2D-COS can be applied to a broad range of spectroscopic probes, for example, IR, near-infrared (NIR), Raman, ultraviolet-visible (UV−Vis), VCD, ROA, fluorescence, THz, mass spectrometry (MS), X-ray, and nuclear magnetic resonance (NMR), as well as those from microscopy or chromatography studies. In 2D-COS, an external perturbation, such as chemical, thermal, biological, mechanical, electrical, and many more, is applied to the system of interest inducing spectral variations. A variety of materials and systems, such as biomolecules including proteins, peptides, and nucleic acids, polymers, nanomaterials and composites, pharmaceuticals and medicals, natural materials, food science, and environmental science have been extensively studied by 2D-COS.

12.2.2 Basic concept of two-dimensional correlation spectroscopy

The detailed mathematical treatment for generalized 2D-COS has been already well described [1−6]. Here, only the basic concept is briefly described. 2D-COS is a broadly applicable tool developed to interpret a set of spectral changes for a system under the influence of an external perturbation. In generalized 2D-COS, a series of so-called dynamic spectra exhibiting perturbation-induced intensity changes are collected systematically as shown in Fig. 12.1. Suppose a set of spectra $A(\nu_j, t_i)$ is obtained as a function of the spectral variable ν_j with $j = 1, 2, \ldots n$ and some perturbation variable t_i with $i = 1, 2, \ldots m$ during a well-defined observation interval between t_1 and t_m. The spectral variable ν can be selected from many different types, including IR wave number, Raman shift, X-ray diffraction angle, or UV wavelength, depending on the specific analytical probe used for the study. Perturbation variable t can also be one of many reasonable measures of the effect of perturbations, such as temperature, time, composition, pressure, or even electric potential in some applications. A dataset collected using the combination of such a probe and perturbation can be readily transformed into 2D correlation spectra based on a simple cross correlation analysis.

The *dynamic spectrum* $\tilde{A}(\nu_j, t_i)$ of a system, that is, spectral intensity changes induced by the application of an external perturbation within the finite observation interval between t_1 and t_m is formally defined as

$$\tilde{A}(\nu_j, t_i) = A(\nu_j, t_i) - \overline{A}(\nu_j) \tag{12.1}$$

where $\overline{A}(\nu_j)$ is the *reference spectrum* of the system. For typical applications, the reference spectrum may be set as the *stationary* or *averaged spectrum* given by

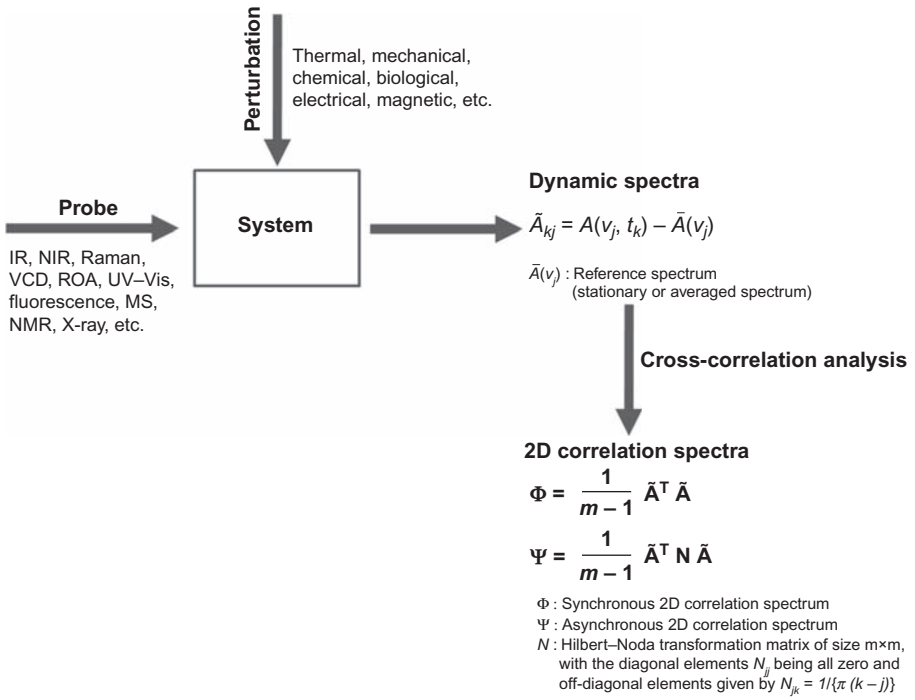


Figure 12.1
A general scheme for constructing generalized two-dimensional (2D) correlation spectra.

$$\overline{A}(\nu_j) = \frac{1}{m}\sum_{i=1}^{m} A(\nu_j, t_i) \tag{12.2}$$

Synchronous correlation intensity $\Phi(\nu_1, \nu_2)$, representing the correlation of simultaneous variations of spectral intensities measured at two different wave numbers, ν_1 and ν_2, becomes

$$\Phi(\nu_1, \nu_2) = \frac{1}{m-1}\sum_{i=1}^{m} \tilde{A}(\nu_1, t_i) \cdot \tilde{A}(\nu_2, t_i) \tag{12.3}$$

It is a quantitative measure of the similarity between spectral intensity variations observed along the perturbation variable t_i. A large value of $\Phi(\nu_1, \nu_2)$ implies the spectral intensities at ν_1 and ν_2 vary in a coordinated or simultaneous manner, inferring the coupled or common origin of the spectral signals.

Asynchronous correlation intensity $\Psi(\nu_1, \nu_2)$, in turn, represents the correlation of nonsimultaneous or sequential variations of spectral intensities measured at ν_1 and ν_2. It is given by

$$\Psi(\nu_1,\nu_2)=\frac{1}{m-1}\sum_{i=1}^{m}\tilde{A}(\nu_1,t_i)\cdot\sum_{k=1}^{m}N_{ik}\tilde{A}(\nu_2,t_k) \tag{12.4}$$

$$N_{ik}=\begin{cases}0 & if\ i=k\\ \dfrac{1}{\pi(k-i)} & otherwise\end{cases} \tag{12.5}$$

This quantity can be used as a convenient measure of dissimilarity or out-of-phase behavior of spectral intensity variations. A large value of $\Psi(\nu_1,\nu_2)$ implies the spectral intensities at ν_1 and ν_2 vary not together but independently of each other. This phenomenon, in turn, suggests that corresponding spectral signals are probably originating from separate molecular moieties which are responding somewhat differently to a common perturbation.

In 2D-COS, it is often useful to represent a spectral dataset in a form of an $m\times n$ matrix $\tilde{A}$ with its element being $\tilde{A}=\tilde{A}(\nu_j,t_i)$. With this notation, the synchronous and asynchronous spectra are conveniently given in the compact form of $n\times n$ matrices

$$\Phi=\frac{1}{m-1}\tilde{\mathbf{A}}^{\mathbf{T}}\tilde{\mathbf{A}} \tag{12.6}$$

$$\Psi=\frac{1}{m-1}\tilde{\mathbf{A}}^{\mathbf{T}}\mathbf{N}\tilde{\mathbf{A}} \tag{12.7}$$

Superscript $^{\mathrm{T}}$ stands for matrix transpose. Eq. (12.5) defines the elements of the so-called Hilbert−Noda transformation matrix $\boldsymbol{N}$.

12.2.3 Interpretation of two-dimensional correlation spectra

A schematic example of synchronous and asynchronous 2D correlation spectra plotted as contour maps is illustrated in Fig. 12.2. The synchronous spectrum represents simultaneous or coincidental changes of spectral intensities at ν_1 and ν_2, while the asynchronous spectrum represents sequential, or unsynchronized, variations [1]. More detailed properties of 2D correlation spectra are now discussed.

The synchronous spectrum is a map of the synchronous correlation intensity defined by the spectral coordinate (ν_1, ν_2). It is a symmetric map with respect to the main diagonal line corresponding to the coordinate at $\nu_1=\nu_2$. The intensity of a peak located at the diagonal position corresponds to the autocorrelation function or variance of spectral intensity variations. These peaks found on the diagonal line are therefore referred to as autopeaks. In the example spectrum shown in Fig. 12.2A, there are four autopeaks located at the different spectral coordinates A−D. The intensities of the autopeaks represent the overall extent of dynamic changes in spectral signals induced by the given perturbation. The greater the change, the stronger the autopeak intensity.

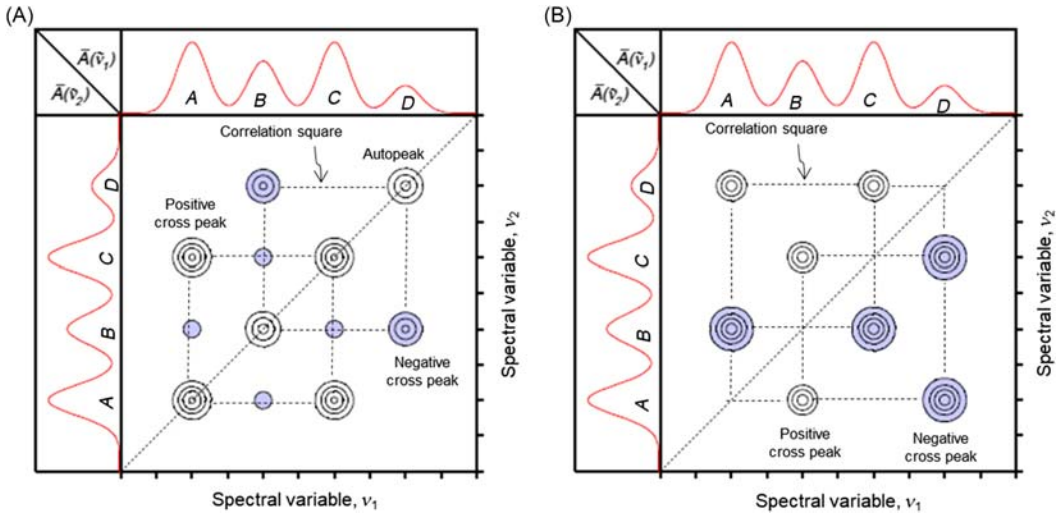


Figure 12.2
(A) Synchronous and (B) asynchronous two-dimensional (2D) correlation spectra. A one-dimensional (1D) reference spectrum is also provided at the top and side of the 2D map.

Cross peaks which are located at the off-diagonal positions of a synchronous spectrum represent the simultaneous changes in signal intensities at two different wave numbers. In the example spectrum, bands A and C are synchronously correlated, so are bands B and D. If the sign of the cross peak is positive, the spectral intensities at the corresponding two spectral variables increase or decrease together. On the other hand, the sign becomes negative (often indicated by the shading of cross peak) if one of the spectral intensities is increasing while the other is decreasing. The example spectrum shown in Fig. 12.2A indicates that the intensities of the bands A and C change in the same directions and the intensities of the bands B and D change in the opposite directions.

The asynchronous counterpart of the 2D correlation spectrum shown in Fig. 12.2B, which consists exclusively of off-diagonal cross peaks, providing information complementary to the synchronous spectrum. It represents sequential or nonsimultaneous changes in spectral intensities measured at ν_1 and ν_2. They are distinguished by different responses to temperature, time, concentration, etc. In other words, asynchronous cross peaks develop only if the intensities of two dynamic spectral intensities vary out of phase with each other. This feature becomes especially useful in differentiating overlapped bands arising from different moieties. The sign of an asynchronous cross peak can also be either negative (as indicated by the shaded areas) or positive. If the signs of synchronous and asynchronous peaks at the same coordinate are the same, then the intensity change at ν_1 occurs predominantly before ν_2 (i.e., in the earlier part of the sequentially arranged spectral dataset). This sequence is reversed if the signs of the two

cross peaks are different. The example spectrum shown in Fig. 12.2 indicates that bands A and C change before bands B and D.

12.2.4 Hetero two-dimensional correlation analysis

The ability to invoke heterocorrelation analysis is one of the most powerful advantages of 2D-COS. In 2D heterocorrelation analysis, three different distinct types are possible: heterospectral correlation [68,89,102,104–107,128], heterosample correlation [116], and heteroperturbation (or hybrid) correlation [152]. Among them, the most active field of research in 2D-COS is 2D heterospectral correlation, where two different types of spectra that are obtained for a system measured by multiple spectroscopic probes under a similar external perturbation are compared.

If there is any commonality between the spectral variations of a system observed under the same perturbation by two different probes, the correlation even between different classes of spectral signals can be detected. 2D heterospectral analysis can be applied to the correlation not only between closely related spectroscopic measurements, such as IR and NIR (or Raman) spectra, but also between completely different types of spectroscopic or physical techniques, such as IR and X-ray. Here we will introduce the basic concept of 2D heterospectral correlation analysis. It compares one dynamic spectrum $\tilde{A}\left(\nu_j, t_i\right)$ measured by one analytical technique with another dynamic spectrum $\tilde{Z}\left(\mu_j, t_i\right)$ measured with a completely different analytical technique. The general form of the heterospectral 2D correlation is given by

$$\mathbf{\Phi} = \frac{1}{m-1}\tilde{\mathbf{A}}^{\mathrm{T}}\tilde{\mathbf{Z}} \tag{12.8}$$

$$\mathbf{\Psi} = \frac{1}{m-1}\tilde{\mathbf{A}}^{\mathrm{T}}\mathbf{N}\tilde{\mathbf{Z}} \tag{12.9}$$

in which $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{Z}}$ are data matrices of *dynamic spectra* $\tilde{A}(\nu_j, t_i)$ and $\tilde{Z}(\mu_j, t_i)$ respectively.

Heterosample correlation, where correlations are examined among spectral intensity variations of different samples by using the same spectroscopic probe and perturbation condition, is a very interesting form of heterocorrelation analysis. The heterosample 2D correlation analysis in protein studies will be discussed later. 2D heterosample correlation analysis can open a much wider range of correlation studies than the conventional 2D correlation spectroscopy. It can also give additional unexpected insights into the studied systems.

Hybrid 2D-COS for two spectral datasets obtained separately was originally proposed by Wu et al. [152]. In hybrid 2D-COS, spectral intensity variations induced by applying different types of external perturbations to the same sample are correlated. It can potentially

explore the latent correlation between different perturbation variables. It is called heteroperturbation correlation, which examines the correlation of the effect of applied external perturbations to the system.

12.3 Applications of two-dimensional correlation spectroscopy to protein study

Vibrational spectroscopy is a powerful analytical technique well suited for protein study. However, the extensively overlapped spectral features occasionally limit the utility of IR spectroscopy for protein studies investigating the structural changes. To overcome this shortcoming of conventional spectral analysis, 2D-COS has often been applied to IR spectra. Analysis of vibrational spectra using 2D-COS can give more information to investigate protein structural changes, the protein denaturation mechanism, and the conformation as well as configuration of protein. Therefore 2D-COS provides new insights into better understanding protein denaturation mechanisms at the molecular level.

It is not surprising to find that within the application of 2D-COS to biological and biomedical research, the most popular sample studied is protein. 2D-COS has many advantages, such as the enhanced spectral resolution, identification of inter- or intramolecular interactions, and the determination of the sequential order of band intensity changes. These advantages provide the improved understanding to characterize subtle structural changes of protein, which are sometimes hardly detected in conventional 1D spectroscopy [5,7,10,12–36]. 2D-COS can open a way to explore spectral changes of protein obtained under a perturbation. Popularly used perturbations in protein studies are temperature [27,37–46,59,71,79,81,108,113,114,119,124–134], concentration of protein [54–56,71,83,115,124,135–140], pH [44,47–49,56,82,85–88,98,102,120–122,147,150], chemical denaturants [50–53,148,149,151], hydrogen/deuterium exchange [58,59,62–65,67,73], and pressure [34,57–61]. Of these, temperature is most commonly used in protein studies. Spectral changes are usually accompanied by structural changes in the protein backbone induced upon cooling or heating. Until now, the most popularly used analytical probe in 2D-COS for protein study is IR spectroscopy [29,31,32,34,38–40, 42,44,46,48,51–84,104–106,109,112–117,124,125,127,129–134,136,138–140,143–150]. In protein studies, Raman [30,49,86–97,104–106,120–123,126,141,142], VCD [98,99,106], and ROA spectroscopy [86,88–92,128] have also been widely applied in 2D-COS.

Furthermore, the incorporation of 2D-COS and multivariate chemometric techniques, such as PCA, partial least squares regression, and multivariate curve resolution, is also very useful in protein research. 2D-COS and PCA are generally used analytical methods in spectroscopy. Both provide complementarity information but have different computational

methods and objectives. Each spectrum in the entire range is considered in PCA to be one integrated object represented by scores and loading vectors. In contrast, the spectral variations at each frequency that contribute individually to the synchronous and asynchronous values in 2D-COS are not based on an integral approach. The chemometrics combined 2D-COS provides great potential to investigate protein structure during the denaturation process [72,113,114,116,119–122,136,137].

Many reviews describing the application of vibrational 2D-COS in protein studies with detailed descriptions of the computational procedures and of the background were previously published [7,10,11,26,29–31]. In this chapter, we will briefly compile the applications of 2D-COS in protein studies describing the improved information about protein structure obtained from vibrational spectra.

12.3.1 Two-dimensional infrared correlation spectroscopy in protein study

IR spectroscopy is very well suited to protein studies, such as protein aggregation and amyloid fibril formation during protein denaturation process. Most protein study by IR spectroscopy is mainly focused on the amide I region, where it is well-established for the study of protein secondary structure.

Park et al. investigated the thermal denaturation mechanism and structural difference between spherical and rod-shaped human insulin nanoparticles produced by a process of supercritical fluids precipitation [119]. They produced spherical and rod-shaped insulin nanoparticles using the supercritical fluids technique with dimethyl sulfoxide and ethanol solutions as the solvents, respectively. Fig. 12.3A and B show temperature-dependent IR spectra for the spherical and rod-shaped insulin nanoparticles, respectively. The changes in the IR spectra of two different types of nanoparticles with increasing temperature were quite similar in the amide I (1800–1580 cm^{-1}) regions.

To better understand the secondary structure changes during thermal denaturation, they applied 2D-COS to the temperature-dependent IR spectra and only focused on the amide I region of the two shapes of insulin nanoparticles. 2D correlation spectra of spherical and rod-shaped insulin nanoparticles are displayed in Figs. 12.4 and 12.5, respectively. As shown in the synchronous 2D IR correlation spectra (Figs. 12.4A and 12.5A), three bands at 1716, 1670, and 1603 cm^{-1} were observed in the spherical insulin nanoparticles, whereas the bands at 1664 and 1630 cm^{-1} were the most changed in the rod-shaped nanoparticles. These synchronous 2D IR correlation spectra clearly revealed that variations in the β structure and the helical structures with increasing temperature predominated in the spherical and rod-shaped nanoparticles, respectively. Based on 2D-COS analysis, the sequential order of intensity changes with heating are determined as follows: disordered structure $\rightarrow$ α-helix $\rightarrow$ β-turn $\rightarrow$ intermolecular β-sheet $\rightarrow$ β-sheet in spherical type of

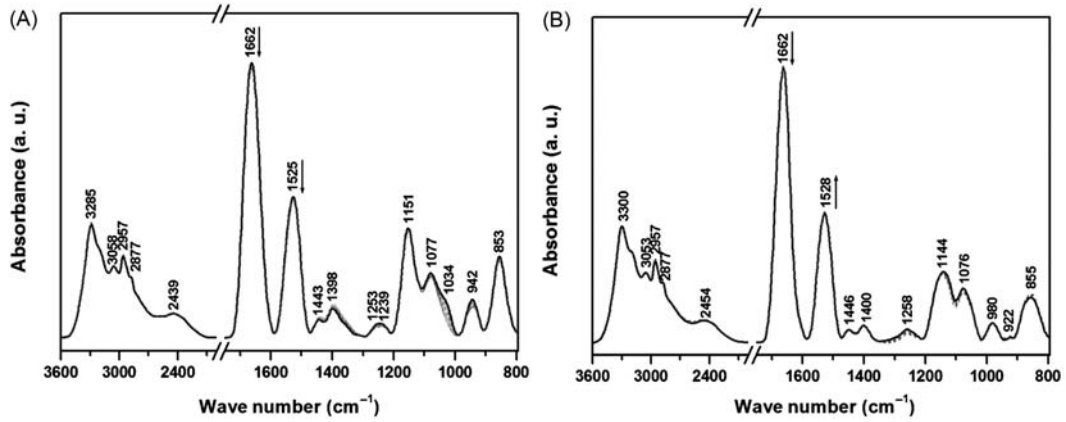


Figure 12.3

Temperature-dependent infrared (IR) spectra in the 800–3600 cm^{-1} region for the (A) spherical and (B) rod-shaped insulin nanoparticles. The arrows indicate the direction of changes in the band intensity with increasing temperature. Source*: Reproduced with permission from Y. Park, Y. Seo, B. Chae, D. Pyo, H. Chung, H. Hwang, et al., Understanding the structural difference between spherical and rod-shaped human insulin nanoparticles produced by supercritical fluids precipitation, ChemPhysChem 16 (2015) 476–482, Copyright 2015, with permission from Wiley.*

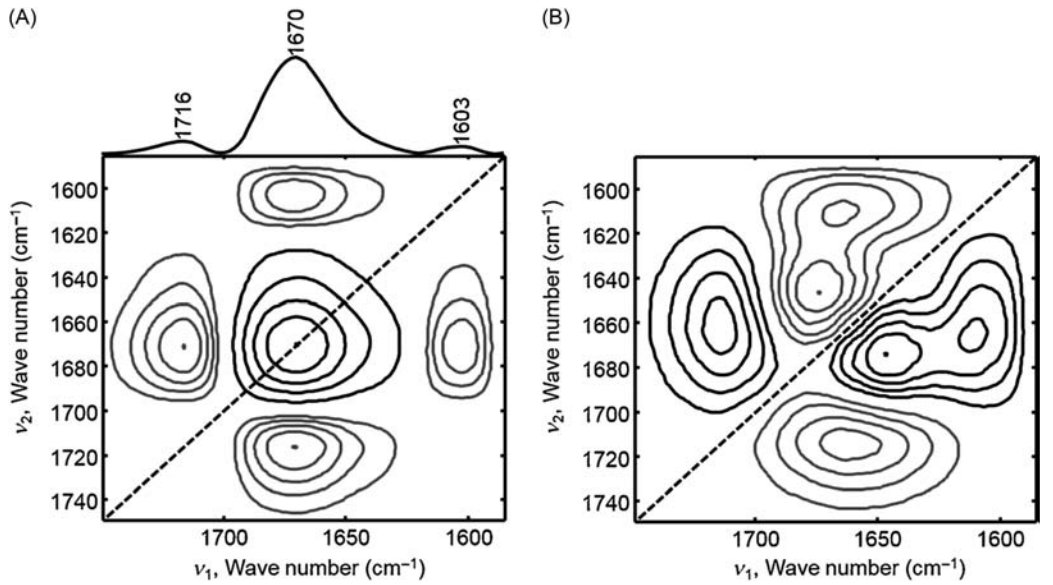


Figure 12.4

(A) Synchronous and (B) asynchronous two-dimensional infrared (2D IR) correlation spectra obtained from the temperature-dependent changes in spectra of the spherical-shaped insulin nanoparticles. The black and gray lines represent positive and negative cross peaks, respectively. Source*: Reproduced with permission from Y. Park, Y. Seo, B. Chae, D. Pyo, H. Chung, H. Hwang, et al., Understanding the structural difference between spherical and rod-shaped human insulin nanoparticles produced by supercritical fluids precipitation, ChemPhysChem 16 (2015) 476–482, Copyright 2015, with permission from Wiley.*

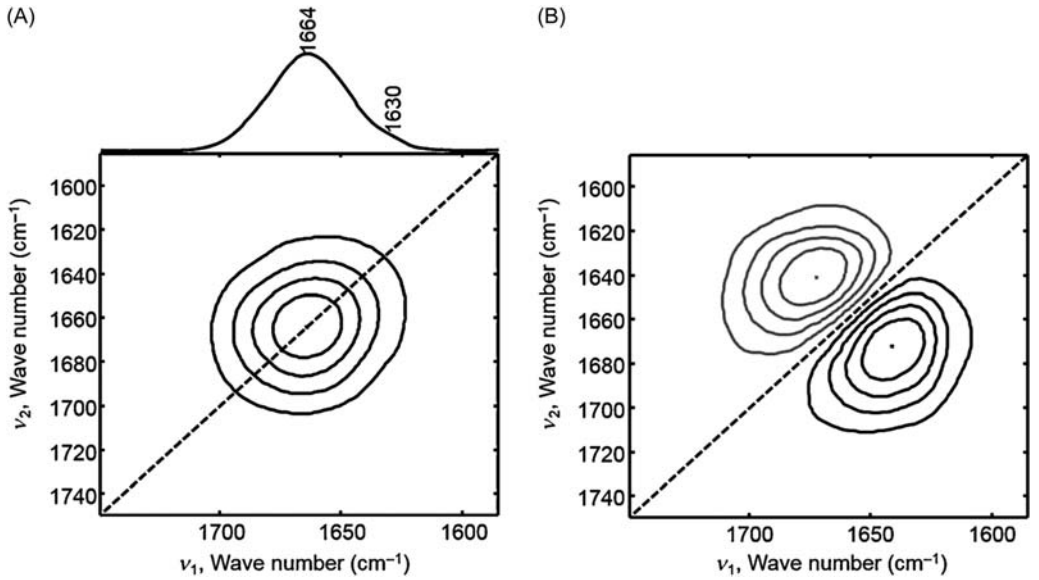


Figure 12.5

(A) Synchronous and (B) asynchronous 2D IR correlation spectra obtained from the temperature-dependent changes in spectra of the rod-shaped insulin nanoparticles. The black and gray lines represent positive and negative cross peaks, respectively. Source: *Reproduced with permission from Y. Park, Y. Seo, B. Chae, D. Pyo, H. Chung, H. Hwang, et al., Understanding the structural difference between spherical and rod-shaped human insulin nanoparticles produced by supercritical fluids precipitation, ChemPhysChem 16 (2015) 476–482, Copyright* 2015, *with permission from Wiley.*

nanoparticles, whereas β-sheet → disordered structure → α-helix → β-turn → β-sheet in the rod-shaped particles. From the 2D correlation analysis, it can be concluded that an intermolecular β-sheet structure exists only in the spherical insulin nanoparticles and that the thermal behavior of the two shapes of nanoparticles are completely different. They clearly revealed by using 2D IR correlation spectroscopy that the thermal denaturation mechanisms of the two types of nanoparticles are completely different and that an intermolecular β-sheet structure exists only in the spherical type insulin nanoparticles.

Litwińczuk et al. investigated the pH-induced transition from the native (N) to the acidic (A) state of bovine α-lactalbumin (bLA) by using 2D IR correlation spectroscopy [117]. They analyzed the pH-induced structural changes of IR spectra obtained in the transmission (TR) and attenuated total reflection (ATR) modes. As shown in Fig. 12.6, the pH-dependent IR spectra of the film and the solution samples of bLA measured in the TA and ATR modes are remarkably similar to each other. In 2D-COS, all Fourier transform infrared (FTIR) spectra treated by Pareto scaling were used to enhance the contribution of the weak spectral features to 2D correlation spectra. The detection of the important spectral changes related to the protein secondary structure in 2D correlation spectra can be improved,

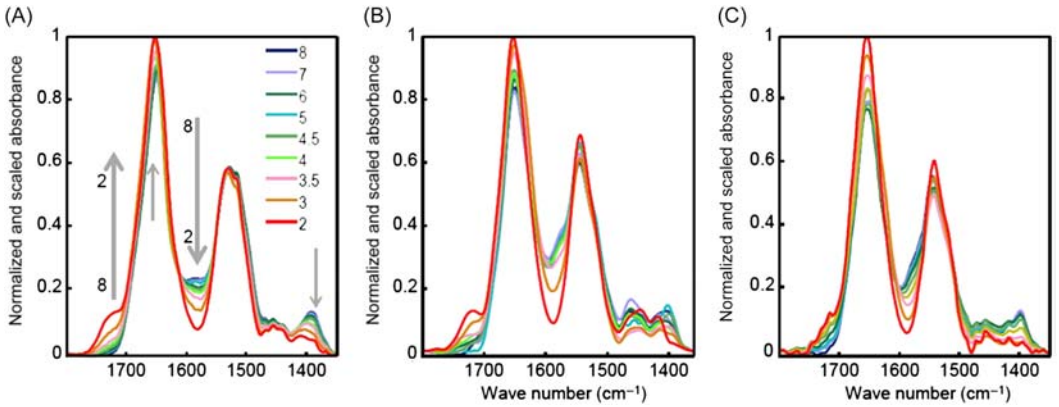


Figure 12.6

The normalized spectra for the holoform of bovine α-lactalbumin (bLA) obtained in the following experiments: (A) Fourier transform infrared (FTIR) in the attenuated total reflection (ATR) mode for dry film samples; (B) FTIR in the ATR mode for solutions; and (C) FTIR in the transmission (TR) for solutions. Source*: Reproduced with permission from A. Litwińczuk, S.R. Ryu, L.A. Nafie, J.W. Lee, H.I. Kim, Y.M. Jung, et al., The transition from the native to the acid-state characterized by multi-spectroscopy approach: Study for the holo-form of bovine α-lactalbumin, Biochim. Biophys. Acta 1844 (2014) 593–606, Copyright 2014, with permission from Elsevier.*

as Pareto scaling emphasizes the smaller absorbance features [153]. Fig. 12.7 shows synchronous 2D correlation spectra of three systems, which reveal that the pH-induced spectral changes are apparently similar. However, the power spectra extracted along the diagonal line of the synchronous 2D correlation spectrum shown in Fig. 12.8 revealed clear differences of three samples. Synchronous 2D correlation spectra in Fig. 12.7 clearly showed the most affected structural changes by the decreasing pH conditions are native α-helices and β-sheets. In synchronous 2D correlation spectra, only positive peaks in the regions 1700–1600 cm^{-1} were observed.

As the film has smaller interactions with water than solution samples, the power spectrum of film shows smaller secondary structure at subsequent stages of the transition from the native to the A-state than those obtained for the solutions. The observed very weak absorbance changes attributed to the β-turn formation and the high-frequency shift of the peak assigned to the β-sheet structures clearly reveal that the differences in hydration affected both the interior (β-sheet changes) and exterior (β-turn changes) of bLA. Furthermore, the power spectra of two solutions should be very similar due to their same properties. However, the power spectra of two solutions were clearly different. These discrepancies of the power spectra in position and in intensity are likely due to the differences between the interaction of the bLA molecules with the surface of the Ge crystal in the ATR experiment and with the CaF_2 windows in the TR experiment.

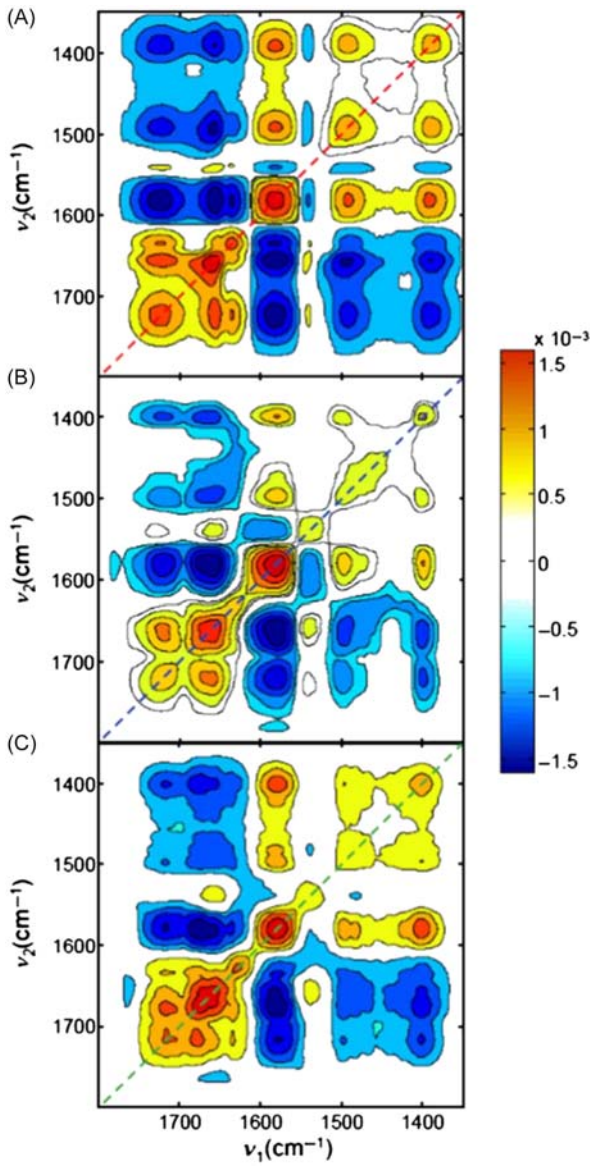


Figure 12.7

The synchronous spectra calculated from the infrared (IR) spectra measured under different conditions: (A) film sample in attenuated total reflection (ATR) mode; (B) solution sample in ATR mode; and (C) solution sample in transmission (TR) mode. The color bar, which is in common for the three maps, is displayed to the right of the plot. Source*: Reproduced with permission from A. Litwińczuk, S.R. Ryu, L.A. Nafie, J.W. Lee, H.I. Kim, Y.M. Jung, et al., The transition from the native to the acid-state characterized by multi-spectroscopy approach: Study for the holo-form of bovine α-lactalbumin, Biochim. Biophys. Acta 1844 (2014) 593–606, Copyright 2014, with permission from Elsevier.*

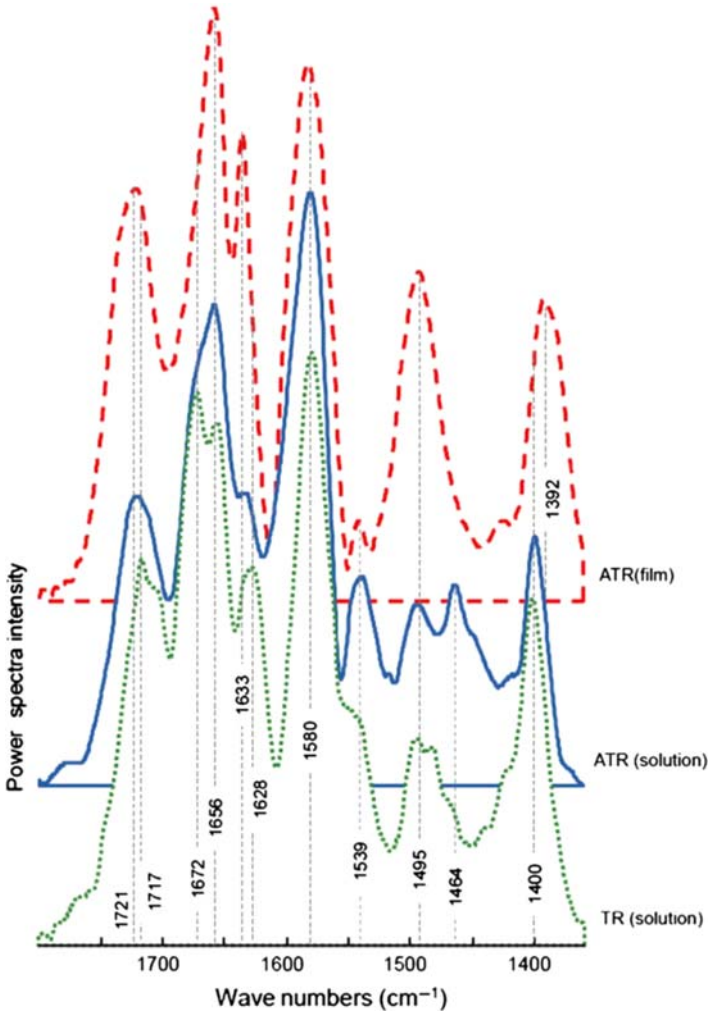


Figure 12.8

Stacking of power spectra extracted from the three synchronous maps: (A) film sample in attenuated total reflection (ATR) mode (*red dashed line*); (B) solution sample in ATR mode (*blue continuous line*); and (C) solution sample in transmission (TR) mode (*green dotted line*). Source: *Reproduced with permission from A. Litwińczuk, S.R. Ryu, L.A. Nafie, J.W. Lee, H.I. Kim, Y.M. Jung, et al., The transition from the native to the acid-state characterized by multi-spectroscopy approach: Study for the holo-form of bovine α-lactalbumin, Biochim. Biophys. Acta 1844 (2014) 593–606, Copyright 2014, with permission from Elsevier.*

The corresponding asynchronous 2D spectra shown in Fig. 12.9 are very different, because the asynchronous 2D spectrum is very sensitive to subtle differences in the experimental conditions. From the analysis of 2D correlation spectra the sequential order of intensity changes during transition from N to A-state were determined: for the film sample,

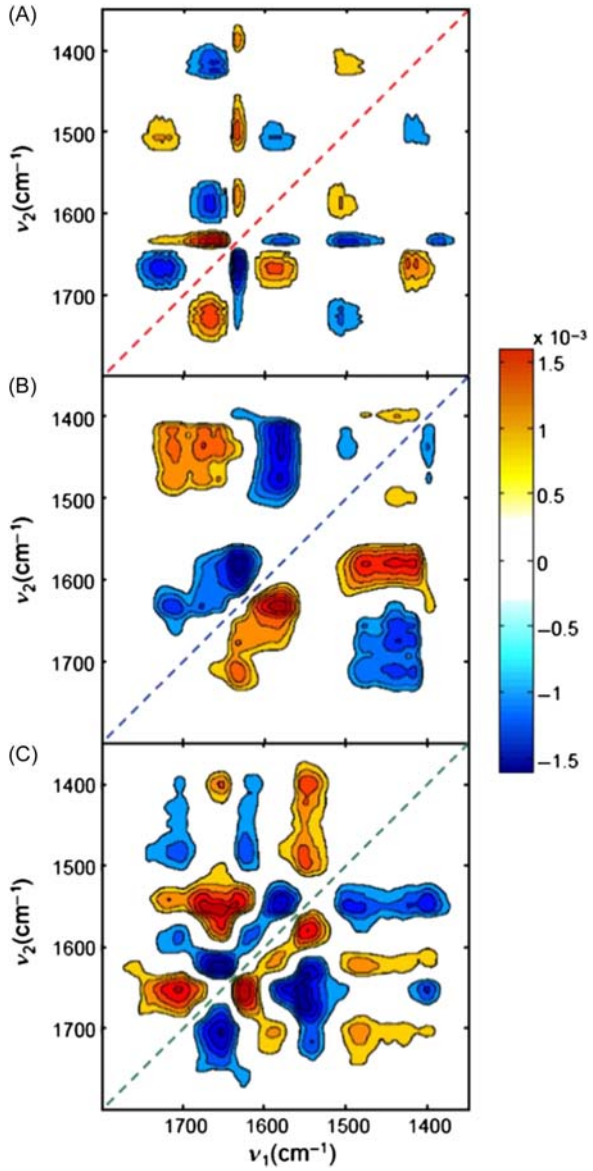


Figure 12.9

The asynchronous spectra calculated from the infrared spectra measured under different conditions: (A) film sample in attenuated total reflection (ATR) mode; (B) solution sample in ATR mode; and (C) solution sample in transmission (TR) mode. The color bar, which is in common for the three maps, is displayed to the right of the plot. Solid and dashed lines represent positive and negative cross peaks, respectively. Source*: Reproduced with permission from A. Litwińczuk, S.R. Ryu, L.A. Nafie, J.W. Lee, H.I. Kim, Y.M. Jung, et al., The transition from the native to the acid-state characterized by multi-spectroscopy approach: Study for the holo-form of bovine α-lactalbumin, Biochim. Biophys. Acta 1844 (2014) 593–606, Copyright 2014, with permission from Elsevier.*

the β-turn elements → the aromatic residues → the dehydrated β-sheet; for the solution sample in the ATR experiment, the aromatic residues → rearrangements in the β-sheet structure → β-turns emerged simultaneously with the changes assigned to the protonation of the aspartic groups; for the solution sample in the TR experiment, the helical fragments → the β-domain → the α-domain → the main protonation process of the aspartic and glutamic groups. It means that subtle changes caused by the transition from N to A-state accompanied by comparable intensity changes induced by hydration and adsorption processes could strongly disturb the overall picture of the formation of the A-state. That is, the A-state intermediate structure determined based on variations of the secondary structure can be varied by the experimental conditions, such as water deficiencies, interactions with the hydrophobic or hydrophilic surfaces of the ATR crystal, and adsorption processes.
It clearly reveals that there were no seemingly competing properties of the secondary elements, such as stability and flexibility during transition of bLA from N to A-state.
It clearly reveals that properties of the secondary structure were very stable during pH-induced transition of bLA from N to A-state; on the other hand, they were highly sensitive to hydration conditions and the interaction with the hydrophobic or hydrophilic surface of the ATR crystal.

12.3.2 Two-dimensional Raman correlation spectroscopy in protein study

The IR study of protein secondary structures is mostly based on only the amide I region. In contrast, the Raman study is focused both on amide I and amide III regions. Raman spectra provide complementary information for investigating the secondary structure of proteins to IR spectra.

Park et al. first investigated the pH-induced formation mechanism of the partially unfolded bovine α-lactalbumin (bLA) and oleic acid (OA) complex by using Raman spectroscopy [120]. This α-lactalbumin (ALA)/OA complex, called bLA made lethal to tumor cells (BAMLET), is a tumoricidal complex. For a better understanding of the formation mechanism of the ALA/OA complex with decreasing pH, 2D-COS was applied to pH-dependent Raman spectra of the ALA/OA complex. Based on the results of the 2D-COS analysis, they successfully elucidated the pH-induced formation mechanism of the ALA/OA complex at the molecular level.

Raman spectra of the ALA/OA complex obtained at pH values from 4.5 to 1.0 are displayed in Fig. 12.10. Most of bands for the amide I, II, III, IV, and V regions in the 650−1800 cm^{-1} range were sensitive to pH changes. Some bands shown in the ALA solution of their previous study [122], such as the bands near 830 and 850 cm^{-1} assigned to tyrosine (Tyr) Fermi doublet bands, were not clearly observed in this system, which may be because changes in these bands are masked by OA. As shown in Fig. 12.10, there were eight significant bands at 770 [threonine (Thr)], 892 [lysine (Lys)], 1218 [Tyr/phenylalanine (Phe)], 1269 (amide III),

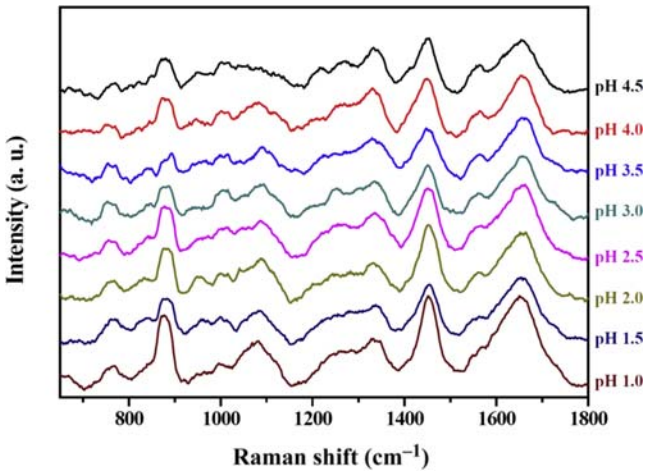


Figure 12.10
pH-dependent Raman spectra of the α-lactalbumin/oleic acid (ALA/OA) complex obtained at pH values from 4.5 to 1.0. Source: *Reproduced with permission from Y. Park, Y. Park, S. Jin, J.W. Kim, Y.M. Jung, Formation mechanism of BAMLET by 2D Raman correlation analysis, J. Mol. Struct. 1171 (2018) 33–37, Copyright 2018, with permission from Elsevier.*

1334 [α-helix/tryptophan (Trp)], 1450 (C–H deformation), 1561 (the indole ring of Trp), and 1654 (amide I) cm^{-1} for pH 4.5 [154–156]. The broad feature at 860–890 cm^{-1} is assigned to Trp hydrogen bonding [157]. The changes of band position and intensity with decreasing pH indicate that the partially buried Trp amino acid residue of ALA may be exposed after the formation of the ALA/OA complex. The intensities of bands at 1218, 1269, and 1561 cm^{-1} decreased with pH decreasing, but those of the bands at 1450 and 1654 cm^{-1} increased. They investigated the details of these changes by using 2D-COS.

Fig. 12.11 shows the 2D Raman correlation spectra of the ALA/OA complex obtained from the pH-dependent Raman spectra at 1520–1800 cm^{-1} for amide I and II regions. As shown in the synchronous 2D correlation spectrum (Fig. 12.11A), the changes of two bands at 1708 and 1728 cm^{-1} are very sensitive to pH, and are hardly observed in the 1D Raman spectra. Based on the analysis of 2D correlation spectra, the following sequence of intensity changes during the formation of the ALA/OA complex was determined: 1591 (Tyr/Trp/Phe) → 1672 (β-turn) → 1564 (Trp (side chain/indole ring)) → 1551 (Trp (amide II/indole ring)) → 1725 (ester C = O stretching (COOH) in amino acid residues) → 1706 (deprotonated ester C = O stretching (COO^-) in amino acid residues) → 1609 (Tyr/Trp/Phe) → 1638 (β-sheet) cm^{-1}. Interestingly, the intensity changes of the α-helix structure during the formation of the ALA/OA complex was not detected, suggesting that the spectral variation of the α-helix structure is less sensitive in Raman spectroscopy than in IR spectroscopy, as in their previous study [121]. They performed a heteroregion 2D correlation analysis to confirm the band assignments of the 1700–1780 cm^{-1} region.

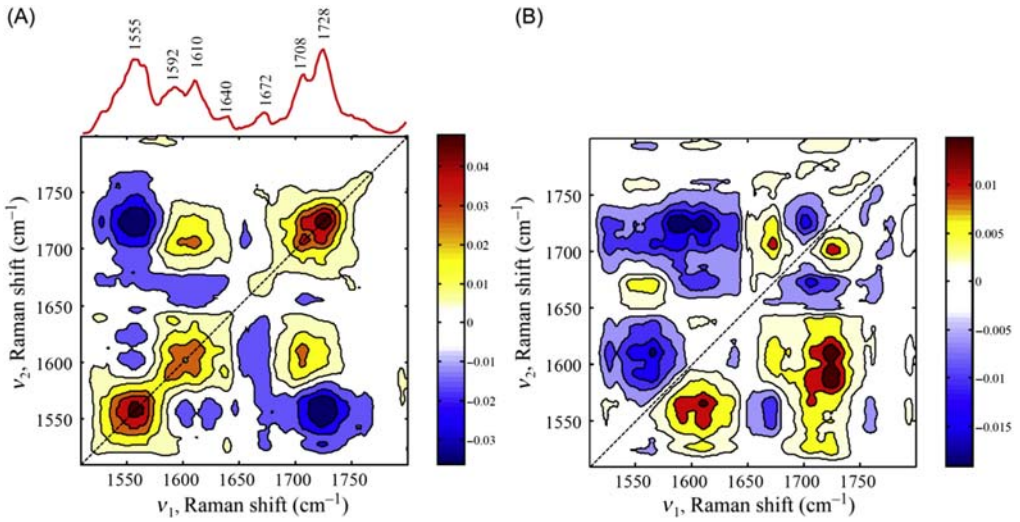


Figure 12.11

Synchronous (A) and asynchronous (B) two-dimensional (2D) Raman correlation spectra in the 1520–1800 cm^{-1} region of the α-lactalbumin/oleic acid (ALA/OA) complex with decreasing pH. The color bar, which also applies to the map, is displayed to the right of each plot. Source: *Reproduced with permission from Y. Park, Y. Park, S. Jin, J.W. Kim, Y.M. Jung, Formation mechanism of BAMLET by 2D Raman correlation analysis, J. Mol. Struct. 1171 (2018) 33–37, Copyright 2018, with permission from Elsevier.*

The heteroregion synchronous 2D Raman correlation spectra of the ALA/OA complex with decreasing pH are displayed in Fig. 12.12. Two bands at 1706 and 1724 cm^{-1} assigned to protonated/deprotonated carboxylic acid have positive cross peaks with the bands at 783, 1108, 1224, 1235, 1250, and 1466 cm^{-1} for the amino acid (Lys, His, Glu, and Thr) residues and the β-structure. It reveals that these two bands at 1706 and 1724 cm^{-1} for carboxylic acid are due to the formation of the ALA/OA complex because these four amino acids are the interaction site with OA [158]. Furthermore, it also confirms that the ALA/OA complex formed at pH 3.0 or below, because the PCA results showed all the bands at 783, 1108, 1224, 1235, 1250, 1466, 1706, and 1724 cm^{-1} changed below pH 3.0. The ambiguous band assignments of the Raman spectra of the ALA/OA complex with decreasing pH were clearly established by heteroregion synchronous 2D Raman correlation spectra.

12.3.3 Two-dimensional vibrational circular dichroism and two-dimensional Raman optical activity correlation spectroscopy in protein study

Unlike conventional IR and Raman spectra, VCD and ROA are sensitive to the stereochemical details of molecular structure. Therefore VCD and ROA can provide more

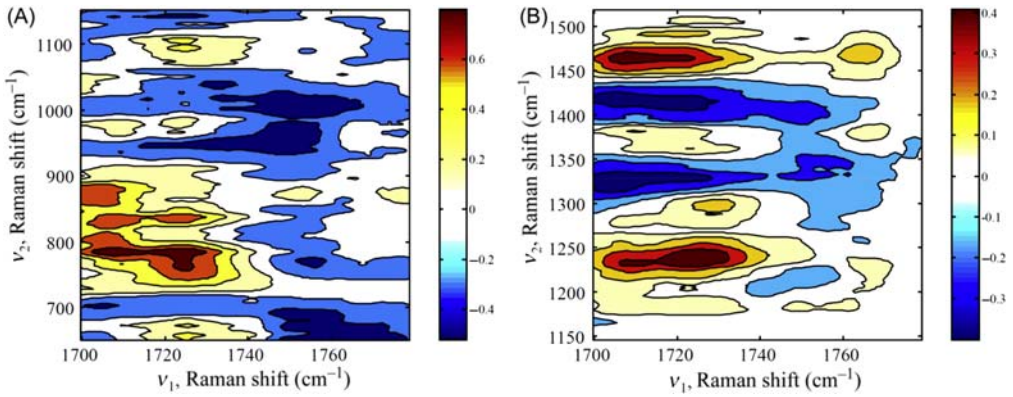


Figure 12.12

Heteroregion synchronous two-dimensional (2D) Raman correlation spectra of the α-lactalbumin/oleic acid (ALA/OA) complex with decreasing pH for the 650−1150 and 1700−1800 cm^{-1} regions (A) and the 1150−1520 and 1700−1800 cm^{-1} regions (B). The color bar, which also applies to the map, is displayed to the right of each plot. Source: *Reproduced with permission from Y. Park, Y. Park, S. Jin, J.W. Kim, Y.M. Jung, Formation mechanism of BAMLET by 2D Raman correlation analysis, J. Mol. Struct. 1171 (2018) 33−37, Copyright 2018, with permission from Elsevier.*

information of the protein structure at the molecular level and detailed structural dynamics in proteins. The application of 2D-COS to VCD and ROA spectra is an attractive and valuable method for the investigation of conformational transitions in proteins. Recently, 2D VCD and 2D ROA correlation spectroscopies have been performed to obtain further details about the changes of biomolecules, such as amino acids and proteins [86,88−92,98,99,106,128].

The nature of ROA and VCD spectra, that is, having both positive and negative bands, should be taken into account to interpret 2D correlation spectra for determination of the sequential order of intensity changes. It is noted that for the interpretation of synchronous cross peaks occurring between positive and negative bands, negative cross peaks indicate that both bands are increasing or decreasing in intensity together, while positive cross peaks indicate that one band is increasing and the other is decreasing in intensity. However, the sequential order of intensity changes can be determined by the same rule of 2D-COS regardless of the sign of the band.

Ryu et al. demonstrated pH-induced VCD spectral changes of ALA aqueous solutions by using 2D correlation spectroscopy [99]. Fig. 12.13 displays the gradual change in the pH-dependent VCD spectra of ALA from pH 8 to pH 1.5 in which the amide I and amide II bands changed the most as pH decreased. In the VCD spectrum of ALA at pH 8, a small positive band at 1685 cm^{-1} was observed that indicates the presence of a β-turn structure. Strong amide I bands at 1659 cm^{-1} clearly suggest a predominant α-helix structure. An intense negative band at 1520 cm^{-1} was observed in the amide II region in VCD spectra.

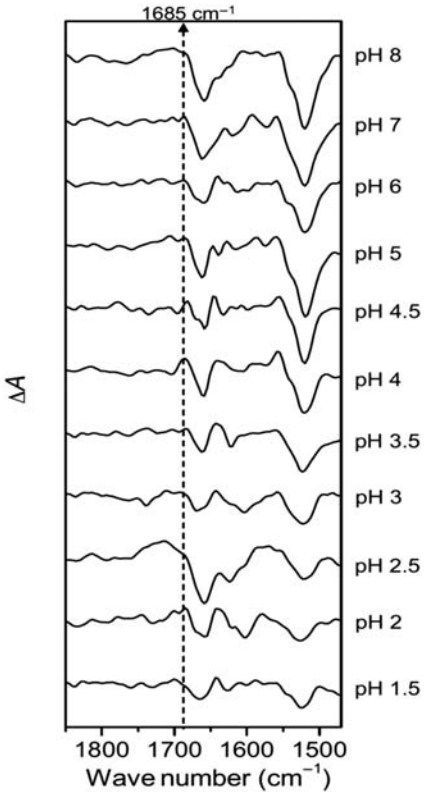


Figure 12.13

pH-dependent vibrational circular dichroism (VCD) spectra of α-lactalbumin (ALA). Source: *Reproduced with permission from S.R. Ryu, B. Czarnik-Matusewicz, R.K. Dukor, L.A. Nafie, Y.M. Jung, Analysis of the molten globule state of bovine α-lactalbumin by using vibrational circular dichroism, Vib. Spectrosc. 60 (2012) 68–72, Copyright 2012, with permission from Elsevier.*

2D VCD correlation spectra for the higher pH region (pH 3–8) are shown in Fig. 12.14. The power spectrum extracted from along the diagonal line on the synchronous 2D correlation spectrum is also given at the top of Fig. 12.14A. The strongest autopeak at around 1520 cm^{-1} is observed in the synchronous 2D correlation spectrum. This band is apparently resolved into two bands at 1543 and 1516 cm^{-1}, which are assigned to the amide II and Tyr bands, respectively, in the asynchronous 2D correlation spectrum. Based on the results of 2D correlation spectroscopy analysis, the following sequence of spectral intensity changes was determined: unordered structure → β-turn → Tyr residue. In 2D VCD correlation spectra for the lower pH region (pH 1.5–3) shown in Fig. 12.15, noticeable spectral changes of amide I band were observed. Autopeaks at 1659 and 1636 (shoulder) cm^{-1} assigned to α-helix and β-sheet, respectively, were also observed in the synchronous 2D correlation spectrum. Analysis of 2D correlation spectra suggested the following sequential order of intensity changes from the premolten globule state to the molten globule state: intermolecular β-sheet → lower-frequency β-sheet structures → α-helices.

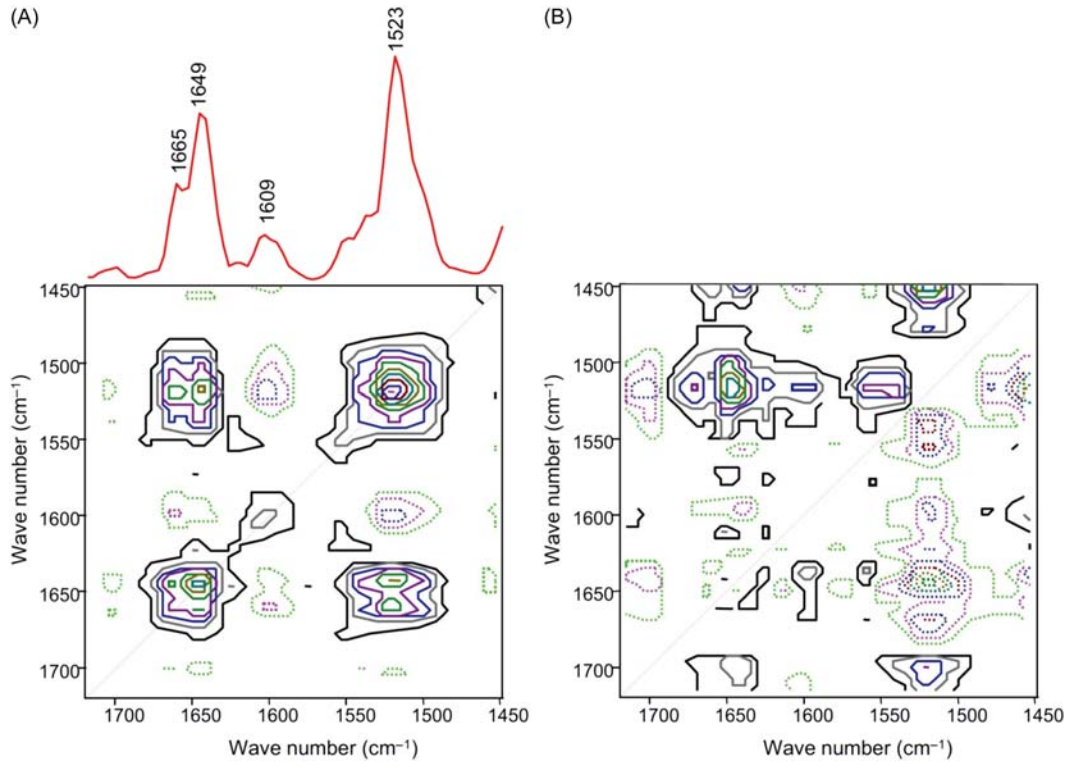


Figure 12.14

The synchronous (A) and asynchronous (B) two-dimensional vibrational circular dichroism (2D VCD) correlation spectra obtained from the pH-dependent VCD spectra of α-lactalbumin (ALA) for pH 8−3. The solid and dashed lines represent positive and negative cross peaks, respectively. Source: *Reproduced with permission from S.R. Ryu, B. Czarnik-Matusewicz, R.K. Dukor, L.A. Nafie, Y.M. Jung, Analysis of the molten globule state of bovine α-lactalbumin by using vibrational circular dichroism, Vib. Spectrosc. 60 (2012) 68−72, Copyright 2012, with permission from Elsevier.*

Ashton et al. [89] applied for the first time 2D-COS to ROA spectra and combined it with 2D Raman correlation analysis to monitor the transition from α-helix to β-sheet in poly(L-lysine), to better understand the underlying mechanism of fibril formation. It was difficult to obtain reliable structural information for the temperature-induced spectral intensity changes of poly(L-lysine) from the raw Raman and ROA spectra, because potentially useful information is obscured by large streaking and elongation of cross peaks. To obtain clear and reliable 2D ROA and 2D Raman correlation spectra and to substantially enhance their interpretation, a simple process of four pretreatments, that is, baseline subtraction, smoothing, normalization, and further baseline subtraction, was performed before 2D-COS [90]. Analysis of 2D ROA correlation spectra can identify numerous bands and determine the sequential order of intensity changes. In addition, they demonstrated the

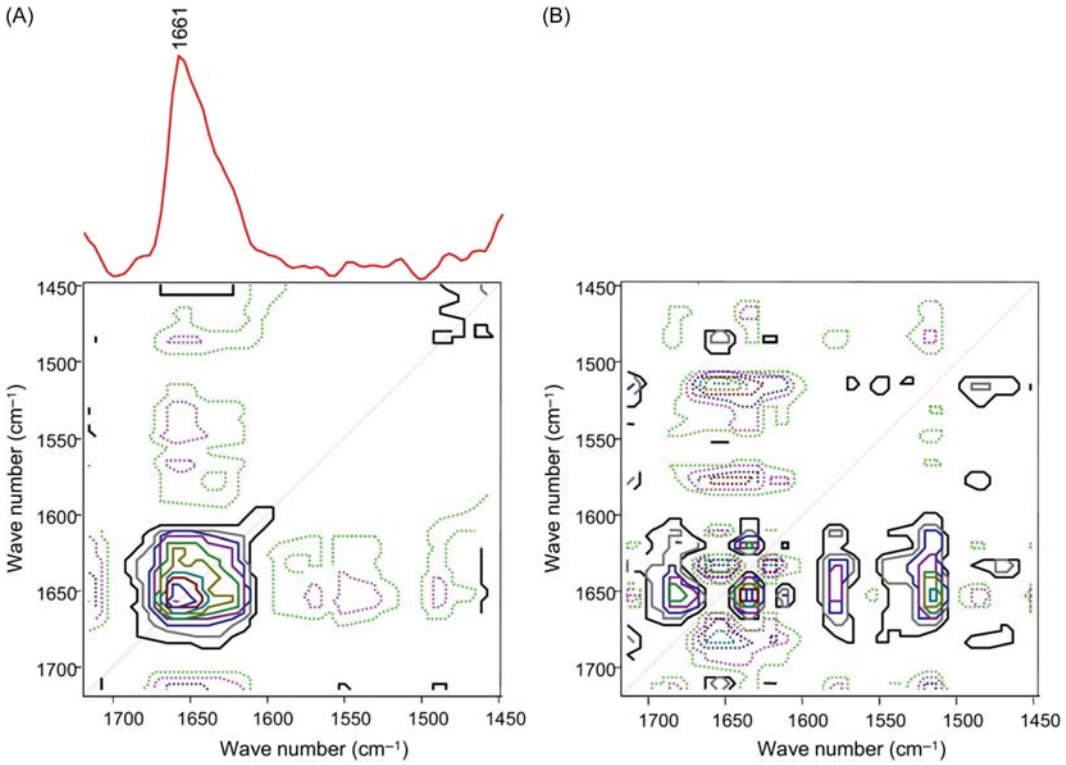


Figure 12.15
The synchronous (A) and asynchronous (B) two-dimensional vibrational circular dichroism (2D VCD) correlation spectra obtained from the pH-dependent VCD spectra of α-lactalbumin (ALA) for pH 3−1.5. The solid and dashed lines represent positive and negative cross peaks, respectively.
Source*: Reproduced with permission from S.R. Ryu, B. Czarnik-Matusewicz, R.K. Dukor, L.A. Nafie, Y.M. Jung, Analysis of the molten globule state of bovine α-lactalbumin by using vibrational circular dichroism, Vib. Spectrosc. 60 (2012) 68−72, Copyright 2012, with permission from Elsevier.*

vast potential of 2D ROA correlation analysis to probe not only changes in conformational transitions but also to identify aggregation prior to the conformational changes.

12.3.4 Combination of two-dimensional correlation spectroscopy and principal component analysis in protein study

A very powerful development in the 2D-COS for protein research is the incorporation of 2D-COS and multivariate chemometric techniques, especially PCA. This combination of 2D-COS and PCA has a great potential in protein study, which provides more reliable and significant interpretation for protein investigations.

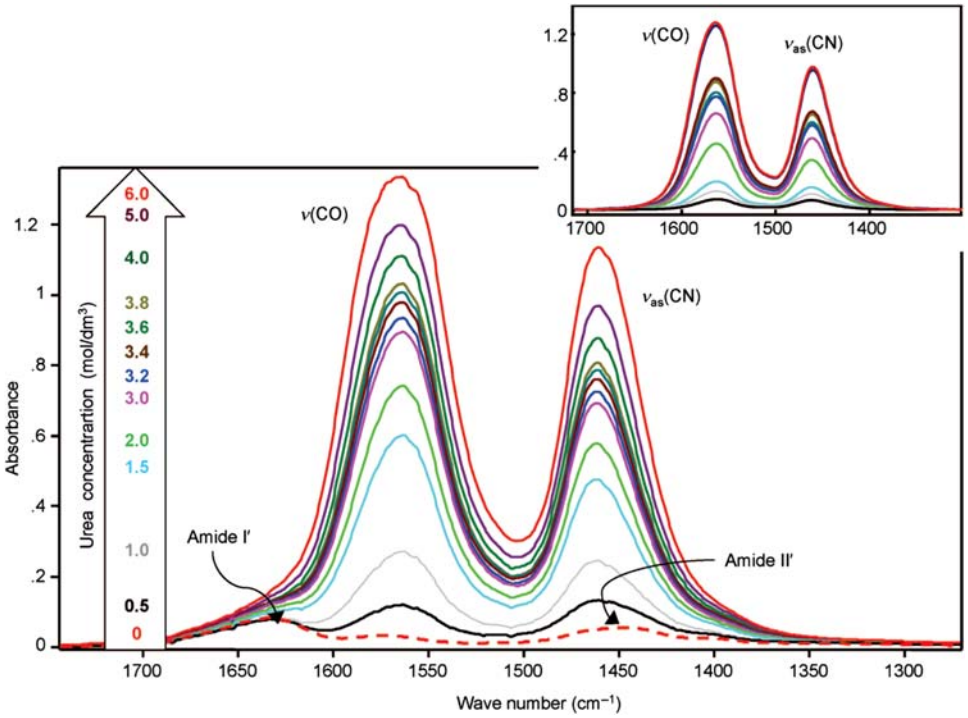


Figure 12.16

Absorbance evolution as a function of urea concentration in the range from 1700 to 1300 cm^{-1} for the ternary system after subtracting the spectrum of buffer, baseline correction, and denoising procedure. The spectrum marked by the dashed line relates to the solution of β-lactoglobulin (BLG) (2 wt.%) without urea. The inset shows the same category of results for the binary system. Source: *Reproduced with permission from B. Czarnik-Matusewicz, S.B. Kim, Y.M. Jung, A study of urea-dependent denaturation of β-lactoglobulin by principal component analysis and two-dimensional correlation spectroscopy, J. Phys. Chem. B 113 (2009) 559–566, Copyright 2009, American Chemical Society.*

Czarnik-Matusewicz et al. examined the role of urea in the unfolding of β-lactoglobulin (BLG) induced by increasing concentration of urea by using both 2D-COS and PCA [116]. They analyzed two bands of urea, ν(C = O) and ν_{as}(CN) bands, instead of the typical analysis of amide bands of protein. Fig. 12.16 shows the urea concentration-dependent IR spectra of BLG (ternary system of BLG–urea–water). For comparison, the urea concentration-dependent IR spectra of urea solution itself (binary system of urea–water) was also shown. A simple analysis for comparing the relative absorbance changes of two bands demonstrated that the absorbance of the urea band for the C = O group is very sensitive to interactions with the BLG molecules.

In order to obtain detailed differences in the urea–water interaction caused by BLG, they applied PCA to the one data matrix that was combined from two sets of spectra for binary

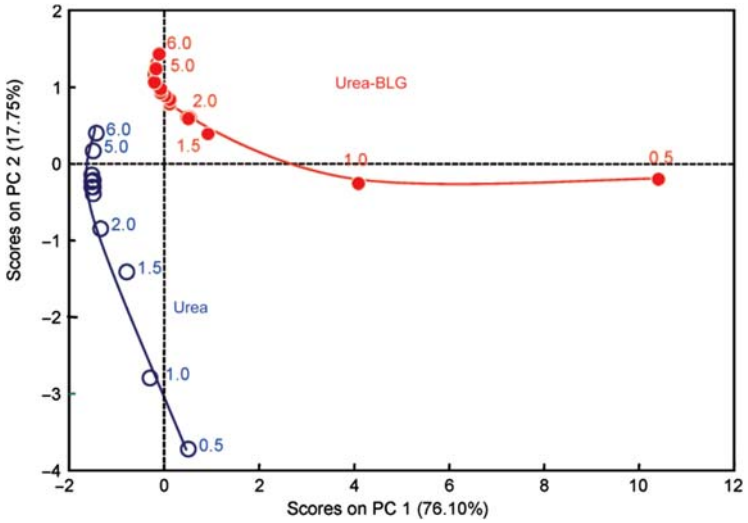


Figure 12.17

Scores plot of the spectra for the system with (·) and without (O) β-lactoglobulin (BLG) combined into one matrix on the first two principal components. Source: *Reproduced with permission from B. Czarnik-Matusewicz, S.B. Kim, Y.M. Jung, A study of urea-dependent denaturation of β-lactoglobulin by principal component analysis and two-dimensional correlation spectroscopy, J. Phys. Chem. B 113 (2009) 559–566, Copyright 2009, American Chemical Society.*

and ternary systems. As shown in Fig. 12.17, the plot of scores of the first two principal components indicated that the concentration-dependent association phenomenon of urea in aqueous solution has been strongly influenced by the interaction with BLG, mainly at concentrations lower than 2 M in which urea was involved in ribbon and chain dimers [115]. The most changes were captured by PC1, and PC2 described the absorbance changes that arose from urea-induced unfolding of BLG mainly occurring at concentrations above 2 M. Fig. 12.18 shows the plot of the loadings for the two components confirming the classification of the changes. In PCA results, the changes that are directly correlated with the amide I′ and amide II′ bands were detected. The denaturation properties in solutions above 3 M are due to the direct interaction between BLG and urea. To understand in more depth the full scenario for the interactions in the tertiary system and the urea-induced denaturation mechanism of BLG, they performed both conventional 2D-COS and heterosample 2D-COS. These results will be described in the following heterosample correlation analysis.

12.3.5 Two-dimensional heterocorrelation analysis in protein study

To provide detailed and useful information of protein structure, heterocorrelation analysis has been widely performed, which is very useful for the analysis of proteins.

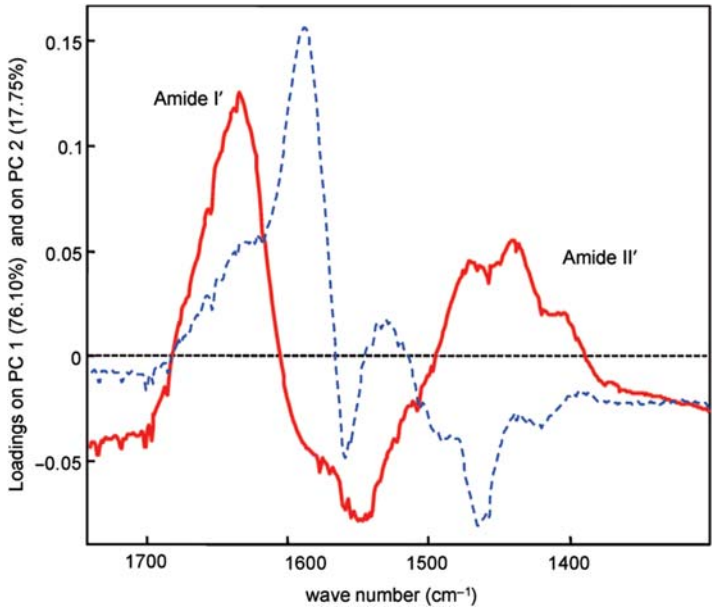


Figure 12.18

Loadings plot of the spectra for the two systems combined into one matrix for the PC1 (*solid line*) and PC2 (*dashed line*). Source: *Reproduced with permission from B. Czarnik-Matusewicz, S.B. Kim, Y.M. Jung, A study of urea-dependent denaturation of β-lactoglobulin by principal component analysis and two-dimensional correlation spectroscopy, J. Phys. Chem. B 113 (2009) 559–566, Copyright 2009, American Chemical Society.*

Heterospectral 2D-COS can compare the spectral changes induced by identical external perturbations and measured by closely related spectroscopic methods as well as by completely different types of spectroscopies. The heterospectral correlation mid-infrared (mid-IR)/Raman [68,104–106], Raman/ROA [89,128], electronic circular dichroism/Raman [106], and mid-IR/VCD [107] spectra of the protein have been performed to better understand the correlation between the bands attributed to the same secondary structure elements in the two types of spectroscopy. Heterospectral 2D-COS is very useful for interpreting the two sets of spectra obtained simultaneously with two different spectrometers under the same experimental perturbation. This approach has the powerful potential to probe conformational changes and to resolve the ambiguous order of structural changes during the transition process of proteins. It also helps to approve the denaturation mechanism of proteins.

Jung et al. reported heterospectral IR/Raman correlation spectroscopy for investigating the secondary structure of BLG in buffer solutions as a function of concentration of protein [68]. The heterospectral IR/Raman 2D-COS provided new insight into the correlation between the IR and Raman bands in the amide III region and confirmed the band

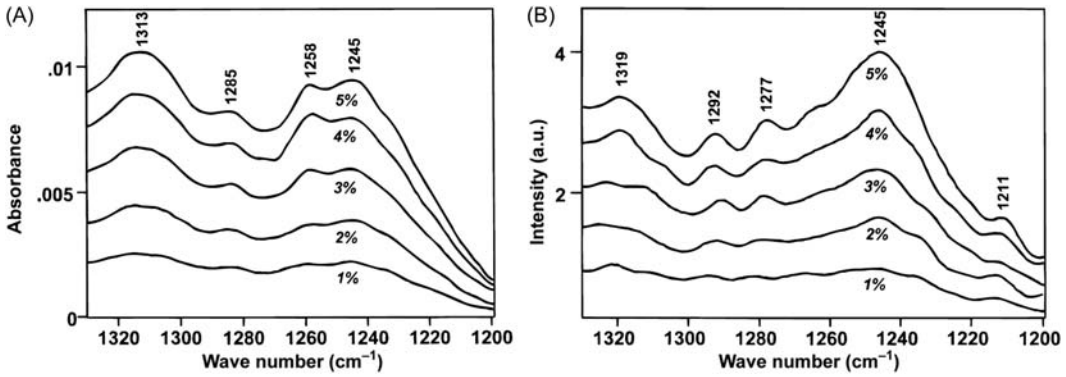


Figure 12.19

Infrared (IR) (A) and Raman (B) spectra in the amide III region of β-lactoglobulin (BLG) in buffer solutions (1, 2, 3, 4, and 5 wt.%). Source*: Reproduced with permission from Y.M. Jung, B. Czarnik-Matusewicz, Y. Ozaki, Two-dimensional infrared, two-dimensional Raman, and two-dimensional infrared and Raman heterospectral correlation studies of secondary structure of β-lactoglobulin in buffer solutions, J. Phys. Chem. B 104 (32) (2000) 7812–7817, Copyright 2000, with permission from American Chemical Society.*

assignments for the unordered form of BLG. Fig. 12.19 shows the concentration-dependent IR and Raman spectra of BLG buffer solutions in the amide III region (1330–1200 cm^{-1}). As shown in Fig. 12.19, four IR bands at 1313, 1285, 1258, and 1245 cm^{-1} and five Raman bands at 1319, 1292, 1277, 1245, and 1211 cm^{-1} are observed.

Figs. 12.20 and 12.21 show 2D IR and 2D Raman correlation spectra of BLG, respectively, as a function of concentration. It is particularly noticeable in the power spectra that the IR band at 1257 cm^{-1} and the Raman band at 1243 cm^{-1} show much greater intensity changes compared with the other bands. It means that the unordered component of BLG is most greatly affected by the association of the protein. Based on the analysis of 2D IR and Raman correlation spectra, the sequential order of intensity changes were determined: for the IR spectra, 1258 (unordered structure) → 1238, 1227 (β-sheet) → 1290, 1278, 1243, 1232, 1221, 1211 (β-turn) → 1265 (α-helixes) cm^{-1}; and for the Raman spectra, 1245, 1225 (unordered structure) → 1277, 1261 (β-sheet) → 1255, 1235 (β-turn) → 1283, 1298, 1268 (α-helixes) → 1214 (TYR residue) cm^{-1}. This reveals that the unordered component of BLG resides on the outer surface of the protein and other components (β-turns, β-sheets, and α-helices) of BLG are hardly accessible to water molecules.

To establish the band assignments in the IR and Raman spectra of BLG buffer solutions, 2D heterospectral IR/Raman analysis was performed. Fig. 12.22 shows 2D heterospectral IR/Raman correlation spectra of BLG. Generally, in the synchronous 2D heterospectral correlation spectrum, the positive cross peak means that two bands sharing the cross peak have the same origin or change in phase with the perturbation, while a negative cross peak means that two bands have different origins. In synchronous 2D heterospectral IR/Raman

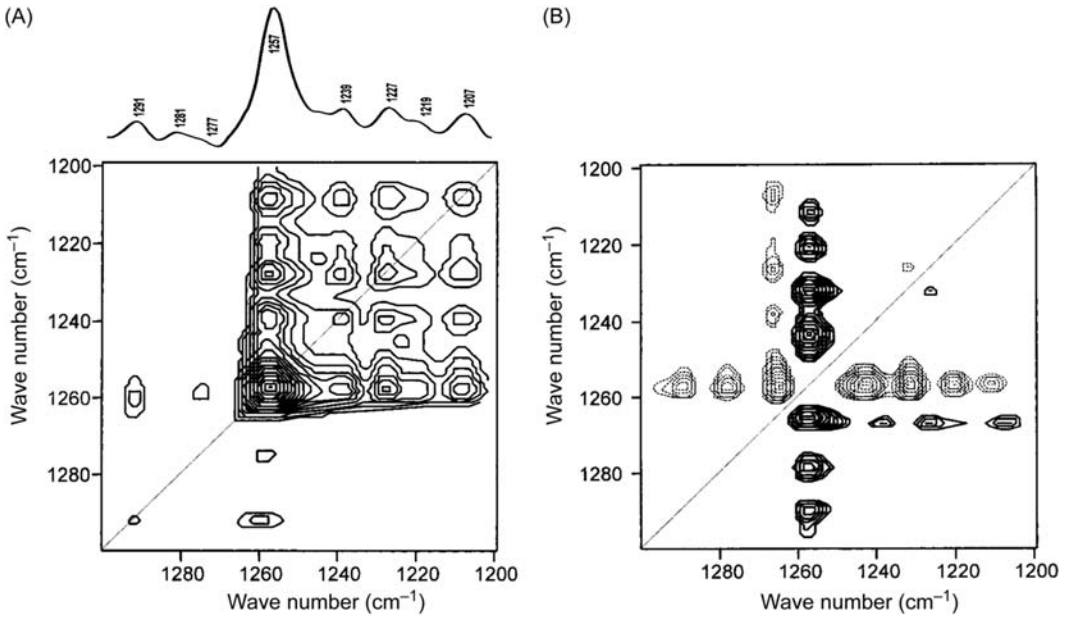


Figure 12.20

(A) Synchronous two-dimensional infrared (2D IR) correlation spectrum generated from concentration-dependent spectral variations of β-lactoglobulin (BLG) in buffer solutions. (B) Power spectrum along the diagonal line in the synchronous spectrum. Source*: Reproduced with permission from Y.M. Jung, B. Czarnik-Matusewicz, Y. Ozaki, Two-dimensional infrared, two-dimensional Raman, and two-dimensional infrared and Raman heterospectral correlation studies of secondary structure of β-lactoglobulin in buffer solutions, J. Phys. Chem. B 104 (32) (2000) 7812–7817, Copyright 2000, with permission from American Chemical Society.*

spectra, the positive cross peak at (1245, 1256) cm^{-1} refers to the Raman band at 1245 cm^{-1} and the IR band at 1256 cm^{-1} originating from the same unordered component. The asynchronous 2D heterospectral correlation spectrum clearly confirms that the unordered structures for the IR band at 1258 cm^{-1} and those for the Raman band at 1255 cm^{-1} have different origins.

Czarnik-Matusewicz et al. also reported on the heterosample 2D-COS to investigate the mechanism of urea-dependent denaturation of BLG and to confirm that the properties of the urea associates are affected by the presence of BLG [116]. The asynchronous 2D heterosample correlation spectrum is shown in Fig. 12.23. They confirmed that the presence of BLG affects the properties of the urea associates. The strong asynchronous peaks at 1563 and 1562 cm^{-1} for the hydrogen-bonded C = O groups in the ternary and binary systems, respectively, indicate that the changes proceed with different kinetics. It also suggests that the presence of BLG makes the urea dimers less stable and more sensitive to changes of urea concentration. Consequently, 2D heterosample correlation analysis would open a wide

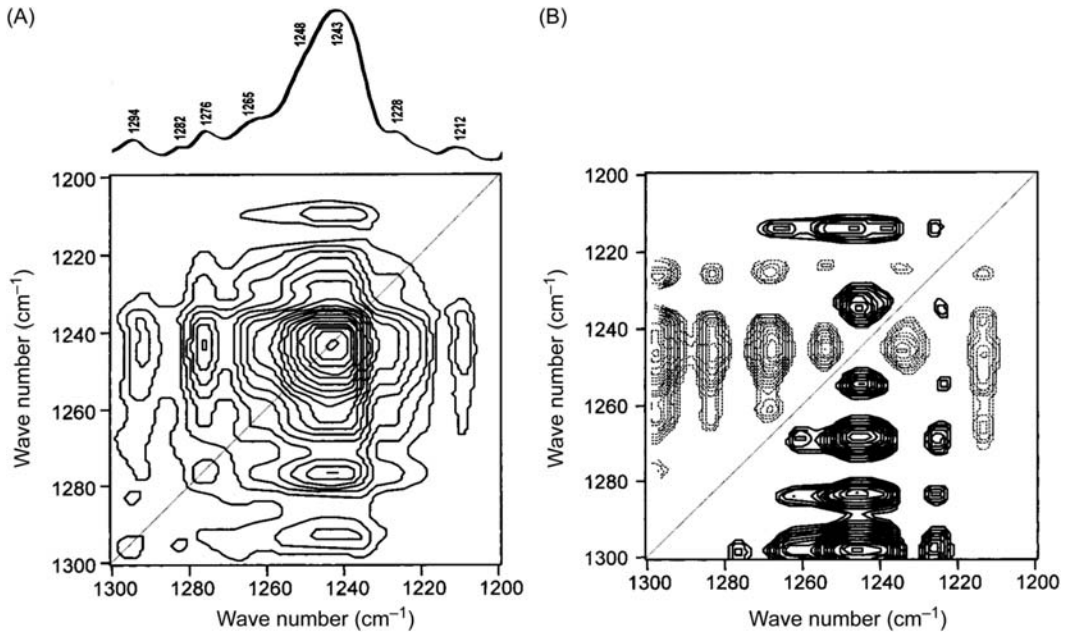


Figure 12.21

(A) Synchronous two-dimensional (2D) Raman correlation spectrum generated from concentration-dependent spectral variations of β-lactoglobulin (BLG) in buffer solutions. (B) Power spectrum along the diagonal line in the synchronous spectrum. Source*: Reproduced with permission from Y.M. Jung, B. Czarnik-Matusewicz, Y. Ozaki, Two-dimensional infrared, two-dimensional Raman, and two-dimensional infrared and Raman heterospectral correlation studies of secondary structure of β-lactoglobulin in buffer solutions, J. Phys. Chem. B 104 (32) (2000) 7812−7817, Copyright 2000, with permission from American Chemical Society.*

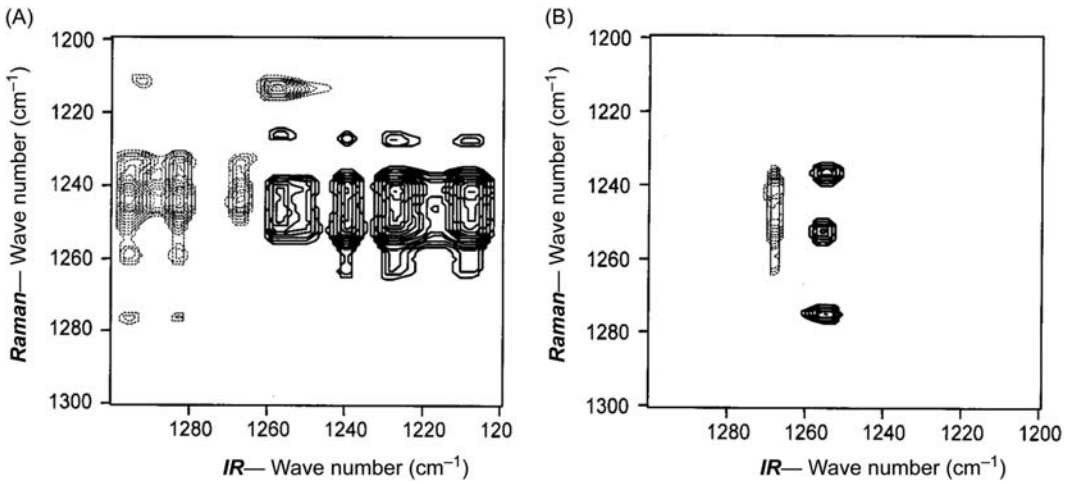


Figure 12.22

(A) Synchronous and (B) asynchronous two-dimensional infrared (2D IR)-Raman heterospectral correlation spectra generated from the concentration-dependent spectral variations of β-lactoglobulin (BLG) in buffer solutions. Source*: Reproduced with permission from Y.M. Jung, B. Czarnik-Matusewicz, Y. Ozaki, Two-dimensional infrared, two-dimensional Raman, and two-dimensional infrared and Raman heterospectral correlation studies of secondary structure of β-lactoglobulin in buffer solutions, J. Phys. Chem. B 104 (32) (2000) 7812−7817, Copyright 2000, with permission from American Chemical Society.*

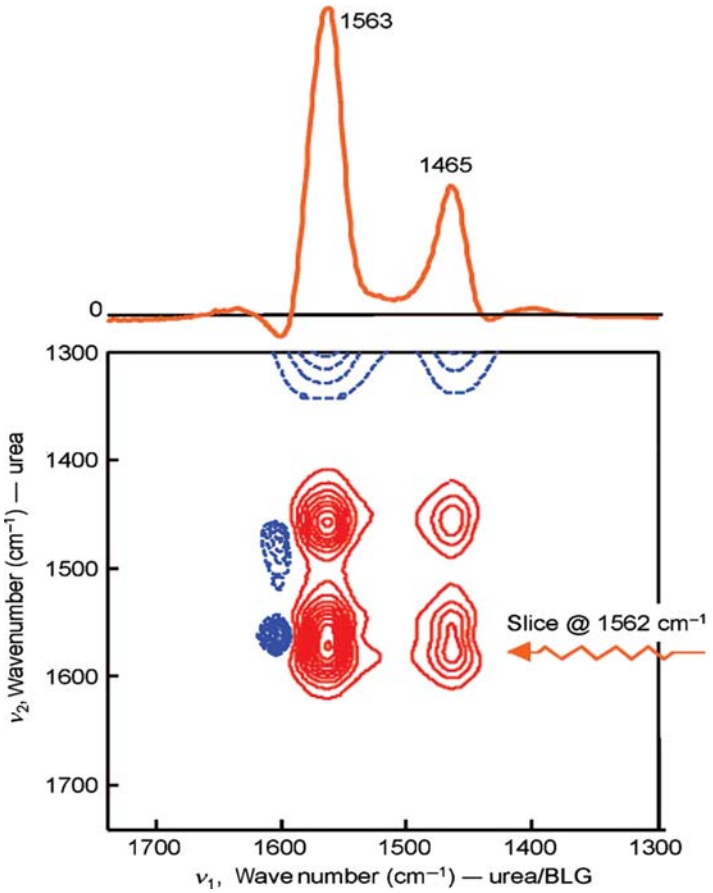


Figure 12.23
Heterosample asynchronous two-dimensional (2D) correlation spectrum for the two systems along with the slice spectrum extracted at the position of their maximal asynchronicity presented on top. Solid and dashed lines represent positive and negative cross peaks, respectively. Source: *Reproduced with permission from B. Czarnik-Matusewicz, S.B. Kim, Y.M. Jung, A study of urea-dependent denaturation of β-lactoglobulin by principal component analysis and two-dimensional correlation spectroscopy, J. Phys. Chem. B 113 (2009) 559–566, Copyright 2009, American Chemical Society.*

range of correlations, which can help give additional insights into the studied systems, compared with the conventional 2D-COS.

12.4 Future aspect

2D-COS has become a well-accepted powerful analytical technique widely used for many studies in various fields. 2D-COS continues to evolve and grow with new developments and successful applications in various scientific fields. Among the applications of 2D-COS in

biological systems, protein is the most actively studied sample. The 2D-COS studies in protein remain healthy, vigorous, and diverse in order to explore this fascinating class of biomolecules.

References

[1] I. Noda, Generalized two-dimensional correlation method applicable to infrared, Raman, and other types of spectroscopy, Appl. Spectrosc. 47 (9) (1993) 1329–1336.
[2] I. Noda, Two-dimensional infrared (2D IR) spectroscopy of synthetic and biopolymers, Bull. Am. Phys. Soc. 31 (1986) 520.
[3] I. Noda, Two-dimensional infrared (2-D IR) spectroscopy: theory and applications, Appl. Spectrosc. 44 (4) (1990) 550–561.
[4] I. Noda, Determination of two-dimensional correlation spectra using the Hilbert transform, Appl. Spectrosc. 54 (7) (2000) 994–999.
[5] I. Noda, Y. Ozaki, Two-dimensional Correlation Spectroscopy-Applications in Vibrational and Optical Spectroscopy, John Wiley & Sons, Chichester, 2004.
[6] I. Noda, Close-up view on the inner workings of two-dimensional correlation spectroscopy, Vib. Spectrosc. 60 (2012) 146–153.
[7] B. Czarnik-Matusewicz, Y.M. Jung, Two-dimensional mid-infrared correlation spectroscopy in protein research, in: M. Baranska (Ed.), Optical Spectroscopy and Computational Methods in Biology and Medicine, Vol. 14, Springer, Dordrecht, Germany, 2014, pp. 213–250.
[8] I. Noda, S.B. Lindsey, D. Carraway, Plastics from bacteria: natural functions and applications, in: G.-Q. Chen (Ed.), Microbiology Monographs, Vol. 14, Springer, Berlin, 2010, pp. 237–255.
[9] I. Noda, Generalized two-dimensional correlation spectroscopy, in: J. Laane (Ed.), Frontiers of Molecular Spectroscopy, Elsevier Science, The Netherlands, 2009, pp. 367–381.
[10] Y. Ozaki, I. Noda, Two-dimensional vibrational correlation spectroscopy in biomedical sciences, in: R.A. Meyers (Ed.), Encyclopedia of Analytical Chemistry, John Wiley & Sons, 2006, pp. 1–18.
[11] Y. Ozaki, S. Šašic, Two-dimensional correlation spectroscopy of biological and polymeric materials, in: M. Braiman, V.G. Gregoriou (Eds.), Vibrational Spectroscopy of Biological and Polymeric Materials, CRC Press, Taylor & Francis Group, 2006, pp. 163–214.
[12] Y. Ozaki, 2D correlation spectroscopy in vibrational spectroscopy, in: J.M. Chalmers, P. Griffiths (Eds.), Handbook of Vibrational Spectroscopy, Wiley & Sons, Chichester, 2002, pp. 2135–2172.
[13] I. Noda, General theory of two-dimensional (2D) analysis, in: J.M. Chalmers, P.R. Griffiths (Eds.), Handbook of Vibrational Spectroscopy, Vol. 3, John Wiley & Sons, Chichester, 2002, pp. 2123–2134.
[14] I. Noda, Two-dimensional correlation spectroscopy-biannual survey 2007-2009, J. Mol. Struct. 974 (2010) 3–24.
[15] I. Noda, Recent Advancement in the field of two-dimensional correlation spectroscopy, J. Mol. Struct. 883-884 (2008) 2–26.
[16] I. Noda, Two-dimensional correlation analysis useful for spectroscopy, chromatography, and other analytical measurements, Anal. Sci. 23 (2007) 139–146.
[17] I. Noda, Progress in two-dimensional (2D) correlation spectroscopy, J. Mol. Struct. 799 (2006) 2–15.
[18] Y.M. Jung, I. Noda, New approaches to generalized two-dimensional correlation spectroscopy and its applications, Appl. Spectrosc. Rev. 41 (2006) 515–547.
[19] I. Noda, Advances in two-dimensional correlation spectroscopy, Vib. Spectrosc. 36 (2004) 143–165.
[20] I. Noda, A.E. Dowrey, C. Marcott, G.M. Story, Y. Ozaki, Generalized two-dimensional correlation spectroscopy, Appl. Spectrosc. 54 (2000) 236A–248A.
[21] I. Noda, Progress in 2D correlation spectroscopy, in: Y. Ozaki, I. Noda (Eds.), Two-Dimensional Correlation Spectroscopy, AIP Press, Melville, 2000, pp. 3–17.

[22] I. Noda, Advances in two-dimensional correlation spectroscopy (2DCOS), in: J. Laane (Ed.), Frontiers and Advances in Molecular Spectroscopy, Elsevier, Amsterdam, 2018, pp. 47–75.
[23] I. Noda, Vibrational two-dimensional correlation spectroscopy (2DCOS) study of proteins, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 187 (2017) 119–129.
[24] Y. Park, I. Noda, Y.M. Jung, Novel developments and applications of two-dimensional correlation spectroscopy, J. Mol. Struct. 1124 (2016) 11–28.
[25] Y. Park, S. Jin, I. Noda, Y.M. Jung, Recent progresses in two-dimensional correlation spectroscopy (2D-COS), J. Mol. Struct. 1168 (2018) 1–21.
[26] Y.M. Jung, B. Czarnik-Matusewicz, Y. Wu, L. Zhang, I. Noda, Application of two-dimensional correlation spectroscopy in protein research, in: R.A. Meyers (Ed.), Encyclopedia of Analytical Chemistry: Applications, Theory and Instrumentation, John Wiley & Sons, Chichester, 2015, pp. 1–27.
[27] I. Noda, Vibrational two-dimensional correlation spectroscopy (2DCOS) study of proteins, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 187 (2017) 119–129.
[28] Y. Tao, Y. Wu, L. Zhang, Advancements of two dimensional correlation spectroscopy in protein researches, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 197 (2018) 185–193.
[29] J.L.R. Arrondo, I. Iloro, M. Garcia-Pacios, F.M. Goñi, Two-dimensional infrared correlation spectroscopy, in: J.L.R. Arrondo, A. Alonso (Eds.), Advanced Techniques in Biophysics, Vol. 10, Springer Series in Biophysics, 2006, pp. 73–88.
[30] V.A. Shashilov, I.K. Lednev, Two-dimensional correlation Raman spectroscopy for characterizing protein structure and dynamics, J. Raman. Spectrosc. 40 (12) (2009) 1749–1758.
[31] Y. Ozaki, K. Murayama, Y. Wu, B. Czarnik-Matusewicz, Two-dimensional infrared correlation spectroscopy studies on secondary structures and hydrogen bondings of side chains of proteins, Spectroscopy 17 (2–3) (2003) 79–100.
[32] A.I. Gomaa, J. Sedman, A.A. Ismail, An investigation of the effect of microwave treatment on the structure and unfolding pathways of β-lactoglobulin using FTIR spectroscopy with the application of two-dimensional correlation spectroscopy (2D-COS), Vib. Spectrosc. 65 (2013) 101–109.
[33] Q. Shao, Y. Qian, P. Wu, H. Zhang, C. Cai, Graphene oxide-induced conformation changes of glucose oxidase studied by infrared spectroscopy, Colloids Surf. B Biointerfaces 109 (2013) 115–120.
[34] M. Zhang, Y.Q. Dang, T.Y. Liu, H.W. Li, Y. Wu, Q. Li, et al., Pressure-induced fluorescence enhancement of the BSA-protected gold nanoclusters and the corresponding conformational changes of protein, J. Phys. Chem. C 117 (2013) 639–647.
[35] V.L. Brewster, L. Ashton, R. Goodacre, Monitoring guanidinium-induced structural changes in ribonuclease proteins using Raman spectroscopy and 2D correlation analysis, Anal. Chem. 85 (2013) 3570–3575.
[36] V. Sikirzhytski, N.I. Toplina, G.A. Takor, S. Higashiya, J.T. Welch, V.N. Uversky, et al., Fibrillation mechanism of a model intrinsically disordered protein revealed by 2D correlation deep UV resonance raman spectroscopy, Biomacromolecules 13 (2012) 1503–1509.
[37] A. Filosa, Y. Wang, A.A. Ismail, A.M. English, Two-dimensional infrared correlation spectroscopy as a probe of sequential events in the thermal unfolding of cytochromes c, Biochemistry 40 (28) (2001) 8256–8263.
[38] Y.B. Yan, Q. Wang, H.W. He, X.Y. Hu, R.Q. Zhang, H.M. Zhou, Two-dimensional infrared correlation spectroscopy study of sequential events in the heat-induced unfolding and aggregation process of myoglobin, Biophys. J. 85 (3) (2003) 1959–1967.
[39] I. Iloro, R. Chehin, F.M. Goni, M.A. Pajares, J.L. Arrondo, Methionine adenosyltransferase α-helix structure unfolds at lower temperatures than β-sheet: a 2D-IR study, Biophys. J. 86 (6) (2004) 3951–3958.
[40] T. Lefèvre, K. Arseneault, M. Pézolet, Study of protein aggregation using two-dimensional correlation infrared spectroscopy and spectral simulations, Biopolymers 73 (6) (2004) 705–715.

[41] Y.B. Yan, Q. Wang, H.W. He, H.M. Zhou, Protein thermal aggregation involves distinct regions: sequential events in the heat-induced unfolding and aggregation of hemoglobin, Biophys. J. 86 (3) (2004) 1682–1690.
[42] I. Iloro, F.M. Goni, J.L. Arrondo, A. 2D-IR study of heat- and [^{13}C] urea-induced denaturation of sarcoplasmic reticulum Ca^{2+}-ATPase, Acta Biochim. Pol. 52 (2) (2005) 477–483.
[43] J. Zhang, Y.B. Yan, Probing conformational changes of proteins by quantitative second derivative infrared spectroscopy, Anal. Biochem. 340 (1) (2005) 89–98.
[44] E. Saguer, P. Alvarez, J. Sedman, H.S. Ramaswamy, A.A. Ismail, Heat-induced gel formation of plasma proteins: new insights by FTIR 2D correlation spectroscopy, Food Hydrocoll. 23 (3) (2009) 874–879.
[45] L.X. Wang, L. Zhang, Y. Wu, Thermally induced early events of ribonuclease a under reducing conditions: evidenced by principal component analysis and two-dimensional correlation, Vib. Spectrosc. 51 (1) (2009) 4–10.
[46] E. Saguer, P. Alvarez, A.A. Ismail, Heat-induced denaturation/aggregation of porcine plasma and its fractions studied by FTIR spectroscopy, Food Hydrocoll. 27 (1) (2012) 208–219.
[47] T. Lefèvre, M. Subirade, Conformational rearrangement of β-lactoglobulin upon interaction with an anionic membrane, Biochim. Biophys. Acta 1549 (1) (2001) 37–50.
[48] K. Murayama, Y. Wu, B. Czarnik-Matusewicz, Y. Ozaki, Two-dimensional/attenuated total reflection infrared correlation spectroscopy studies on secondary structural changes in human serum albumin in aqueous solutions: pH-dependent structural changes in the secondary structures and in the hydrogen bondings of side chains, J. Phys. Chem. B 105 (20) (2001) 4763–4769.
[49] B.W. Hu, P. Zhou, I. Noda, Q.X. Ruan, Generalized two-dimensional correlation analysis of NMR and Raman spectra for structural evolution characterizations of silk fibroin, J. Phys. Chem. B 110 (36) (2006) 18046–18051.
[50] R. Dluhy, S. Shanmukh, S.I. Morita, The application of two-dimensional correlation spectroscopy to surface and interfacial analysis, Surf. Interface Anal. 38 (11) (2006) 1481–1496.
[51] N.L. Sefara, N.P. Magtoto, H.H. Richardson, Structural characterization of beta-lactoglobulin in solution using two-dimensional FT mid-infrared and FT near-infrared correlation spectroscopy, Appl. Spectrosc. 51 (4) (1997) 536–540.
[52] S. Shanmukh, P. Howell, J.E. Baatz, R.A. Dluhy, Effect of hydrophobic surfactant proteins SP-B and SP-C on phospholipid monolayers. Protein structure studied using 2D IR and βν correlation analysis, Biophys. J. 83 (4) (2002) 2126–2141.
[53] S. Shanmukh, N. Biswas, A.J. Waring, F.J. Walther, Z. Wang, Y. Chang, et al., Structure and properties of phospholipid-peptide monolayers containing monomeric SP-B1-25: II. Peptide conformation by infrared spectroscopy, Biophys. Chem. 113 (1) (2005) 233–244.
[54] B. Czarnik-Matusewicz, K. Murayama, Y. Wu, Y. Ozaki, Two-dimensional attenuated total reflection/infrared correlation spectroscopy of adsorption-induced and concentration-dependent spectral variations of β-lactoglobulin in aqueous solutions, J. Phys. Chem. B 104 (32) (2000) 7803–7811.
[55] K. Murayama, B. Czarnik-Matusewicz, Y. Wu, R. Tsenkova, Y. Ozaki, Comparison between conventional spectral analysis methods, chemometrics, and two-dimensional correlation spectroscopy in the analysis of near-infrared spectra of protein, Appl. Spectrosc. 54 (7) (2000) 978–985.
[56] K. Murayama, Y. Ozaki, Two-dimensional near-IR correlation spectroscopy study of molten globule-like state of ovalbumin in acidic pH region: simultaneous changes in hydration and secondary structure, Biopolymers 67 (6) (2002) 394–405.
[57] L. Smeller, K. Heremans, 2D FT-IR spectroscopy analysis of the pressure-induced changes in proteins, Vib. Spectrosc. 19 (2) (1999) 375–378.
[58] L. Smeller, K. Heremans, Hydrogen-deuterium exchange versus conformational changes in proteins. A two-dimensional FTIR approach, in: H. Ludwig (Ed.), Advances in High Pressure Bioscience and Biotechnology, Springer-Verlag, Berlin, 1999, pp. 223–226.

[59] W. Dzwolak, M. Kato, A. Shimizu, Y. Taniguchi, Comparative two-dimensional Fourier transform infrared correlation spectroscopic study on the spontaneous, pressure-, and temperature-enhanced H/D exchange in alpha-lactalbumin, Appl. Spectrosc. 54 (7) (2000) 963–967.
[60] L. Smeller, P. Rubens, J. Frank, J. Fidy, K. Heremans, Two dimensional Fourier-transform infrared correlation spectroscopy study of the high-pressure tuning of proteins, Vib. Spectrosc. 22 (2000) 119–125.
[61] W. Dzwolak, M. Kato, Y. Taniguchi, Fourier transform infrared spectroscopy in high-pressure studies on proteins, Biochim. Biophys. Acta 1595 (1-2) (2002) 131–144.
[62] A. Nabet, M. Pézolet, Two-dimensional FT-IR spectroscopy: a powerful method to study the secondary structure of proteins using H-D exchange, Appl. Spectrosc. 51 (4) (1997) 466–469.
[63] Y. Wu, K. Murayama, Y. Ozaki, Two-dimensional infrared spectroscopy and principle component analysis studies of the secondary structure and kinetics of hydrogen-deuterium exchange of human serum albumin, J. Phys. Chem. B 105 (26) (2001) 6251–6259.
[64] J.A. Richard, I. Kelly, D. Marion, M. Auger, M. Pézolet, Structure of β-purothionin in membranes: a two-dimensional infrared correlation spectroscopy study, Biochemistry 44 (1) (2005) 52–61.
[65] L. del Valle Sosa, E. Alfaro, J. Santiago, D. Narváez, M.C. Rosado, A. Rodríguez, et al., The structure, molecular dynamics, and energetics of centrin-melittin complex, Proteins Struct. Funct. Bioinform 79 (11) (2011) 3132–3143.
[66] H. Fabian, H.H. Mantsch, C.P. Schultz, et al., IR correlation spectroscopy: sequential events in the unfolding process of the λ Cro-V55C repressor protein, Proc. Natl Acad. Sci. U.S.A. 96 (23) (1999) 13153–13158.
[67] S. Meskers, J.M. Ruysschaert, E. Goormaghtigh, Hydrogen-deuterium exchange of streptavidin and its complex with biotin studied by 2D-attenuated total reflection Fourier transform infrared spectroscopy, J. Am. Chem. Soc. 121 (22) (1999) 5115–5122.
[68] Y.M. Jung, B. Czarnik-Matusewicz, Y. Ozaki, Two-dimensional infrared, two-dimensional Raman, and two-dimensional infrared and Raman heterospectral correlation studies of secondary structure of β-lactoglobulin in buffer solutions, J. Phys. Chem. B 104 (32) (2000) 7812–7817.
[69] M. Sonoyama, T. Nakano, Infrared rheo-optics of bombyx mori fibroin film by dynamic step-scan FT-IR spectroscopy combined with digital signal processing, Appl. Spectrosc. 54 (7) (2000) 968–973.
[70] M.J. Paquet, M. Laviolette, M. Pézolet, M. Auger, Two-dimensional infrared correlation spectroscopy study of the aggregation of cytochrome c in the presence of dimyristoylphosphatidylglycerol, Biophys. J. 81 (1) (2001) 305–312.
[71] Y. Wu, K. Murayama, B. Czarnik-Matusewicz, Y. Ozaki, Two-dimensional attenuated total reflection/infrared correlation spectroscopy studies on concentration and heat-induced structural changes of human serum albumin in aqueous solutions, Appl. Spectrosc. 56 (9) (2002) 1186–1193.
[72] A. Domínguez-Vidal, M.P. Saenz-Navajas, M.J. Ayora-Cañada, B. Lendl, Detection of albumin unfolding preceding proteolysis using Fourier transform infrared spectroscopy and chemometric data analysis, Anal. Chem. 78 (10) (2006) 3257–3264.
[73] I. Iloroa, B. Pastrana-Rios, Simulation of FT-IR spectra and 2D-COS analysis for the H/D exchange of two related ligands, J. Mol. Struct. 799 (1-3) (2006) 153–157.
[74] R.E. Sallach, M. Wei, N. Biswas, V.P. Conticello, S. Lecommandoux, R.A. Dluhy, et al., Micelle density regulated by a reversible switch of protein, J. Am. Chem. Soc. 128 (36) (2006) 12014–12019.
[75] S. Sánchez-Bautista, A. Kazaks, M. Beaulande, A. Torrecillas, S. Corbalán-García, J.C. Gómez-Fernández, Structural study of the catalytic domain of PKCzeta using infrared spectroscopy and two-dimensional infrared correlation spectroscopy, FEBS J. 273 (14) (2006) 3273–3286.
[76] A. Bernabeua, L.M. Contrerasb, J. Villalaín, Two-dimensional infrared correlation spectroscopy study of the interaction of oxidized and reduced cytochrome c with phospholipid model membranes, Biochim. Biophys. Acta Biomembr. 1768 (10) (2007) 2409–2420.

[77] I. Arada, J.P. Julien, B.G. Torre, N. Huarte, D. Andreu, E.F. Pai, et al., Structural constraints imposed by the conserved fusion peptide on the HIV-1 gp41 epitope recognized by the broadly neutralizing antibody 2F5, J. Phys. Chem. B 113 (41) (2009) 13626–13637.
[78] Y. Ozaki, K. Murayama, Y. Wang, Application of two-dimensional near-infrared correlation spectroscopy to protein research, Vib. Spectrosc. 20 (2) (1999) 127–132.
[79] Y. Wu, B. Czarnik-Matusewicz, K. Murayama, Y. Ozaki, Two-dimensional near-infrared spectroscopy study of human serum albumin in aqueous solutions: using overtones and combination modes to monitor temperature-dependent changes in the secondary structure, J. Phys. Chem. B 104 (24) (2000) 5840–5847.
[80] B. Czarnik-Matusewicz, Near-infrared spectroscopy combined with two-dimensional correlation spectroscopy in studies of biomolecules, in: S. Tsuchikawa (Ed.), Useful and Advanced Information in the Field of Near Infrared Spectroscopy, Research Signpost, 2002, pp. 189–216.
[81] B. Yuan, K. Murayama, Y. Wu, R. Tsenkova, X. Dou, S. Era, Temperature-dependent near-infrared spectra of bovine serum albumin in aqueous solutions: spectral analysis by principal component analysis and evolving factor analysis, Appl. Spectrosc. 57 (10) (2003) 1223–1229.
[82] L. Yu, B. Xiang, Two-dimensional near-IR correlation spectroscopy study the interaction of protein and famotidine in acidic pH region, Spectrochim. Acta A 69 (2) (2008) 599–603.
[83] T.J. Kamerzell, S. Kanai, J. Liu, S.J. Shire, Y.J. Wang, Increasing IgG concentration modulates the conformational heterogeneity and bonding network that influence solution properties, J. Phys. Chem. B 113 (17) (2009) 6109–6118.
[84] C. Mo, P. Wu, X. Chen, Z. Shao, The effect of water on the conformation transition of *Bombyx mori* silk fibroin, Vib. Spectrosc. 51 (1) (2009) 105–109.
[85] P. Zhou, X. Xie, D.P. Knight, X.H. Zong, F. Deng, W.H. Yao, Effects of pH and calcium ions on the conformational transitions in silk fibroin using 2D Raman correlation spectroscopy and 13C solid-state NMR, Biochemistry 43 (35) (2004) 11302–11311.
[86] L. Ashton, L.D. Barron, L. Hecht, J. Hyde, E.W. Blanch, Two-dimensional Raman and Raman optical activity correlation analysis of the α-helix-to-disordered transition in poly(L-glutamic acid), Analyst 132 (5) (2007) 468–479.
[87] L. Ashton, E.W. Blanch, pH-induced conformational transitions in α-lactalbumin investigated with two-dimensional Raman correlation variance plots and moving windows, J. Mol. Struct. 974 (1-3) (2010) 132–138.
[88] L. Ashton, C. Johannessen, R. Goodacre, The importance of protonation in the investigation of protein phosphorylation using Raman spectroscopy and Raman optical activity, Anal. Chem. 83 (20) (2011) 7978–7983.
[89] L. Ashton, L.D. Barron, B. Czarnik-Matusewicz, L. Hecht, J. Hyde, E.W. Blanch, Two-dimensional correlation analysis of Raman optical activity data on the α-helix-to-β-sheet transition in poly(L-lysine), Mol. Phys. 104 (9) (2006) 1429–1445.
[90] L. Ashton, B. Czarnik-Matusewicz, E.W. Blanch, Application of two-dimensional correlation analysis to Raman optical activity, J. Mol. Struct. 799 (1-3) (2006) 61–71.
[91] L. Ashton, E.W. Blanch, Investigation of polypeptide conformational transitions with two-dimensional Raman optical activity correlation analysis, applying autocorrelation and moving window approaches, Appl. Spectrosc. 62 (5) (2008) 469–475.
[92] T. Pazderka, V. Kopecký Jr., Two-dimensional correlation analysis of Raman optical activity—basic rules and data treatment, Vib. Spectrosc. 60 (2010) 193–199.
[93] L. Ma, Z. Ahmed, S.A. Asher, Ultraviolet resonance Raman study of side chain electrostatic control of poly-l-lysine conformation, J. Phys. Chem. B 115 (14) (2011) 4251–4258.
[94] V. Shashilov, I.K. Lednev, 2D correlation deep UV resonance Raman spectroscopy of early events of lysozyme fibrillation: kinetic mechanism and potential interpretation pitfalls, J. Am. Chem. Soc. 130 (1) (2008) 309–317.
[95] V. Shashilov, M. Xu, V.V. Ermolenkov, L. Fredriksen, I.K. Lednev, Probing a fibrillation nucleus directly by deep ultraviolet Raman spectroscopy, J. Am. Chem. Soc. 129 (22) (2007) 6972–6973.

[96] I.K. Lednev, M. Xu, V. Shashilov, Ultraviolet Raman spectroscopy is uniquely suitable for studying amyloid diseases, Curr. Sci. 97 (2) (2009) 180–185.
[97] S.A. Oladepo, K. Xiong, Z. Hong, S.A. Asher, J. Handen, I.K. Lednev, UV resonance Raman investigations of peptide and protein structure and dynamics, Chem. Rev. 112 (5) (2012) 2604–2628.
[98] S. Ma, T.B. Freedman, X. Cao, L.A. Nafie, Two-dimensional vibrational circular dichroism correlation spectroscopy: pH-induced spectral changes in L-alanine, J. Mol. Struct. 799 (1-3) (2006) 226–238.
[99] S.R. Ryu, B. Czarnik-Matusewicz, R.K. Dukor, L.A. Nafie, Y.M. Jung, Analysis of the molten globule state of bovine α-lactalbumin by using vibrational circular dichroism, Vib. Spectrosc. 60 (2012) 68–72.
[100] G. Wang, L. Geng, Statistical and generalized two-dimensional correlation spectroscopy of multiple ionization states. Fluorescence of neurotransmitter serotonin, Anal. Chem. 77 (1) (2005) 20–29.
[101] H. Fukuma, K. Nakashima, Y. Ozaki, I. Noda, Two-dimensional fluorescence correlation spectroscopy IV: resolution of fluorescence of tryptophan residues in alcohol dehydrogenase and lysozyme, Spectrochim. Acta A 65 (3-4) (2006) 517–522.
[102] G. Wang, Y. Gao, M.L. Geng, Generalized two-dimensional heterocorrelation analysis of spectrally resolved and temporally resolved fluorescence of the 8-anilino-1-naphthalenesulfonate-apomyoglobin complex with pH perturbation, J. Phys. Chem. B 110 (16) (2006) 8506–8512.
[103] C. Chen, B. Xiang, L. Yu, T. Wang, B. Zhao, The application of two-dimensional fluorescence correlation spectroscopy on the interaction between bovine serum albumin and paeonolum in the presence of Fe(III), Spectrosc. Lett. 41 (8) (2008) 385–392.
[104] Y. Jung, B. Czarnik-Matusewicz, Y. Ozaki, An IR vs. Raman 2D heterospectral correlation study of the secondary structure of β-lactoglobulin in aqueous solutions, in: Y. Ozaki, I. Noda (Eds.), Two-Dimensional Correlation Spectroscopy, Kobe-Sanda, 29 August–1 September 1999, AIP Conference Proceedings, vol. 503, AIP Press, Melville, NY, 1999, pp. 275–278.
[105] P. Pancoska, J. Kubelka, T.A. Keiderling, Novel use of a static modification of two-dimensional correlation analysis. Part I: comparison of the secondary structure sensitivity of electronic circular dichroism, FT-IR, and Raman spectra of proteins, Appl. Spectrosc. 53 (6) (1999) 655–665.
[106] J. Kubelka, P. Pancoska, T.A. Keiderling, Novel use of a static modification of two-dimensional correlation analysis. Part II: hetero-spectral correlations of protein Raman, FT-IR, and circular dichroism spectra, Appl. Spectrosc. 53 (6) (1999) 666–671.
[107] J. Kubelka, P. Pancoska, T.A. Keiderling, Spectroscopy of biological molecules: new directions, in: J. Greve, G.J. Puppels, C. Otto (Eds.), 8th European Conference on the Spectroscopy of Biological Molecules, Enschede, 29 August–2 September 1999, Kluwer Academic Publishers, 1999, pp. 67–68.
[108] H.W. He, J. Zhang, H.M. Zhou, Y.B. Yan, Conformational change in the C-terminal domain is responsible for the initiation of creatine kinase thermal aggregation, Biophys. J. 89 (4) (2005) 2650–2658.
[109] A. Mezzetti, L. Blanchet, A. de Juan, W. Leibl, C. Ruckebusch, Ubiquinol formation in isolated photosynthetic reaction centers monitored by time-resolved differential FTIR in combination with 2D correlation spectroscopy and multivariate curve resolution, Anal. Bioanal. Chem. 399 (2011) 1999–2014.
[110] L. Meng, J. Jin, G. Yang, T. Lu, H. Zhang, C.X. Cai, Nonenzymatic electrochemical detection of glucose based on palladium-single-walled carbon nanotube hybrid nanostructures, Anal. Chem. 81 (17) (2009) 7271–7280.
[111] D.L. Elmore, R.A. Dluhy, βν-Correlation analysis: a modified two-dimensional infrared correlation method for determining relative rates of intensity change, J. Phys. Chem. B 105 (45) (2001) 11377–11386.
[112] S. Shanmukh, R.A. Dluhy, kν correlation analysis. a quantitative two-dimensional IR correlation method for analysis of rate processes with exponential functions, J. Phys. Chem. A 108 (26) (2004) 5625–5634.
[113] L. Szyc, S. Pilorz, B. Czarnik-Matusewicz, FTIR-ATR investigations of an α-helix to β-sheet conformational transition in poly(l-lysine), J. Mol. Liq. 141 (3) (2008) 155–159.

[114] S. Navea, A. de Juan, R. Tauler, Modeling temperature-dependent protein structural transitions by combined near-IR and Mid-IR spectroscopies and multivariate curve resolution, Anal. Chem. 75 (20) (2003) 5592−5601.
[115] Y.M. Jung, B. Czarnik-Matusewicz, S.B. Kim, Characterization of concentration-dependent infrared spectral variations of urea aqueous solutions by principal component analysis and two-dimensional correlation spectroscopy, J. Phys. Chem. B 108 (2004) 13008−13014.
[116] B. Czarnik-Matusewicz, S.B. Kim, Y.M. Jung, A study of urea-dependent denaturation of β-lactoglobulin by principal component analysis and two-dimensional correlation spectroscopy, J. Phys. Chem. B 113 (2009) 559−566.
[117] A. Litwińczuk, S.R. Ryu, L.A. Nafie, J.W. Lee, H.I. Kim, Y.M. Jung, et al., The transition from the native to the acid-state characterized by multi-spectroscopy approach: study for the holo-form of bovine α-lactalbumin, Biochim. Biophys. Acta 2014 (1844) 593−606.
[118] F. Song, F. Wu, B. Xing, T. Li, W. Feng, J.P. Giesy, et al., Protonation-dependent heterogeneity in fluorescent binding sites in sub-fraction of fulvic acid using principle component analysis and two-dimensional correlation spectroscopy, Sci. Total Environ. 616-617 (2018) 1279−1287.
[119] Y. Park, Y. Seo, B. Chae, D. Pyo, H. Chung, H. Hwang, et al., Understanding the structural difference between spherical and rod-shaped human insulin nanoparticles produced by supercritical fluids precipitation, ChemPhysChem 16 (2015) 476−482.
[120] Y. Park, Y. Park, S. Jin, J.W. Kim, Y.M. Jung, Formation mechanism of BAMLET by 2D Raman correlation analysis, J. Mol. Struct. 1171 (2018) 33−37.
[121] Y. Park, Y. Kim, Y. Park, S. Jin, H. Hwang, Y.M. Jung, Formation mechanism of α-lactalabumin/oleic acid complex characterized by 2D correlation analysis, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 185 (2017) 93−97.
[122] Y. Park, Y. Kim, K. Vikram, B. Czarnik-Matusewicz, Y.M. Jung, Two-dimensional correlation analysis of pH-induced Raman spectral changes of α-lactalbumin, Bull. Korean Chem. Soc. 37 (2016) 783−785.
[123] A. Wesełucha-Birczyńska, P. Moskal, M. Dużyja, E. Stodolak-Zych, E. Długoń, S. Kluska, et al., 2D correlation Raman spectroscopy of model micro- and nano-carbon layers in interaction with albumin, human and animal, J. Mol. Struct. 1171 (2018) 587−593.
[124] A. Rodrígyez Nassif, I. de la Arada, J. Luis Arrondo, B. Pastrana-Rios, 2D IR correlation spectroscopy in the determination of aggregation and stability of KH domain GXXG loop peptide in the presence and absence of trifluoroacetate, Anal. Chem. 89 (11) (2017) 5765−5775.
[125] M. Sánchez, A. Scirè, F. Tanfani, A. Ausili, The thermal unfolding of the ribosome-inactivating protein saporin-S6 characterized by infrared spectroscopy, Biochim. Biophys. Acta 1854 (2015) 1357−1364.
[126] Y. Suzuki, H. Furuya, Insights into the helix-sense inversion of poly(β-phenethyl-L-aspartate) by two-dimensional Raman correlation spectroscopy, J. Pept. Sci. 24 (4-5) (2018) 1−6.
[127] Z. He, T. Zhao, X. Zhou, Z. Liu, H. Huang, Sequential order of the secondary structure transitions of proteins under external perturbations: regenerated silk fibroin under thermal treatment, Anal. Chem. 89 (2017) 5534−5541.
[128] G. Ramer, L. Ashton, Two-dimensional codistribution spectroscopy applied to UVRR and ROA investigations of biomolecular transitions, J. Mol. Struct. 1124 (2016) 173−179.
[129] N. Andraka, L. Sánchez-Magraner, M. Garćia-Pecios, F.M. Goñi, J.L.R. Arrondo, The conformation of human phospholipid scramrase 1, as studied by infrared spectroscopy. Effects of calcium and detergent, Biochim. Biophys. Acta 1859 (2017) 1019−1028.
[130] A. Díaz Casas, W.J. Chazin, B. Pastrana-Rios, Prp40 homolog A is a novel centrin target, Biophys. J. 112 (2017) 2529−2539.
[131] B. Pastrana-Rios, L. del Valle Sosa, J. Santiago, Trifluoroacetic acid as excipient destabilizes melittin causing the selective aggregation of melittin within the centrin-melittin-trifluoroacetic acid complex, Struct. Dyn. 2 (2015) 041711.

[132] R. Lu, W.W. Li, A. Katzir, Y. Raichlin, B. Mizaikoff, H.Q. Yu, Fourier transform infrared spectroscopy on external perturbation inducing secondary structure changes of globin, Analyst 141 (2016) 6061–6067.
[133] R. Lu, W.W. Li, A. Katzir, Y. Raichlin, H.Q. Yu, B. Mizaikoff, Probing the secondary structure of bovine serum albumin during heat-induced denaturation using mid-infrared fiberoptic sensors, Analyst 140 (2015) 765–770.
[134] B. Pastrana-Rios, L. del Valle Sosa, J. Santiago, Trifluoroacetic acid as excipient destabilizes melittin causing the selective aggregation of melittin within the centrin-melittin-trifluoroacetic acid complex, Struct. Dyn. 2 (2015) 01711/1–01711/11.
[135] P.F. Gao, S. Zhang, H.W. Li, T. Zhang, Y. Wu, L. Wu, et al., Process of Eu-containing polyoxometalates to bovine serum albumin, Langmuir 31 (2015) 10888–10896.
[136] R. Yang, G. Dong, X. Sun, Y. Yu, H. Liu, Y. Yang, et al., Synchronous-asynchronous two-dimensional correlation spectroscopy for the discrimination of adulterated milk, Anal. Methods 7 (2015) 4302–4307.
[137] T. Maqbool, J. Hur, Changes in fluorescent dissolved organic matter upon interaction with anionic surfactant as revealed by EEM-PARAFAC and two-dimensional correlation spectroscopy, Chemosphere 161 (2016) 190–199.
[138] M. Bonsaii, K. Gholivand, K. Abdi, A.A.E. Valmoozi, M. Khosravi, A combined experimental and computational study on the interaction of nitrogen mustards with DNA, MedChemComm. 7 (2016) 2003–2015.
[139] L. Duan, J. Yuan, X. Yang, X. Cheng, J. Li, Interaction study of collagen and sericin in blending solution, Int. J. Biol. Macromol. 91 (2016) 468–475.
[140] M. Zhang, C. Ding, J. Yang, S. Lin, L. Chen, L. Huang, Study of interaction between water-soluble collagen and carboxymethyl cellulose in neutral aqueous solution, Carbohydr. Polym. 137 (2016) 410–417.
[141] J. Sacharz, A. Wesełucha-Birczyńska, C. Paluszkiewicz, P. Chaniecki, M. Błażewicz, A 2D correlation Raman spectroscopy analysis of a human cataractous lens, J. Mol. Struct. 1124 (2016) 71–77.
[142] J. Sacharz, A. Wesełucha-Birczyńska, J. Zięba-Palus, et al., Epileptic rat brain tissue analyzed by 2D correlation Raman spectroscopy, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 188 (2018) 581–588.
[143] M. Mecozzi, E. Sturchio, Effects of essential oil treatments on the secondary protein structure of *Vicia faba*: a mid-infrared spectroscopic study supported by two-dimensional correlation analysis, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 137 (2015) 90–98.
[144] W. Yan, H. Wang, C. Jing, Adhesion of *Shewanella oneidensis* MR-1 to goethite: a two-dimensional correlation spectroscopic study, Environ. Sci. Technol. 50 (2016) 4343–4349.
[145] Z. Ren, L. Zhang, Y. Wu, S. Mebarek, R. Buchet, Determination of phosphatase activity in osteoblasts by IR and two-dimensional correlation IR spectroscopy, Vib. Spectrosc. 86 (2016) 206–211.
[146] E.E. Najbauer, B. Bazsó, R. Apóstlo, R. Fausto, M. Biczysko, V. Barone, et al., Identification of serine conformers by matrix-isolation IR spectroscopy aided by near-infrared lase-induced conformational change, 2D correlation analysis, and quantum mechanical anharmonic computations, J. Phys. Chem. B 119 (2015) 10496–10510.
[147] B. Chae, S.H. Son, Y.J. Kwak, Y.M. Jung, Two-dimensional (2D) infrared correlation study of the structural characterization of a surface immobilized polypeptide film stimulated by pH, J. Mol. Struct. 1124 (2016) 192–196.
[148] P. Pieta, M. Majewska, Z. Su, M. Grossutti, B. Wladyka, M. Piejko, et al., Physicochemical studies on orientation and coenformation of a new bacteriocin BacSp222 in a planar phospholipid bilayer, Langmuir 33 (2016) 5653–5662.
[149] M.P. Schmidt, C.E. Martínez, Kinetic and conformational insights of protein adsorption onto montmorillonite revealed using in situ ATR-FTIR/2D-COS, Langmuir 32 (2016) 7719–7729.
[150] S. Cao, Y. Zeng, B. Cheng, W. Zhang, B. Liu, Effect of pH on Al/Zr-binding sites between collagen fibers in tanning process, J. Am. Leather Chem. Assoc. 111 (2016) 242–249.

[151] A. Hughes, Z. Liu, M. Reeves, Scikit-spectra: explorative spectroscopy in Python, J. Open Res. Softw. 3 (2015) e6/1–e6/6.
[152] Y. Wu, J.-H. Jiang, Y. Ozaki, A. new possibility of generalized two-dimensional correlation spectroscopy: hybrid two-dimensional correlation spectroscopy, J. Phys. Chem. A 106 (2002) 2422–2429.
[153] I. Noda, Scaling techniques to enhance two-dimensional correlation spectra, J. Mol. Struct. 883–884 (2008) 216–227.
[154] O. Tarhan, E. Tarhan, S. Harsa, Investigation of the structure of alphalactalbumin protein nanotubes using optical spectroscopy, J. Dairy Res. 81 (2014) 98–106.
[155] G. Zhu, X. Zhu, Q. Fan, X. Wan, Raman spectra of amino acids and their aqueous solutions, Spectrochim. Acta Mol. Biomol. Spectrosc. 78 (2011) 1187–1195.
[156] B.M. Bussian, C. Sander, How to determine protein secondary structure in solution by Raman spectroscopy: practical guide and test case DNase I, Biochemistry 28 (10) (1986) 4271–4277.
[157] Malvern Instruments Limited, Understanding the Conformational Stability of Protein Therapeutics Using Raman Spectroscopy, Malvern Panalytical, Malvern, UK, 2014.
[158] T. Nakamura, T. Aizawa, R. Kariya, S. Okada, M. Demura, K. Kawano, et al., Molecular mechanisms of the cytotoxicity of human a-lactalbumin made lethal to tumor cells (HAMLET) and other protein-oleic acid complexes, J. Biol. Chem. 288 (20) (2013) 14408–14416.

This page intentionally left blank

CHAPTER 13

Resonance Raman spectroscopy of hemoglobin in red blood cells

Bayden R. Wood[1], Kamila Kochan[1] and Katarzyna M. Marzec[2]

[1]Centre for Biospectroscopy, School of Chemistry, Monash University, Clayton, VIC, Australia,
[2]Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

13.1 Introduction

Resonance Raman spectroscopy (RRS) has proven to be the go-to technique for elucidating the structure of hemoglobin (Hb) and its derivatives. The high symmetry and chromophoric structure of the central heme group (Fig. 13.1) provides a spectrum rich in detail, especially when the excitation laser wavelength is in the vicinity of an electronic transition associated with the heme. Tobin [1] reported the first Raman spectrum of proteins, namely crystalline lysozyme, pepsin, and alpha chymotrypsin. Strekas and Spiro [2] and, independently, Brunner and coworkers [3] recorded the first resonance Raman spectra of Hb in both the oxygenated and deoxygenated state. The spectra showed dramatic changes between the two states associated with porphyrin distortions as the oxidation and spin state of the iron ion changed due to oxygen ligation. Brunner [4] was the first to assign the $\nu(\mathrm{Fe-O_2})$ mode to a band at $\sim 567\ \mathrm{cm}^{-1}$. This assignment was questioned by Benko and Yu [5] who suggested that this band was from the $\delta\ (\mathrm{Fe-O_2})$. However, another study by Bajdor investigating oxyphthalocyanto-iron (II) noted two isotopic sensitive bands in this region and presented an argument supporting Brunner's original assignment [6]. Interestingly no Raman studies had reported the $\delta(\mathrm{Fe-O_2})$ mode in oxygenated Hb (oxyHb) and the assignment had to be inferred to from Bajdor et al.'s [6] study on oxyphthalocyanto-iron (II). It was only recently that both ligand modes were observed in the spectra of Hb in red blood cells (RBCs) using a highly optimized 785-nm Raman skin analyzer system and finally the $\delta(\mathrm{Fe-O_2})$ mode from Hb was assigned to a band at $419\ \mathrm{cm}^{-1}$ [7]. The detection of the Fe−CO bond in carbonmonoxyHb (HbCO) was also difficult because of the photodissociation of the CO group at longer excitation wavelengths. The Fe−CO stretching and Fe−C−O bending vibrations were assigned much later by Tsubaki et al. [8] and independently by Armstrong et al. [9]. Strekas and Spiro went on to discover unusual polarization characteristics of hemes including inverse polarization observed in cytochrome C, where the polarization

Vibrational Spectroscopy in Protein Research.
DOI: https://doi.org/10.1016/B978-0-12-818610-7.00013-X

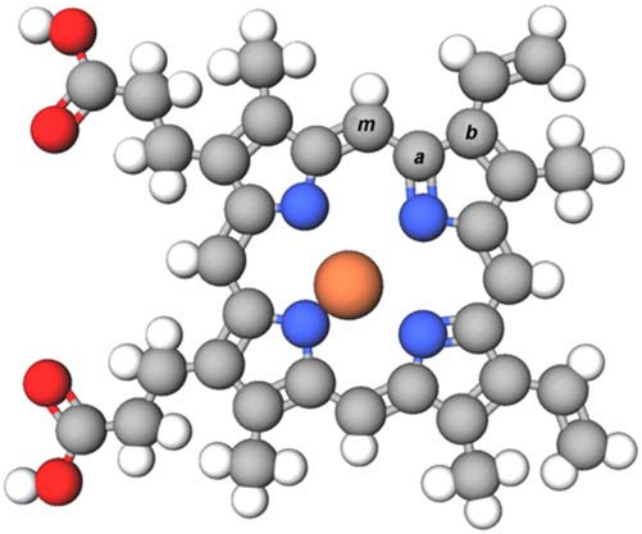


Figure 13.1
Iron protoporphyrin IX the with Fe (*orange*), N (*blue*), O (*red*), C (*gray*), and H (*white*). Carbon atoms are labeled C_a, C_b, and C_m for band assignments in Table 13.1.

vector of the incident radiation is rotated through 90 degrees for 90 degrees scattering, giving infinite depolarization ratios [10]. Further studies investigated the oxidation and spin state marker bands of Hb derivatives and identified six bands which offered promise as oxidation and/or spin state markers by analyzing oxyHb and HbCO as low-spin Fe(III) hemes [11]. A band assigned to the pyrrole half ring stretching vibration (ν_4) became known as the oxidation state marker band. Yamamoto et al. [12] noted the appearance of ν_4 in isolated heme derivatives correlated well with the oxidation state of the central Fe ion when using excitation into the Soret region. In oxyHb, where the Fe ion is in the ferric state, ν_4 appeared at approximately 1370−1378 cm^{-1}, while in deoxygenated Hb (deoxyHb), where it is in the ferrous state, it appeared at around 1356−1361 cm^{-1}. According to Yamamoto et al. [12] this supported the Weiss model [13] that describes how the iron−oxygen bond changes upon ligation. Weiss hypothesized that deoxyHb withdraws an electron from the ferrous ion resulting in a low-spin ferric state where the oxygen is converted to superoxide [13]. The resulting diamagnetism results in antiferromagnetic exchange interaction between the low-spin ferric ($S = 1/2$) and the O_2^- ($S = 1/2$) of the ligand.

Spaulding et al. [14] questioned the correlation of ν_4 as the ferric/ferrous marker band on the investigation of a wide range of metalloporphyrins and found ν_4 to be relatively invariant despite the large anticipated differences in charge density migration from the various metals to the conjugated porphyrin ring. The result implied that the charge distribution around the central iron atom is not a fundamental factor in the positioning of this band.

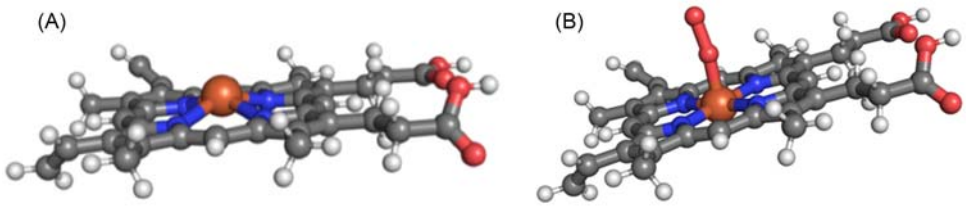


Figure 13.2
(A) Structure of deoxygenated Hb and (B) oxygenated Hb.

The work in the 1970s and early 1980s led to a plethora of knowledge about understanding the dynamical structure and Raman scattering processes of hemes and heme derivatives, which have been extensively reviewed [15–23]. While a myriad of studies focused on Hb isolated from RBCs, no Raman studies investigated Hb in its native environment of the red blood prior to the 21st century. Interestingly Raman spectroscopy had been extensively applied to investigate RBCs ghost membranes [24–30], however, the first publication by our group on applying Raman spectroscopy to functional RBCs did not appear until 2001 [31]. Since then a number of works on the subject have appeared in the literature investigating the unusual properties of Hb and its derivatives in the RBC along with a host of new applications in diagnosing pathological processes and monitoring blood storage for quality control.

13.2 Molecular structure of oxygenated and deoxygenated hemoglobin

Fig. 13.2 shows the structure of oxyHb and deoxyHb. In the ferrous high-spin state the Fe is approximately 4 nm out of the porphyrin plane and hence domed and exhibits idealized C_{4v} symmetry [32]. On the other hand, in the ferric low-spin state the porphyrin becomes more planar due to the translocation of the Fe into the plane resulting in idealized D_{4h} symmetry with the inclusion of the inversion symmetry element. As the Fe atom moves closer to the porphyrin plane it can distort thus reducing its planarity and in the process increasing nontotally symmetric components and shown with other metalloporphyrins [33].

13.3 Electronic structure of hemoglobin

Central to explaining the resonance Raman spectra of heme molecules is an understanding of the major electronic transitions that characterize the Hb molecule. Fig. 13.3 shows the UV–vis spectra of oxyHb and deoxyHb. The spectra of oxyHb are dominated by the Soret band at $\sim$413 nm along with two weaker bands including the Q_0 (or α) at 550 nm and the Q_v (or β) at 525 nm. The electronic transitions on the Gouterman 4 orbital model is shown in Fig. 13.3. Based on the idealized D_{4h} symmetry the of the tetrapyrrole ring the lowest lying unoccupied π^* orbitals are degenerated with e_g symmetry, while the highest occupied

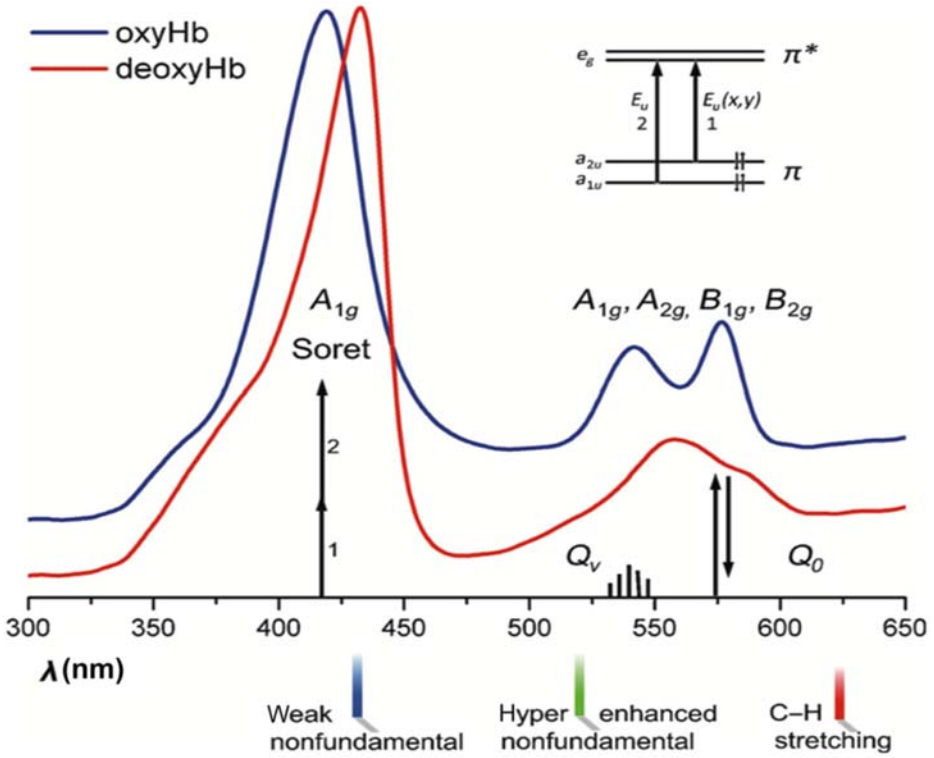


Figure 13.3

Comparison between average UV–vis electronic absorption spectra collected from 10 different oxygenated and deoxygenated RBCs in the spectral regions of 300–650 nm. The figure shows the allowed electronic transitions based on the Gouterman four-orbital model along with the laser lines used in this experiment (*color coded*) and the symmetry terms of the bands that are enhanced at the three major electronic transitions. Source*: Reproduced from K.M. Marzec, et al., Red blood cells polarize green laser light revealing hemoglobin's enhanced non-fundamental Raman modes. Chemphyschem 15(18) (2014) 3963–3968 [34] with permission from John Wiley & Sons.*

orbitals have a_{1u} and a_{2u} symmetry and also degenerate. The equivalent energy of the occupied orbitals gives rise to strong configuration interaction, which results in the adding of the $a_{1u} \rightarrow e_g$ and $a_{2u} \rightarrow e_g$ transition dipoles to form the Soret (or B band) and almost canceling each other out to form the Q_0 band. Some of the intensity is regained by the Q_0 band through vibronic mixing with the Soret band giving rise to the vibronic Q_v side band, which shows vibrational structure at low temperatures.

DeoxyHb is a ferrous high-spin derivative that in addition to the $\pi \rightarrow \pi^*$ transitions has four additional transitions that were deduced by Mössbauer and magnetic susceptibility studies. These are known as Bands I–IV and have been assigned based on single-crystal

polarization data. Band I is a charge transfer band that appears at 910 nm and is assigned to the almost purely z-polarized transition $d_{xz} \rightarrow e_g$. Band II appears at 813 nm and is mostly z-polarized and is assigned to the $d_{xz} \rightarrow d_{z^2}$. Band III is a charge transfer band that is in-plane polarized and occurs at 758 nm $a_{2u}(\pi) \rightarrow d_{yz}$, while band IV is also an in-plane polarized charge transfer band assigned to $a_{1u}(\pi) \rightarrow d_{yz}$.

Oxygenated Hb is a ferric low-spin heme. In the near-IR it exhibits a very broad transition centered at 925 nm. Circular dichroism and magnetic circular dichroism have shown this band to be made up of four bands, designated bands I–IV. Unlike deoxyHb there are no longer three orbitals designated d_{xz}, d_{xy}, and d_{yz} because of the interaction with the oxygen ligand. The d_{xy} orbital is relatively pure but the d_{xz} and d_{yz} are strongly mixed with the oxygen π_g orbitals producing four low-lying molecular orbitals.

Band I is assigned to the $d_{yz} + O_2(\pi_g) \rightarrow d_{xz} + O_2(\pi_g)$ and appears at 1300 nm, while band II appears at 1150 nm and is assigned to the $d_{x^2-y^2} + O_2(\pi_g) \rightarrow d_{xz} + O_2(\pi_g)$ charge transfer transition. Band III appears at 980 nm and is assigned to the x polarized band $a_{2u} \rightarrow d_{xz} + O_2(\pi_g)$, while band IV is x,y polarized and is assigned to the $a_{1u} \rightarrow d_{xz} + O_2(\pi_g)$ transition.

Exciting into the Soret band region gives rise to Type A scattering and the enhancement of totally symmetric modes. In this type of enhancement, the change in the nuclear coordinate results in maximal overlap between the wave function of the ground and excited state giving rise to totally symmetric modes A_{1g} (ν_1–ν_9). Excitation into the Q_0 band at $\sim$575 nm gives rise to the enhancement of both symmetrical and nonsymmetrical modes through Type B scattering, which involves vibronic coupling of excited states. These bands are labeled $A_{2g}(\nu_{10}-\nu_{17})$, $B_{1g}(\nu_{18}-\nu_{26})$, $B_{2g}(\nu_{27}-\nu_{35})$, and $E_u(\nu_{36}-\nu_{53})$. The "g" or gerade modes are Raman active except for A_{2g}, which is both infrared and Raman active, while the E_u modes are only infrared active, where "u" refers to the ungerade modes. As will be discussed, the enhancement into the Q_v band at $\sim$525 nm gives C-term enhancement and a rise to intense overtone and combination bands, while excitation wavelengths into the near-IR results in the protein and CH stretching modes becoming more pronounced because of the weaker resonant enhancement from the $\pi \rightarrow \pi^*$ transitions at these longer wavelengths.

13.4 Methods of recording hemes in cells

Raman microspectroscopy has been applied to detect and monitor heme groups within single cells. Some of the first applications of the technique were investigating myeloperoxidase in living neutrophils [35] and eosinophil peroxidase in living eosinophilic granulocytes [36]. Raman microscopy was also used to detect intracellular NADPH-oxidase activity in eosinophilic and neutrophilic granulocytes after mitogenic stimulation [37].

Raman imaging has also been applied to investigate the spatial distribution of hemes within cells [38]. Care must be taken when recording spectra of living cells to avoid photothermal damage and substrate-induced alterations [39,40]. Using 514-nm excitation Ramser et al. [40] reported substrate-induced effects that influenced Raman images and spectra of single functional erythrocytes. These included changes in the cell membrane mitigated by the poly-L-lysine used to attach the cells, and a photoinduced fluorescence effect that was hypothesized to be from the conversion of oxyHb to metHb, which masked any chance of obtaining reliable images of cells in the oxygenated and deoxygenated state [40]. Wood et al. [39] using 632.8-nm excitation showed that high laser power (>3 mW) and increasing temperature resulted in heme aggregation leading to an increase in five bands designated heme aggregation bands at 1396, 1365, 1248, 972, and 662 cm^{-1}. The formation of heme aggregates was confirmed by UV−vis spectra of erythrocytes, which showed the Soret band to be broader and red shifted ($\sim$3 nm) at temperatures between 45°C and 55°C, indicative of excitonic interactions between aggregated hemes. It was hypothesized that the enhancement of the aggregation marker bands results mainly in excitonic interactions between the aggregated hemes in response to protein denaturation [39]. The effects of fixation and dehydration on the distribution of hemes inside RBCs were investigated using Raman mapping and atomic force microscopy (AFM) topographic imaging [41]. When cells were dried in phosphate-buffered saline hemichrome appeared close to the periphery of the cell. Formaldehyde was found to be inadequate for complete structural preservation resulting in Hb diffusion into the surrounding area. The best results in terms of preserving cell morphology and minimizing autofluorescence was found to be a mixture of formaldehyde (3%) and glutaraldehyde (0.1%) in buffer [41].

13.5 Resonant Raman spectroscopy of hemoglobin in red blood cells

A detailed description of resonance Raman of hemes is beyond the scope of this chapter and one should refer to a number of reviews on the subject [16,19,42−46]. In the context of Hb inside RBCs the resonance Raman scattering profiles are significantly different to isolated Hb and more resemble solid-state spectra of the molecule. This is due in part to the high concentration of Hb and also the intrinsic ordering of hemes within the red cell [39]. It was previously postulated that the RBCs are filled with a highly concentrated solution Hb and it constitutes about 98% of the RBCs' solids [47]. The molar concentration of a single Hb subunit is in the range of 19−22 mM, while the tetrameric Hb molecule is in the range of 4.7−5.6 mM [47]. Fig. 13.4A shows the spectra of oxygenated RBCs at various excitation wavelengths, while Fig. 13.4B shows the corresponding deoxygenated RBCs spectrum using the same excitation wavelengths [48].

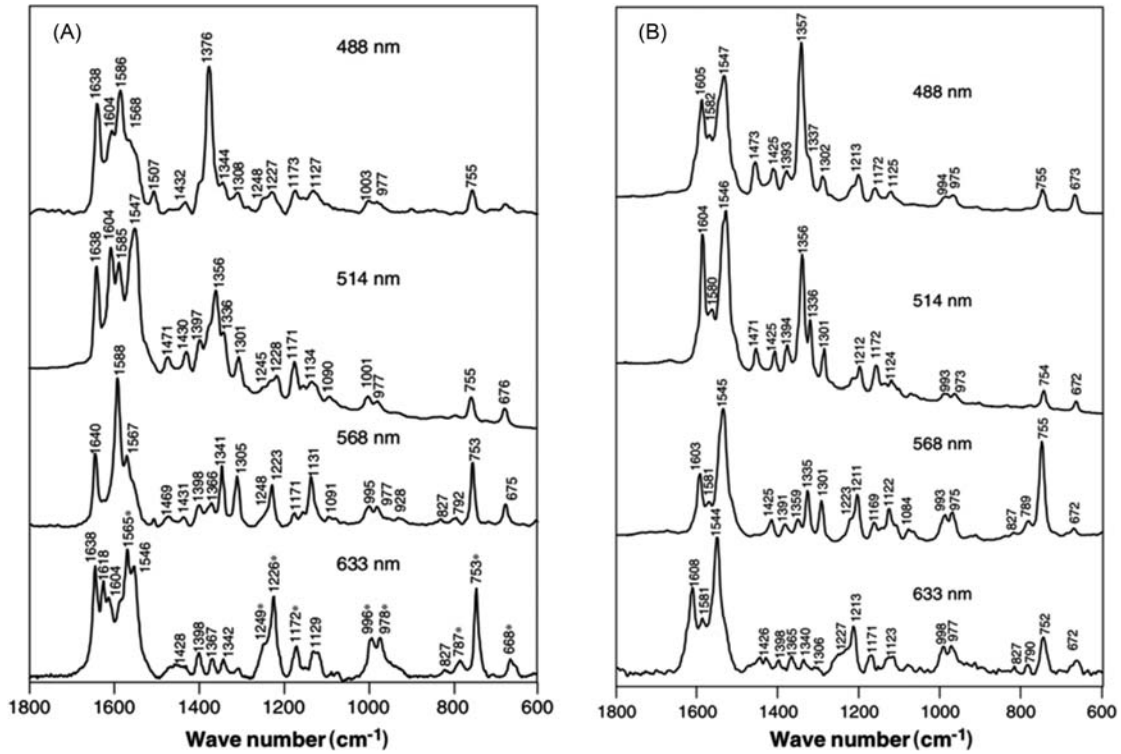


Figure 13.4

(A) Raman spectra of oxygenated single erythrocytes recorded using different excitation wavelengths. Bands marked with asterisks are those which appear relatively enhanced at 632.8 nm compared with other excitation wavelengths. The spectra presented are averaged from four spectra recorded of different RBCs under the same conditions. For each spectrum, 10 intermittent scans were accumulated with a 10 s exposure time for each accumulation. (B) Same as (A) but for cells deoxygenated with sodium dithionite. Source*: The figure was reproduced from B.R. Wood, D. McNaughton, Raman excitation wavelength investigation of single red blood cells in vivo, J. Raman Spectrosc. 33 (2002) 517–523 [48] with permission from Willy & Sons, Inc.*

13.5.1 Raman band assignments of hemoglobin in red blood cells

By treating the hydrogen atoms as point masses, then the 37-atom heme has 71 in-plane modes ($2N-3$), which constitute the following irreducible representation:

$$\Gamma_{\text{in-plane}} = 9A_{1g} + 8A_{2g} + 9B_{1g} + 9B_{2g} + 18_{Eu}$$

Raman spectra of metalloporphyrins can also show out-of-plane modes. There are 34 out-of-plane vibrations, which are

$$G_{out\text{-}of\text{-}plane} = 3A_{1u} + 6A_{2u} + 5B_{1u} + 4B_{2u} + 8E_g.$$

Strictly, for D_{4h} symmetry only the E_g modes are Raman active, but the other modes become weakly allowed for general porphyrin systems. The symbol "ν" is given to in-plane stretching modes, "δ" for in-plane bending modes, and the "γ" symbol is used to designate out-of-plane modes. The latter are grouped into various categories including out-of-plane wagging modes, tilting, and internal folding of the pyrrole rings.

Detailed band assignments for Hb inside RBCs are found in Table 13.1. The spectrum of both oxyRBCs and deoxyRBCs when exciting with the 488-nm laser line shows a very strong band at 1376 cm^{-1}, which is assigned to the pyrrole half ring symmetric stretching vibration (ν_4). Interestingly, in the case of Hb inside RBCs, the position of ν_4 is dependent on the excitation wavelength. Fig. 13.5 shows the wave number shift versus excitation wavelength for oxyHb and deoxyHb inside RBCs. It can be seen that as the excitation wavelength gets longer, the position of ν_4 for oxyHb and deoxyHb becomes closer, approaching a value of 1365 cm^{-1} in both states, whereas at 488 nm they are almost 20 cm^{-1} apart. These results provide further proof that ν_4 is not a definitive indicator of Fe oxidation state as originally proposed by Yamamoto et al. [12] but rather is highly dependent on the excitation wavelength and not charge density factors due to oxidation state. This is in agreement with Spaulding et al.'s [14] observations that they did not observe a relationship between charge density of the central metal ion and the position of ν_4 for a variety of metalloporphyrins.

The intensity of ν_4 and other symmetric modes of vibration are particularly strong in resonance or preresonance excitation into the strong Soret electronic transition at approximately 400 nm and is consistent with Raman spectrum of isolated Hb. The bands between 1650 and 1500 cm^{-1} are from porphyrin C−C and C = C stretching vibrations. The general ferrous high-spin hemes, like deoxyHb, show the characteristic three-band profile as shown in Fig. 13.4B. The ferric low-spin hemes, such as oxyHb and carboxyHb, show more bands (four to five bands) in this region, possibly due to the extra symmetry associated with the state where the Fe^{3+} ion moves closer to the porphyrin plane upon reacting with an additional molecule (in case of oxyHb with oxygen, Fig. 13.2B). Bands between 1300 and 1200 cm^{-1} show significant differences between oxygenated and the

Table 13.1: Observed wavenumber (cm^{-1}), assignments and local coordinates for hemoglobin.

Oxy 488 nm	Deoxy 488 nm	Oxy 514 nm	Deoxy 514 nm	Oxy 568 nm	Deoxy 568 nm	Oxy 632.8 nm	Deoxy 632.8 nm	Assignment[a]	Local coordinate[b]
1638	Absent	1638	Absent	1640	Absent	1638	Absent	ν_{10}	$\nu(C_aC_m)_{asym}$
Absent	Absent	1627[c]	Absent	Absent	Absent	1618	Absent	$\nu_{(C=C)}$	$\nu(C_a{=}C_b)$
1604	1605	1604	1604	Absent	1603	1604	1608	ν_{19}	$\nu(C_aC_m)_{asym}$
1586	1582	1585	1580	1588	1581	1581[c]	1585	ν_{37}	$\nu(C_aC_m)_{asym}$
1568	Absent	1557[c]	Absent	1567	Absent	1565	Absent	ν_2	$\nu(C_{\beta\beta})$
Absent	1547	1547	1546		1545	1546	1544	ν_{11}	$\nu(C_{\beta\beta})$
1507	Absent	Absent	Absent	1505	Absent	Absent	Absent	$2\nu_{15}$	ν(pyr breathing)
	1473	1471	1471	1469	Absent	Absent	Absent	$-CH_2$ (scissor)	$-CH_2$ (scissor)
1432	1425	1430	1425	1431	1425	1428	1426	ν_{28}	$\nu(C_aC_m)_{sym}$
1397	1394	1397	1394	1398	1391	1398	1398	ν_{20}	ν(pyr quarter-ring)
1376	1357	1371 (m) 1356 (s)	1356	1366	1359	1367	1365	ν_4[d]	ν(pyr half-ring)$_{sym}$
1344	1337	1336	1336	1341	1335	1342	1340	ν_{41}	ν(pyr half-ring)$_{sym}$
1308	1302	1301	1301	1305	1301	1306	1306	ν_{21}	$\delta(C_mH)$
1248	Absent	1245	Absent	1248	Absent	1249	Absent	ν_{13}	
1227	1221	1228	1220	1223	1223	1226	1223	ν_{13} or ν_{42}	$\delta(C_mH)$
Absent	1213	Absent	1212	Absent	1211	Absent	1213	$\nu_5 + \nu_{18}$	$\delta(C_mH)$
1173	1172	1171	1172	1171	1169	1172	1171	ν_{30}	ν(pyr half-ring)$_{asym}$
1127	1124	1134	1124	1131	1122	1129	1123	ν_{22}	ν(pyr half-ring)$_{asym}$
Absent	Absent	1090	1082	1091	1084	1090	1084	ν_{23}	$\nu(C_\beta C_1)_{asym}$
1003	994	1001	993	995	998	996	996	ν_{47}	$\nu(C_\beta C_1)_{asym}$
977	975	972	973	977	977	978	972	ν_{46}	δ(pyr deform)$_{asym}$ and/or γ $(=C_bH_2)_{sym}$
Absent	Absent	Absent	Absent	827	827	827	827	γ_{10}	$\gamma(C_mH)$
Absent	Absent	Absent	Absent	792	790	787	790	ν_6	ν(pyr breathing)
755	755	755	754	753	752	753	752	ν_{15}	ν(pyr breathing)
675	673	676	672	675	672	668	672	ν_7	δ(pyr deform)$_{sym}$

[a]Assignments are based mainly on the labeling scheme originally devised by Abe et al. [49] for octaethylporphyrinato-Ni(II).
[b]Local coordinates based mainly on studies by Hu et al. [50].
[c]Only observed after calculating second derivative.
[d]Position influenced by excitonic enhancement.

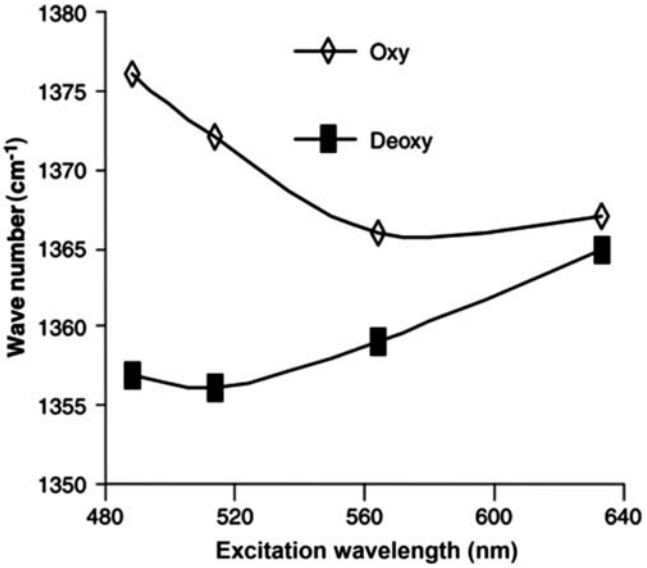


Figure 13.5
Plot depicting the oxyHb and deoxyHb wave number for ν_4 mode versus excitation wavelength applied. Source: *The figure was reproduced from B.R. Wood, D. McNaughton, Raman excitation wavelength investigation of single red blood cells in vivo, J. Raman Spectrosc. 33 (2002) 517–523 [48] with permission from John Wiley & Sons, Inc.*

deoxygenated state. This is most notably observed in the spectra obtained with the 632.8 nm excitation, where the bands at 1249 and 1226 cm^{-1} in the oxygenated RBCs shift to 1227 and 1213 cm^{-1} in the deoxygenated cells. Second derivative analysis of oxyRBCs revealed the region between 1250 and 1200 cm^{-1} to comprise three bands at approximately 1249, 1224, and 1215 cm^{-1}. The 1224 and 1215 cm^{-1} are assigned to methyne deformation modes from the porphyrin macrocycle. The shift in these modes may be associated with the expanding and contracting porphyrin macrocycle. It is tempting to assign the band at 1249 cm^{-1} to amide III mode which is often quite intense in the Raman spectra of proteins. However, this band was found to be particularly sensitive to temperature and laser exposure and was hypothesized to be the result of heme aggregation following denaturation of protein [39]. Given there is no other evidence of protein bands in the spectra obtained with 632.8 nm, then the 1249 cm^{-1} band is most likely originated from the porphyrin macrocycle. It is interesting to note that at the 632.8-nm excitation a number of bands appear enhanced compared to other excitation wavelengths, which are indicated by the asterisks in Fig. 13.4A. In UV—vis spectra of oxyHb there is no evidence of an optical charge transfer band between 600 and 700 nm yet the Raman bands appear enhanced at 632.8-nm excitation [51]. To further investigate this unusual enhancement we recorded the spectra of dehydrated oxyRBCs at 10-nm intervals between 573 and 633 nm using an

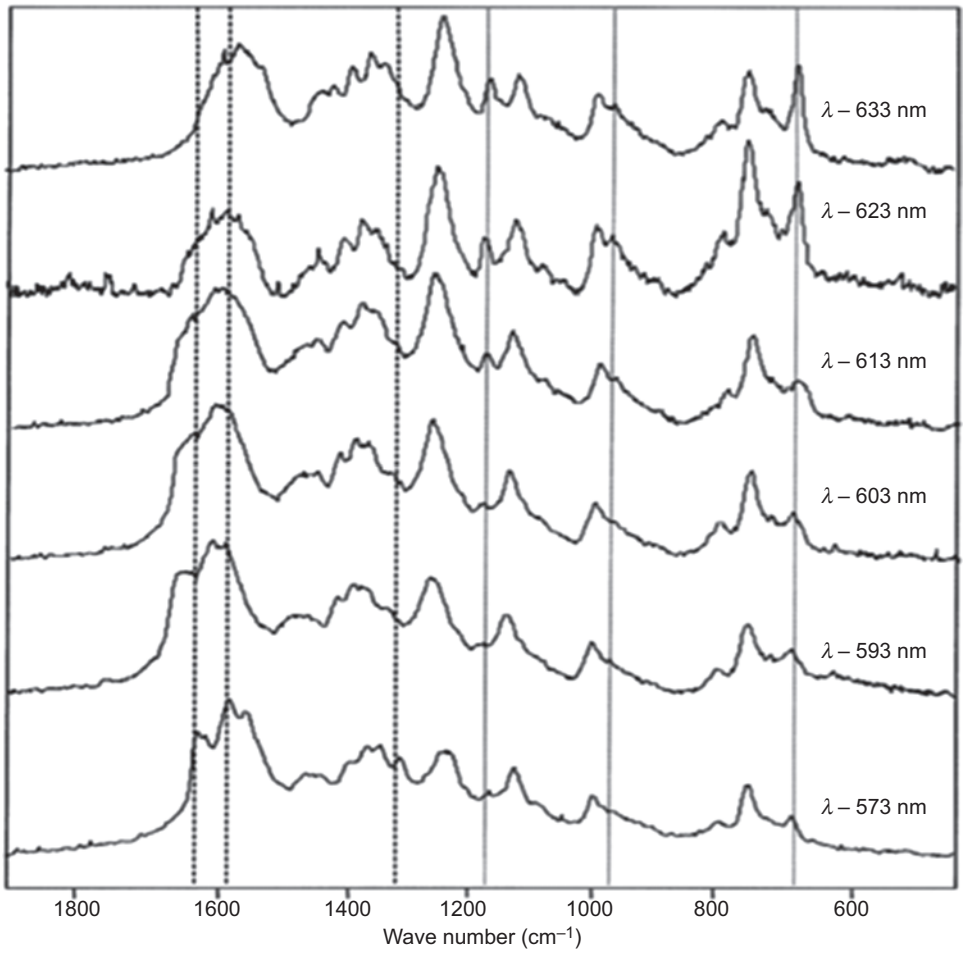


Figure 13.6

Spectra recorded at various excitation wavelengths of dried oxyRBCs in the 573-633 nm region using a tunable all-line argon-ion laser to pump a rhodamine 6G dye laser. Each spectrum is made up of eight spectral windows and the laser exposure time for each window is 180 s. Between each measurement the laser is calibrated to the 520.5 cm^{-1} silicon band and the power adjusted to approximately 1 mW. Dashed vertical lines highlight the bands that become enhanced approaching the Q_0 electronic transition, while dotted vertical lines indicate bands that appear resonance-enhanced at approximately 633 nm. Source*: The figure was reproduced from B.R. Wood, B. Tait, D. McNaughton, Micro-Raman characterisation of the R to T state transition of haemoglobin within a single living erythrocyte, Biochim. Biophys. Acta (BBA) Mol. Cell Res. 1539(1) (2001) 58–70 [51] with permission from Elsevier.*

argon-ion laser to pump a 6G rhodamine dye laser (Fig. 13.6) [51]. Between each measurement, the system was calibrated for intensity using the 520.5 nm band from a Si wafer. A number of bands appear enhanced at 633 nm compared to the other excitation

wavelengths including those at 1245 cm^{-1} (ν_{13}–B_{1g}), 1172 cm^{-1} (ν_{30}–B_{2g}), 976 cm^{-1} (ν_{46}–B_{1g}), and 672 cm^{-1} (ν_7–A_{1g}) and are clearly indicative of resonance enhancement showing a mixture of totally symmetric and nontotally symmetric modes [51]. The enhancement of bands at 632.8 nm, where there is no observable optical electronic transition, is greater than that observed in preresonance with the Q_0 band at $\sim$575 nm. Spectra of metHb erythrocytes, where there is a distinct charge transfer band at 633 nm, gave a similar enhancement profile to oxyHb indicating that the enhancement observed is independent of excitation into a charge transfer transition [51]. It is also important to note that the enhancement only occurs in cells and crystal deposits of Hb and not from solutions of Hb. This indicates that the enhancement observed is the result of the high concentration of heme in RBCs/crystals resulting from long-range excitonic interactions generated from the superposition of electronic transitions in the highly concentrated heme environments [34,39,44,48,52].

In another set of experiments we reported that successive additions of sodium dithionite to a suspension of RBCs increased the capacity of the cells to uptake oxygen [53]. This effect was not observed when the RBCs were exposed to gaseous N_2, which instead induced short episodes of hypoxia. The results indicate that successive additions of sodium dithionite to a suspension of RBCs results in an increase in the rate and capacity of the RBCs to uptake oxygen, while the dithionite suspension did not cause Hb from lysed RBCs to uptake more oxygen after successive additions. Hence the effect was only observed in functional intact RBCs. We hypothesized that the sodium dithionite increases the disorder of Hb in the RBC, thus facilitating oxygen diffusion [53].

13.5.2 Polarized resonance Raman spectra reveal highly ordered heme groups in red blood cell

To further investigate the hypothesis that the unusual enhancement observed in RBCs and crystallized deposits compared to aqueous Hb is the result of long-range interactions in the highly concentrated environment, we performed Raman polarization spectroscopy which revealed intrinsic ordering of hemes in RBCs [39]. To prove the intrinsic ordering of hemes in a single RBC we examined functional erythrocytes orientated horizontally and vertically using 632.8 nm excitation (Fig. 13.7A). The experimental setup entailed using an aluminum-coated Petri dish that was further coated with poly-L-lysine to affix the cells (Fig. 13.7B). Parallel and perpendicular spectra were recorded using a combination of a polarizer and a half-wave plate [39].

Fig. 13.8 shows the Raman spectra recorded of horizontal (8A) and vertical orientated (8B) RBCs using 632.8-nm excitation and normalized to the depolarized at 996 cm^{-1} band and baseline corrected. Totally symmetric vibrations have depolarization values of $p < 0.75$, while nontotally symmetric vibrations, which include depolarized bands and anonymously

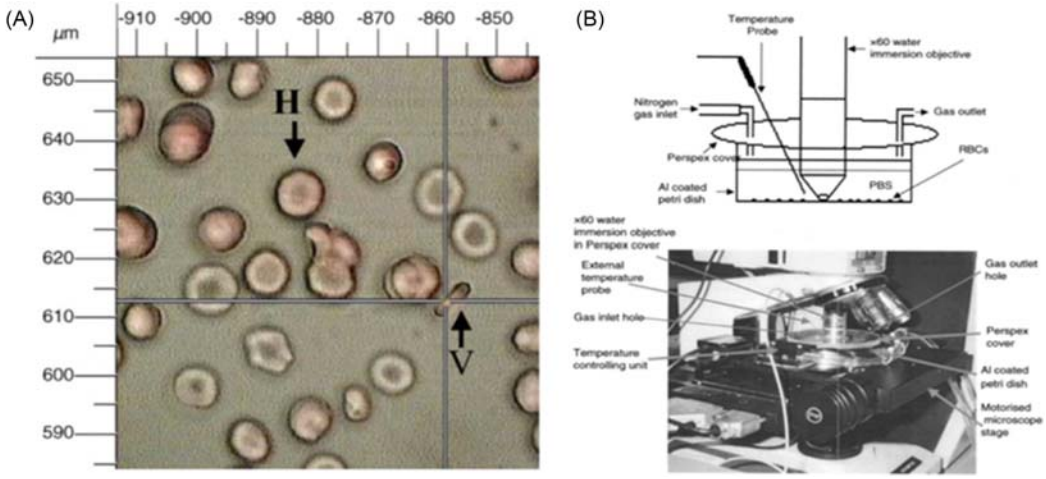


Figure 13.7

(A) Photomicrograph of RBCs affixed in both the vertical and horizontal orientations relative to the Petri dish surface. (B) Experimental setup for recording Raman spectra of functional RBCs. Source*: The figure was reproduced from B.R. Wood, et al., Raman microspectroscopy and imaging provides insights into heme aggregation and denaturation within human erythrocytes, J. Biomed. Opt. 10(1) (2005) 14005 [39] with permission from SPIE.*

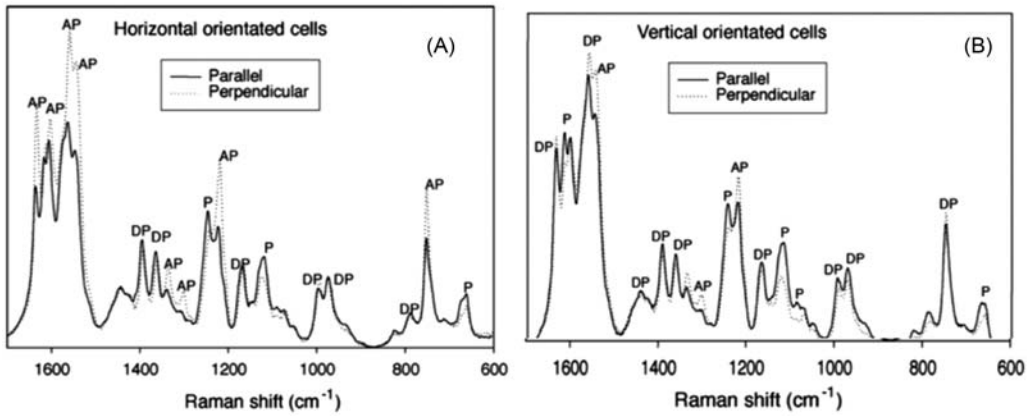


Figure 13.8

(A) Averaged spectra of the parallel and perpendicular scattered light component recorded from 30 oxygenated RBCs orientated horizontally. (B) The corresponding averaged spectra for vertically orientated RBCs. Source*: The figure was reproduced from B.R. Wood, et al., Raman microspectroscopy and imaging provides insights into heme aggregation and denaturation within human erythrocytes, J. Biomed. Opt. 10(1) (2005) 14005 [39] with permission from Elsevier.*

polarized bands have values of $p = 0.75$ and $p \gg 0.75$, respectively. The normalization process results in our values being closer to $p \ll 1.0$, $p = 1.0$, and $p \gg 1.0$ for the three types of depolarization ratio. It should be noted that depolarization ratios in hemes are approximate due to many overlapping bands and anomalous dispersion. The averaged spectra depicted in Fig. 13.5A are dominated by a mixture of totally symmetric and nontotally symmetric modes typical of Type A and Type B scattering. Vertically orientated cells exhibit four totally symmetric polarized bands at 1616, 1244, 1121, and 666 cm^{-1}. Horizontally orientated cells produce bands at 1121 and 662 cm^{-1}. RBCs orientated horizontally exhibit more anomalously polarized than bands compared to spectra of the vertical cells. These additional bands include those observed at 1636, 1607, 1561, 1546, 1221, and 752 cm^{-1}. The intrinsic differences in polarization between vertical and horizontal orientated cells can only be explained by a high degree of heme order in the RBCs. The concentration of heme in the RBC makes it behave like a liquid crystal with Raman scattering properties more similar to single-crystal spectra than to aqueous solutions.

13.5.3 Hemoglobin ligand modes in red blood cells revealed with near-infrared excitation

In general, the (Fe−O−O) bend and the ν(Fe−O_2) stretching vibrations from Hb are very weak and difficult to detect in resonance Raman spectra of RBCs. In fact both the (Fe−O−O) bend and the ν(Fe−O_2) were never observed together in isolated Hb and were only observed as a pair in oxy(phthalocyanato)iron(II) [6]. This is possibly attributed to the extreme resonant enhancement of the porphyrin modes. However, using a highly sensitive near-infrared River Diagnostics Model 3510 Skin Composition Analyzer system with a confocal Raman microspectrometer that is fully optimized to perform rapid *in vivo* measurements on the human skin it was possible to detect the important ligand modes [7]. With this instrumental configuration and using the experimental setup shown in Fig. 13.9A, consisting of a squeeze bottle container filled with liquid N_2, it was possible to deoxygenate and reoxygenate the cells. The resulting spectra (Fig. 13.9B) show the evolution of both the δ(Fe−O−O) bend at 419 cm^{-1} and the ν(Fe−O_2) at 567 cm^{-1} as the cell goes from the deoxygenated to oxygenated state after exposure with atmospheric oxygen. This was the first Raman study to report both ligand modes in oxyHb and the first study to definitely assign the (Fe−O−O) band in Hb.

13.5.4 Enhanced overtone and combination modes observed in hemes using 514-nm excitation

Recently we discovered an entire new series of hyperenhanced overtone and combination modes in the Raman spectrum of Hb [34]. Mostly the first overtone and combination bands are very weak bands and they appear at approximately twice the wave number value of the

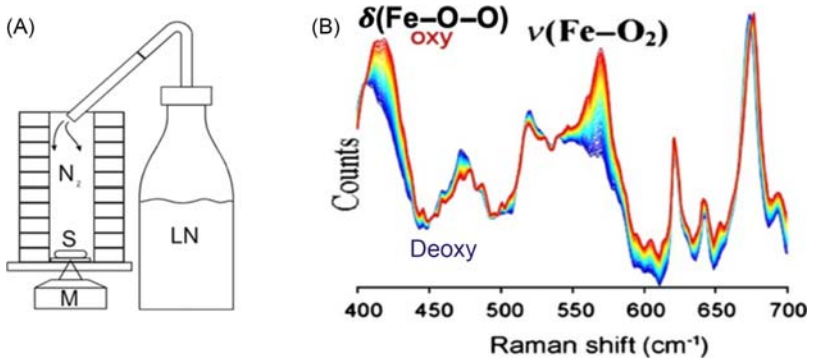


Figure 13.9

(A) Apparatus used in oxygenation/deoxygenation experiments showing a squeeze bottle filled with liquid nitrogen. LN, liquid nitrogen; N_2, nitrogen gas; M, microscope objective; S, blood sample. (B) A series of Raman spectra recorded in the low wave number region (400−700 cm^{-1}) using 785-nm excitation going from the deoxygenated state to the oxygenated state. Source: *The figure was reproduced from B.R. Wood, et al., Resonance Raman spectroscopy of red blood cells using near-infrared laser excitation, Anal. Bioanal. Chem. 387 (5) (2007) 1691−1703 [7] with permission from Springer Nature.*

corresponding fundamental bands. However, due to high intracellular Hb concentration, these bands observed in the RRS spectra obtained from RBCs are more intense than the fundamentals especially when 514-nm excitation and collection of the parallel polarized light components were applied. Fig. 13.10 compares spectra of oxygenated and deoxygenated cells recorded with 413.1, 514.5, and 632.8 nm excitation wavelengths. The spectra recorded at 514.5 nm show extraordinary enhancement of bands between 3600 and 2300 cm^{-1} within the combination and overtone region for heme fundamentals. The observation of intense overtones and combination bands has previously been reported by Aramaki et al. [54] for copper tetraphenyl porphyrin. Excitation into the vibronic Q_v band at $\sim$525 nm can result in a Type C Raman scattering mechanism, which occurs between forbidden electronic transitions that are prohibited at the equilibrium geometry of the molecule [55]. The observed enhancement in RBCs is even greater than that observed in single crystals of Hb [34]. The highly organized and aligned intracellular Hb dominates the optical properties of the RBCs enabling these interactions to occur. It was shown that the intensity of the overtone bands enabled the recording of high contrast Raman images [34,56]. An assignment of observed overtone and combination bands within the RBCs spectrum was possible due to systematic polarization measurements and symmetry arguments. As expected, the overtone bands were found to be sensitive to ligand exchange and imitated the fundamental spectrum in response to changes in the oxidation and spin state of the iron ion. Such bands in the nonfundamental region provide additional spectral

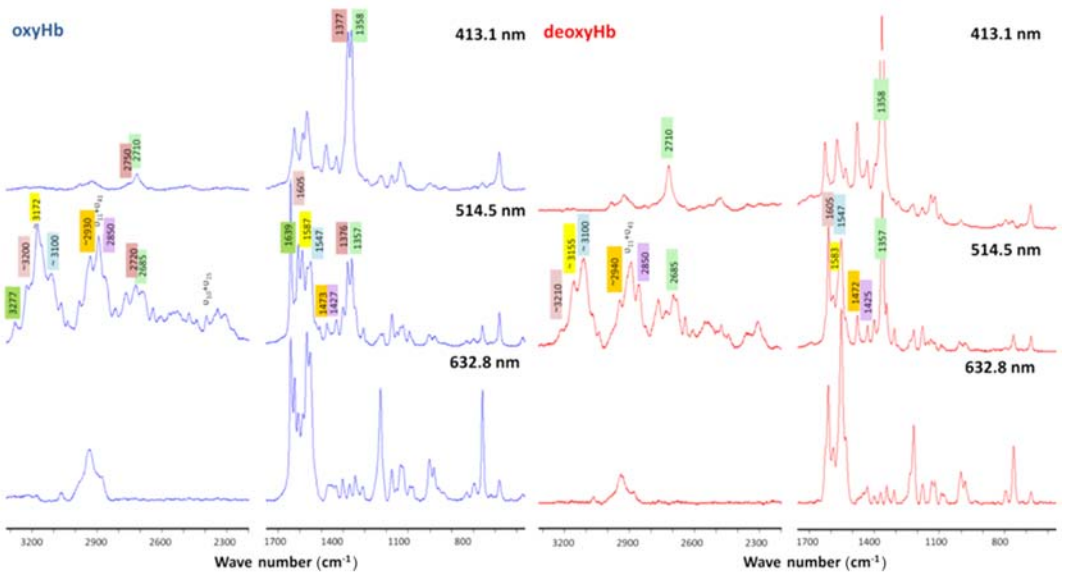


Figure 13.10

Comparison between average resonance Raman spectra collected from 10 different oxygenated and deoxygenated RBCs obtained using 413.1, 514.5, and 632.8 nm laser wavelengths in the spectral regions of 3300−2200 and 1700−600 cm^{-1}. The color assignments of the fundamental bands have corresponding colors for their respective overtones. The intensity of the bands in the 3300−2200 cm^{-1} have not been scaled or normalized and appear on the same scale as the fundamental region. Source: *This figure was reproduced from K.M. Marzec, et al., Red blood cells polarize green laser light revealing hemoglobin's enhanced non-fundamental Raman modes, Chemphyschem 15 (18) (2014) 3963−3968 [34] with permission from Willy & Sons, Inc.*

information, which improves the potential diagnostic capability of the Raman technique for various hemopathies. This is due to the more detailed molecular fingerprint of the heme structure that can be used to discern Hb from hemozoin in malaria-infected RBCs [34,56].

13.6 Applications of Raman spectroscopy in red blood cell research

13.6.1 Resonance Raman spectroscopy in malaria research

Malaria remains one of most significant threats to global health worldwide, with almost 40% of the population at risk [42,52,57,58]. It affects approximately 500 million people each year, leading to death in 1−3 million cases [42,57−59]. Its prevalence is highest in Africa and most of the lethal cases of malaria occur in children under the age of 5 [42,57,58]. These dramatic statistics highlight the importance of malaria research and the urgent need for rapid diagnostic tools.

Malaria infection occurs when the parasite *Plasmodium* sp. is injected into bloodstream by female *Anopheles* mosquitoes [57,58]. The sporozoites introduced through a mosquito bite quickly invade liver cells. Over the next 14 days, the liver stage parasites multiply, resulting in thousands of merozoites. Merozoites invade erythrocytes and within the next 48 hours (intraerythrocytic cycle) undergo several morphological changes, related to various developmental stages. These stages include rings, trophozoites, and schizonts. The presence of malaria parasites significantly changes the chemical composition of the erythrocytes. In particular, the trophozoites catabolize large amounts of Hb, releasing ferrous protoporphyrin IX (Fe(II)PPIX) and denatured globin. Free Fe(II) is thought to be toxic to parasites resulting in lipid oxidation and membrane disruption [42]. To avoid the toxic heme buildup, the parasites have developed a specific defense mechanism, which involves oxidation of the protoporphyrin to form a dimeric heme complex consisting of two Fe(III)PPIX units linked via a propionate linkage, which forms an extended array of hydrogen-bonded dimers known as hemozoin—or malaria pigment [42,52,57,58,60].

The current diagnostic methods for malaria parasites involve microscopy, rapid diagnostic tests (RDTs), and molecular detection methods such as polymerase chain reaction (PCR) [57,59,60]. Malaria is endemic in the poorer regions of the world, where access to modern laboratories and healthcare providers is limited. Consequently, there is a strong need for an effective diagnostic tool that is simple to use and easily adaptable to field work with high sensitivity to detect asymptomatic carriers of the disease. PCR—although highly efficient in determining the presence of *Plasmodium falciparum*—is not only expensive, but also requires the use of costly equipment and facilities, along with extensively trained personnel. Microscopy and rapid diagnosis tests, on the other hand, are easily adaptable for use in remote locations, but still have significant limitations from the diagnostic perspective [60]. Microscopy requires trained and experienced personnel and more importantly is to a high extent subjective and thus prone to human error [60]. The RDT, on the other hand, suffers from rather low sensitivity [59].

Vibrational spectroscopy is extremely promising in this diagnostic landscape because it is rapid, minimally invasive, and requires minimal sample preparation. Currently, there are several attempts to adapt both IR and Raman to rapid diagnosis of malaria. Several broad reviews of this aspect can be found in the literature [42,46] with the most recent by Perez-Guaita et al. [46]. Here, we will focus on the use of Raman spectroscopy specifically and highlight several important findings in the context of malaria research.

Out of all parasitic diseases studied by means of Raman spectroscopy, malaria has definitely been explored most comprehensively. An advantage of Raman spectroscopy for malaria studies, especially those involving the late erythrocytic stages, lies in the strong resonant enhancement of hemozoin, which is spectroscopically identical to its synthetic analogue β-hematin [61], and Hb. These highly symmetric and chromophoric compounds

produce strong scattering, enhanced by several magnitudes of order, compared to the spontaneous Raman effect. This enhancement together with the substantial differences between the spectral profiles of malaria pigment and Hb (including large, unmistakable band shifts) makes the Raman technique extremely promising for malaria diagnosis.

The first Raman studies on malaria involved spectral characterization of hemozoin [62] and murine erythrocytes infected by *Plasmodium berghei* [63]. Subsequently, Wood et al. [58] demonstrated the first application of Raman spectroscopy for *P. falciparum* and provided a broad study of the relationship between spectral profile of hemozoin and excitation wavelength. Spectra of hemozoin were recorded directly from the food vacuole of infected human red blood cells (iRBCs) [58] demonstrating the potential of this technique for diagnostic applications. The mechanism of enhancement of hemozoin signals was later explored by the same group [52]. The authors showed dramatic enhancement of symmetric modes (A_{1g}) observed when using 785- and 830-nm excitations and hypothesized that the enhancement was due to excitonic long-range effects resulting from the extended array of dimeric hemes. This effect is particularly important for diagnostics, because Hb scatters poorly at these longer excitation wavelengths compared to hemozoin [52,58].

In another study, Wood et al. [57] discovered a partial dark-field effect, related to the presence of parasites within erythrocytes. This effect is illustrated in Fig. 13.11, and becomes a powerful diagnostic for malaria when the effect is combined with Raman imaging and unsupervised hierarchical cluster analysis (UHCA) to unequivocally detect late ring stage and trophozoites in thick and thin blood smears. The hemozoin crystals "light up" when in the dark field. In dark-field microscopy, the illumination rays do not enter the microscope aperture, leading to dark (black) background and providing good contrast. The objects appear bright due to the collection of scattered light. The more efficient the light scattering of the material (usually particles), the brighter the object appears. In this case, the hemozoin crystals scatter substantially more compared to other erythrocyte components. Hence, their presence can be easily visualized [57].

Raman imaging combined with UHCA performed on the collected map (Fig. 13.11B–E) clearly confirmed that the bright objects are indeed hemozoin crystals (Fig. 13.11E). The partial dark-field effect is a great contribution to the potential application of Raman imaging of erythrocytes as a diagnostic tool. An imaging-based approach, although undoubtedly allows to identify the presence of a parasite, can be significantly more time-consuming, compared to an approach based on collection of a single spectrum. The latter, however, requires a selective choice of spot for data collection. This entails a risk of choosing a "wrong" spot (not containing a parasite, even though it is present within an erythrocyte), particularly when using high-magnification microscopes. However, the combination of the partial dark-field effect with Raman spectroscopy minimizes that risk,

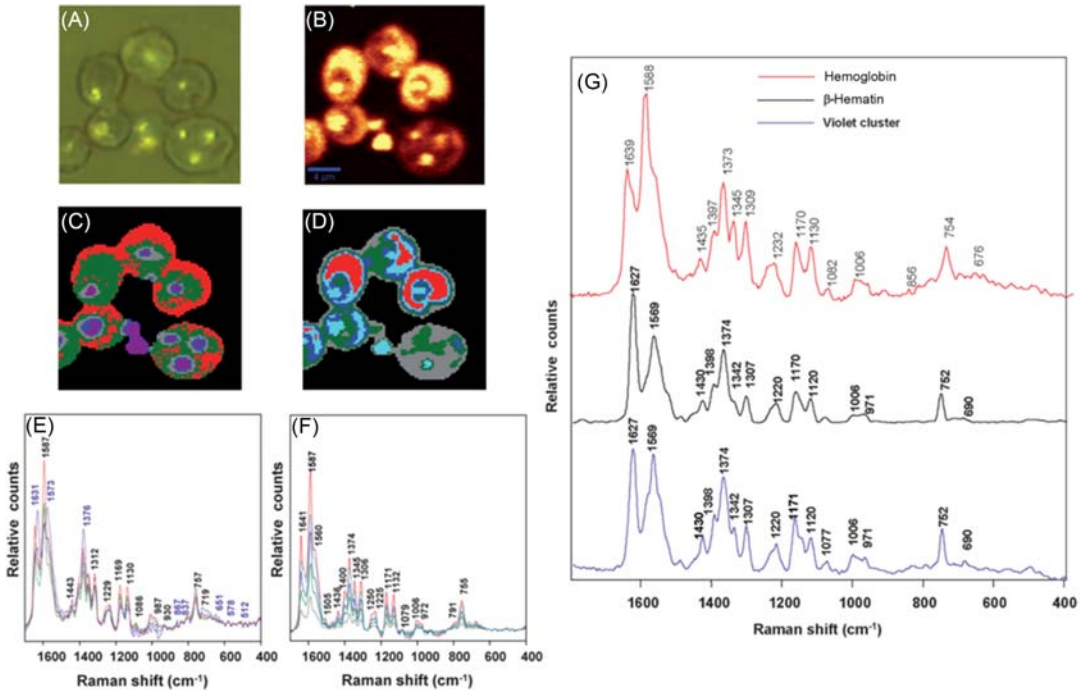


Figure 13.11

Partial dark-field effect in RBCs infected with *Plasmodium falciparum*. (A) Visible microphotograph of infected human red blood cells (iRBCs), with brighter spots corresponding to hemozoin crystals. (B) A chemical map, showing the integral intensity of Raman signal in the range of 1680–1620 cm^{-1}. Brighter color corresponds to higher value. (C) Results of unsupervised hierarchical cluster analysis (UHCA) performed in the spectral range 1700–1300 cm^{-1} with *D*-values. (D) Euclidean distance algorithms. (E) A direct comparison of Raman spectra of the violet cluster in 1C (hemozoin) with Raman spectra of standards: Hb and β-hematin (synthetic analogue of hemozoin). Source*: The figure was reproduced from B.R. Wood, et al., Resonance Raman microscopy in combination with partial dark-field microscopy lights up a new path in malaria diagnostics, Analyst 134 (6) (2009) 1119–1125 [57] with permission of Royal Society of Chemistry.*

while preserving the ability to collect data within seconds (single spots) and confirm the presence/absence of hemozoin (e.g., using the band at 1569 cm^{-1}).

In a novel approach Raman acoustic levitation spectroscopy (RALS) in combination with a portable fiber optic Raman system was applied to investigate oxygenated and deoxygenated RBCs and to detect malaria parasites from RBCs lysed in water [64]. The cells and lysate were levitated using an ultrasonic levitator (Dantec/Invent Measurements Technology, Erlangen, Germany, ultrasonic frequency = 58 kHz, l = 5.9 mm). The experimental setup is shown Fig. 13.12A while the levitated cells are shown in Fig. 13.12B. The resulting spectra

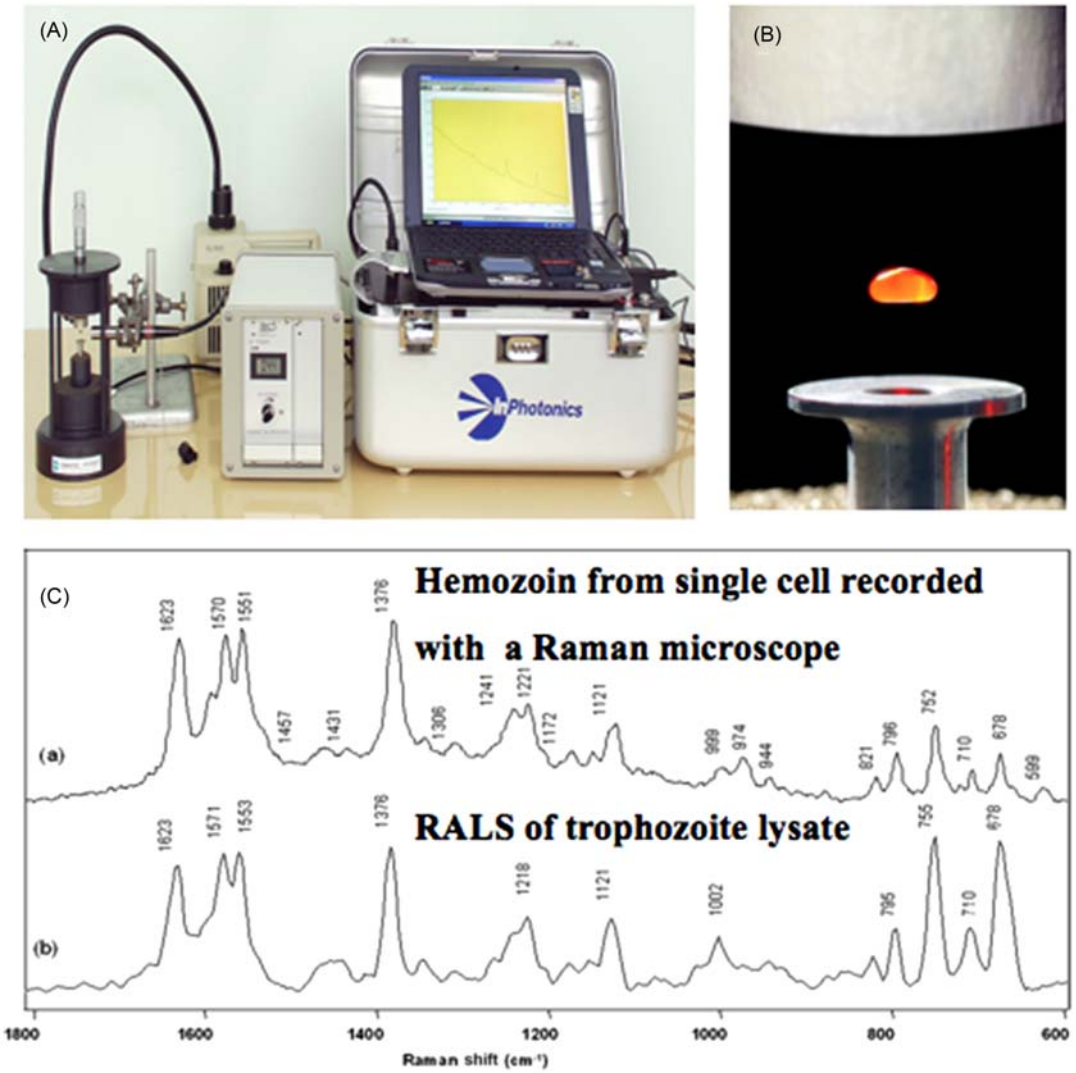


Figure 13.12

(A) Photograph of the experimental set-up showing the portable Raman spectrometer with fiber optic along with the acoustic levitation device with a suspension of cells in the central acoustic node. (B)lose up photograph of a suspension levitating red blood cells (RBCs). (C) (a) Spectrum recorded of hemozoin inside a single RBC using a Raman microscope and 785-nm excitation. (C) (b) RALS spectrum recorded of a 10 μL suspension of trophozoite lysate at 6% parasitemia showing characteristic peaks of hemozoin. Source*: This figure was modified from L. Puskar, et al., Raman acoustic levitation spectroscopy of red blood cells and Plasmodium falciparum trophozoites, Lab. Chip. 7 (9) (2007) 1125–1131 [64] with permission from the Royal Society of Chemistry.*

are shown in Fig. 13.12C where 13.12C(a) shows a spectrum of hemozoin recorded using Renishaw microscope with 785 nm excitation while Fig. 13.12C(b) shows the spectrum recorded of the levitated lysate. It is clear that the RALS approach produces a high-quality spectrum of the hemozoin that closely matches the spectrum of hemozoin recorded using the microscope configuration. The advantage of using this approach is that no substrate is required and the spectra are recorded in a containerless environment and consequently are devoid of surface-induced effects. The disadvantage of the approach is that the red cell droplets and lysates can become unstable and explode which represents a major safety hazard to the operator and strict care must be taken to avoid blood splatter.

Besides resonant Raman spectroscopy, several malaria studies utilized surface-enhanced Raman spectroscopy (SERS) [59,65–67]. For instance, Chen et al. [59] applied SERS to detect the presence of malaria parasites, however, not through the detection of hemozoin, but rather membrane-related changes in RBCs. Although hemozoin is spectroscopically distinct from other cellular components, its content in parasites at the ring stage is very low. It is well established that the host cell plasma membrane undergoes several changes during the intracellular development of *P. falciparum*, including a decrease in cholesterol and sphingomyelin and an increase in the ratio of phospholipid/cholesterol [68]. In addition, parasites modify the membrane composition via secretion of proteins [68,69]. The authors demonstrated the ability to differentiate between parasite stages using SERS spectra, with the observed changes consistent with the expected evolution of RBCs plasma membrane and membrane proteins expressed during the *Plasmodium* cycle. Although the results indicated a lack of discrimination between iRBCs and RBCs within the first 8 hours of the intracell cycle, prominent changes were visible after 16 hours, with the limit of detection after 40 hours reaching 5×10^7/mL [59].

SERS studies focused on hemozoin utilized the surface enhancement (characteristic for SERS) with resonant enhancement (characteristic for RRS). This SERRS technique was developed by Chen et al. [66], using silver nanoparticles synthesized "conventionally" (outside of the studied object/cell) and within the *Plasmodium* parasite, to achieve closer contact. For the later approach, the authors demonstrated an exceptionally high sensitivity of 0.00005% parasitemia level in the ring stage.

Yuen and Liu [67] further combined the SERRS approach with magnetic field enrichment. The SERS-active nanoparticles, containing iron oxide core and silver shell, in addition to surface enhancement of Raman signal provided also magnetic enrichment for β-hematin. This enabled a limit of detection of 5×10^{-9} M for β-hematin (dissolved in aqueous NaOH, with NAOH concentrations ranging from 10^{-4} to 10^{-11} M), which in theory is equivalent to 30 parasites per μL (assuming 0.22 pg/cell as the concentration of hemozoin at ring stage). However, the approach was never tested on iRBCs.

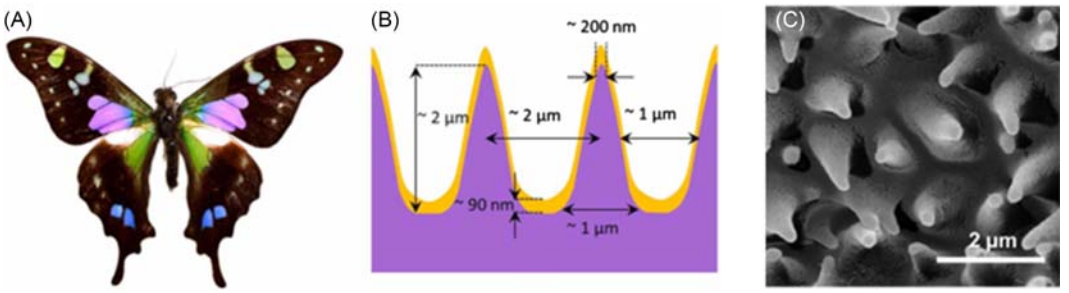


Figure 13.13

Graphium weiskei butterfly wings as SERS substrates. (A) *G. weiskei* butterfly sample. (B) Schematic cross-sectional view of a gold-coated wing. Typical chitinous conical protrusion dimensions and spacing are based on visual investigation of SEM images. (C) SEM image of chitinous nano-structured conical arrays found on the wings of the *G. weiskei* butterfly. Source: *Reproduced from Garrett, N.L., et al., Bio-sensing with butterfly wings: naturally occurring nano-structures for SERS-based malaria parasite detection, Phys. Chem. Chem. Phys. 17 (33) (2015) 21164–21168 [70] with permission from the Royal Society of Chemistry.*

Garrett et al. [70] utilized the SERRS properties of a *Graphium weiskei* butterfly wing to detect hemozoin in malaria-infected RBC suspensions (Fig. 13.13) using 785-nm excitation. The resulting spectra showed resonant enhancement of the totally symmetric mode ν_4 compared to conventional Raman at the same wavelength (Fig. 13.14). The enhancement enabled a detection limit of 0.005% parasitemia, which is within clinical symptomatic range.

Although most Raman-based malarial studies utilize RBCs, some researchers explored the potential of plasma as a diagnostic medium [60,71]. Hobro et al. [60] collected Raman data from plasma and whole blood for 7 consecutive days, following *Plasmodium* infection. The authors were able to identify spectral differences between plasma samples within 1 day postinfection. These included changes related to Hb, hematin, and hemozoin content, together with alterations in lipid composition. Significant differences in plasma content of Hb and hemozoin between malaria-infected and healthy patients were also reported by Bilal et al. [71]. The authors included samples from 60 patients with suspected malaria and dengue fever together with 10 healthy patients. A multivariate model based on PLS regression provided 86% accuracy of prediction [71]. These results are very encouraging, considering the possibility of more advanced analytics and increasing the number of samples to improve accuracy.

In addition to the diagnostic studies, Raman spectroscopy has been applied to other aspects of malaria research, such as parasite–drug interactions [56,72] and tissue studies [73] of malaria infections. These provide a valuable insight into the parasite morphology and chemical composition, potential treatment options, and host's responses. For instance, Webster et al. [72] explored the effect of chloroquine (common antimalaria drug) in RBCs

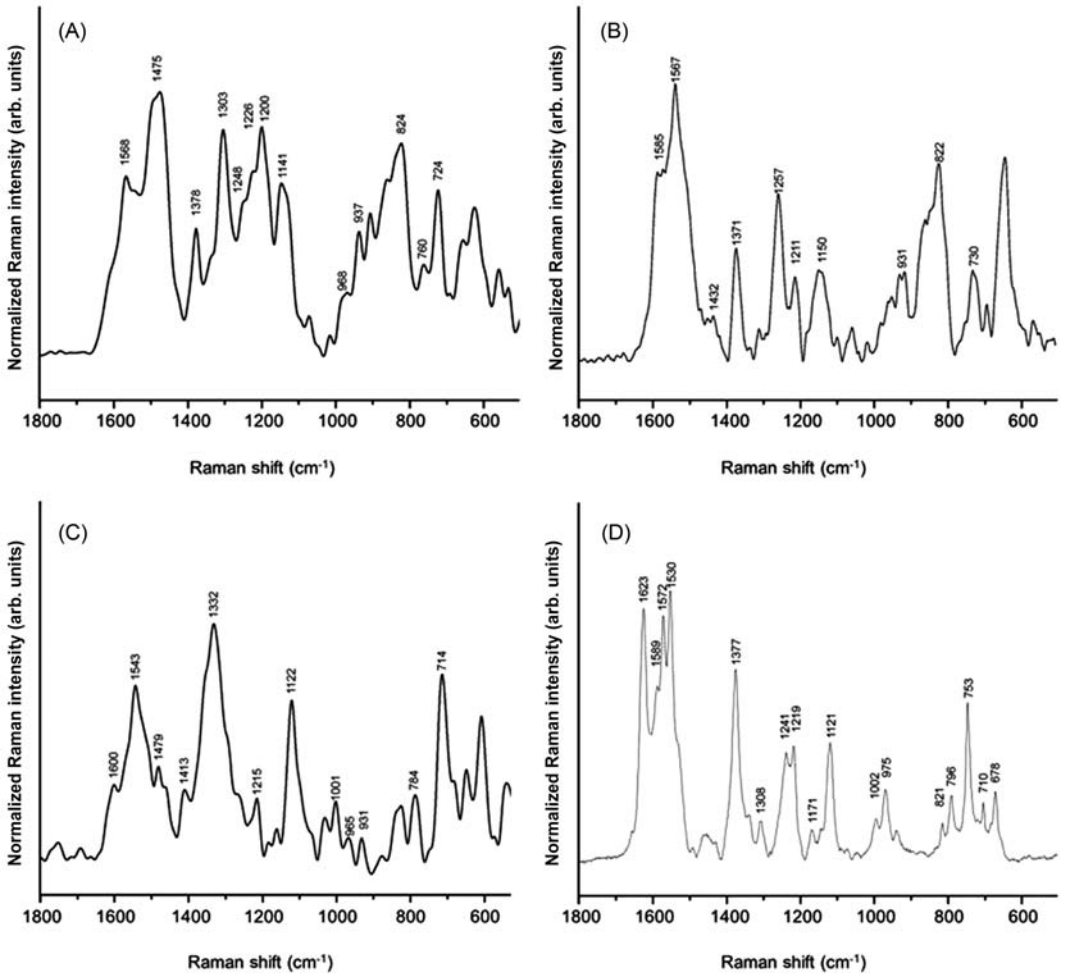


Figure 13.14

SERS spectra of (A) 0.005%, (B) 0.05%, and (C) 0% (control) malarial-infected RBCs lysate, respectively. (D) Conventional Raman spectrum of hemozoin at 785 nm. Source*: Reproduced from Garrett, N.L., et al., Bio-sensing with butterfly wings: naturally occurring nano-structures for SERS-based malaria parasite detection, Phys. Chem. Chem. Phys. 17 (33) (2015) 21164–21168 [70] with permission from the Royal Society of Chemistry.*

infected with *P. falciparum*. They observed a reduction in hemozoin bands when cells were inoculated with chloroquine. Kozicki et al. [56] reported similar changes together with increased levels of oxyHb in the treated erythrocytes, compared to nontreated ones. In addition, multimodal approaches, based on a combination of Raman spectroscopy with other techniques for malaria research are becoming increasingly popular. Examples of this include combining Raman spectroscopy with quantitative phase microscopy [74] and infrared spectroscopy [75].

13.6.2 Application of Raman spectroscopy to blood storage

A second potential application of Raman (and IR) spectroscopy, broadly explored by several research groups, is related to storage of blood and its derivatives. The need for blood storage, particularly RBCs, is an essential aspect of modern medicine.

Current studies suggest that transfusions with older RBCs not only lead to poorer clinical outcomes, but may actually be associated with posttransfusion illnesses, due to the potential for toxicity of RBCs with prolonged storage [76–78].

During storage, even within that short time frame, cells undergo various morphological and physiological changes (known as *storage lesions* [78]), which can gradually degrade their functionality. The term "storage lesion" describes the progressive degradation of RBCs structure and function that occurs during conventional red cell storage. The components of the storage lesion are well described but the clinical relevance of these storage-related changes remains uncertain [79]. The storage lesions may include the changes of the RBCs' structure and function as well as changes in red cell storage medium. The first one may include cellular membrane changes, which affects shape and decreased survivability of the RBC; decrease in 2,3-diphsophoglycerate leading to increased oxygen affinity and decreased oxygen delivery; decrease of the adenosine triphosphate resulting from increased cell fragility and RBCs becoming less resistant to oxidative stress [79]. RBCs exhibit oxidative damage through characteristic changes in the size and shape of the cell morphology along with alterations in the biochemical structure [80]. These changes are associated with, for example, the loss of RBC membrane integrity and oxidative stress [81,82] that can lead to the generation of smaller, fragile RBCs, formation of microvesicles, and the release of free Hb and heme into the supernatant (hemolysis) [76,78,83].

Information on the quantity of metHb, modification of hemoprotein structure, structural changes in RBC membrane connected with its peroxidation may serve as general markers for oxidative stress in RBCs [84]. Both clinical and animal experimental studies have indicated that storage of RBC diminishes their ability to oxygenate the tissue [85]. The main changes in red cell storage medium include accumulation of bioactive substances such as cytokines, histamines, lipids, or enzymes. This increases the oxidative environment for RBCs and has an impact on the febrile transfusion reactions and immunologic activation/suppression [79]. All of these changes have impact or are connected with changes in other blood fractions and general blood properties like increased blood viscosity and impaired flow [86].

Both the effectiveness of the preservation method and the quality of preserved RBCs require rigorous testing. Currently, there are only a few methods available for this purpose. A universal check of RBCs quality includes the following variables: (1) glucose concentration; (2) ATP levels; (3) extracellular Hb (hemolysis) levels; and

(4) posttransfusion survival levels of RBCs within 24 hours [87]. There is no single assay test available for comprehensive profiling of RBCs [87]. Furthermore, the current tests require sample processing, which potentially can introduce artifacts (e.g., centrifugation may promote hemolysis) [88], are time-consuming and often expensive (e.g., metabolomic assays) [78]. More importantly, none of the available methods are noninvasive and enable testing of the content inside blood bags [77,78]. This is particularly important, since RBCs from the segments attached to a bag do not necessarily correlate with RBCs within the bag [89].

Several studies utilized IR and Raman spectroscopy to identify major changes in RBCs during storage. Pistorius et al. [87] used ATR-FTIR spectroscopy to determine a variety of biochemical changes in intact RBCs and RBC lysates, over the period of 50 days. These were associated with proteins, lipids, adenine nucleotides, and glucose. The pattern of protein changes was shown to differ for RBC lysates and intact RBCs, indicating that the cytosolic protein content and characteristics remain the same over 50 days of storage. However, substantial changes were shown for intrinsic membrane proteins and membrane-associated proteins. These undergo a temporary increase, possibly reflecting binding of cytosolic components (e.g., damaged Hb) and subsequently decrease. Furthermore, a decrease in the content of α-helical membrane proteins was reported, indicating a significant loss of membrane cytoskeletal proteins (e.g., spectrin). These changes are accompanied by a decrease in the membrane lipid content, reflecting lower packing density. In fact, lipid-related changes in erythrocyte membranes were previously observed using FTIR and were reported to occur within the first week of RBCs storage, preceding any changes in protein content and/or secondary structure [90]. In addition, ATR was able to easily monitor ATP levels, revealing an increase after 29 days of storage, followed by a decrease. Importantly, after 35 days of storage the ATP levels reached 76% of the initial concentration (with the required minimum according to current standards being 75%) and further decreased to 63% after 50 days. In the same manner, the ability to determine glucose levels using IR spectroscopy was shown numerous times, using various chemometric tools, for example, PLS [91] and linear discriminant analysis (LDA) [87].

The mechanistic insights into the storage-related alterations in RBCs was also broadly investigated by Raman spectroscopy. RS is particularly useful for studies of RBCs because Hb and its derivatives exhibit a strong resonant Raman scattering. Specifically, RS provides a rather unique ability to distinguish between oxyHb and metHb, compared to other available research tools [44,51]. The major oxyHb bands are located at 1638, 1565, 1225, and 569 cm^{-1}, whereas for metHb marker bands include ones at 1610, 1516, 1216, and 506 cm^{-1}. Gautam et al. [78], applied these bands to demonstrate an increase in both oxyHb and metHb during storage over 21 days for washed RBCs. The authors compared the results of oxyHb and metHb obtained using washed and unwashed pellets of RBCs. The washed pellets showed intracellular changes in the RBCs, and while the unwashed cells

showed a cumulative effect for hemolysis products, the same trend was observed in both cases for oxyHb, suggesting that the changes in oxygen affinity reflect the intracellular RBCs environment. A correlation between Hb oxygenation and storage time was also reported by Atkins et al. [92]. However, for metHb significant changes were observed in the identified trends, when using washed RBCs, compared to unwashed ones [78]. This indicates that RS measurements of metHb could be indicative of hemolysis. Furthermore, using the bands at 1132 and 1064 cm^{-1} (C−C stretching, all-*trans* lipids) and at 721 cm^{-1} (C−N stretching, lipid head groups) a change in membrane lipids leading to formation of more rigid RBCs over storage was shown [78]. This was accompanied by changes in the globin polypeptide chain of globin and denaturation (961 cm^{-1}) increasing with storage time. Another marker of RBCs' degradation, which can potentially be used to assess the quality of RBCs, are the lactate levels. Increasing lactate levels can result from ongoing glucose metabolism. Buildup of lactate reduced the pH in blood bags and affected the integrity of cells. A combination of RS, principal component analysis (PCA) and band target entropy minimization (aimed at recovering a pure component spectrum from a mixture spectrum) was used to determine the lactate levels. Atkins et al. [77] used for this purpose a Raman band at 853 cm^{-1} and the supernatant obtained after pelletizing RBCs (as opposed to intact RBCs). The lactate levels were already measurable within the first 2 days of storage, highlighting the potential to apply this approach for RBC quality testing. The same group explored the possibility of using the increasing oxygenation of Hb as a quality test for RBCs directly within the blood bags using spatially offset Raman spectroscopy [81]. The increased oxygenation was demonstrated several times for samples obtained by probing the blood bags [78,92]. This effect is also expected to occur within the bags, due to the semi gas-permeable nature of the PVC transfusion bag. A good correlation between Hb oxygenation and storage time was demonstrated for data acquired directly from PVC bags, despite the PVC contributions to the spectra. Of note is that even with PVC contributions to the spectra, several regions of interest where the bag and RBC contributions do not overlap could be found (including ones containing markers of Hb oxidation) [81]. At the same time, it is important to underline that multiple studies highlighted significant differences between donors [77,78,81,92].

13.6.3 Application to thalassemia and sickle cell anemia

Aside from identifying the presence of parasites and quality control related to storage and functionality of blood, vibrational spectroscopy has found several applications in the characterization and potential detection of RBC disorders. A normal, healthy RBC is a disk-like doughnut shape cell [93] with a flexible cell membrane surrounding a viscous solution of Hb [18]. Hb is a metalloprotein, typically consisting of four globular protein subunits (two α and two β chains), each of which is attached to the heme containing an iron group, which can temporarily link the oxygen [18]. Two common RBCs disorders are related

either to alteration of the shape and mechanical properties of cells (sickle cell disease, SCD) or to abnormal Hb structure (β-thalassemia). These affect the functionality and/or survival rate of RBCs, ultimately leading to severe anemia and the need for regular blood transfusions, often resulting in further complications [94].

β-Thalassemia is a genetic disorder, affecting Hb synthesis and resulting in abnormal Hb formation [94,95]. It can be found in the Middle East and Mediterranean region [95], however, its particularly high prevalence and incidence (up to 24% of population identified as gene careers) occurs in Southern China [96]. Genetic mutations associated with β-thalassemia (approximately. 200 types of point mutations or small deletions/insertions [95]) lead to the reduction or absence of β-globulin chains, unbalanced globin chains synthesis, and finally an excess of α-chains [94]. Often β-thalassemic patients rely on lifelong blood transfusions and iron-chelating drugs. The most effective means to reduce the prevalence of β-thalassemia is prenatal diagnosis and screening [94–96]. Screening methods include complete blood count (with emphasis on presence of anemia, low mean corpuscular volume, hypochromia, etc.), high-performance liquid chromatography, cellulose acetate electrophoresis (at pH 9.2), etc. [94–96]. These tests are relatively complicated, often time-consuming and require multiple reagents and various measuring instruments. Hence, there is a strong need for a new rapid screening approach.

The possibility of utilizing either Raman or IR spectroscopy for rapid screening is important in understanding the biochemical changes (and reflecting them as spectral markers) associated with β-thalassemia.

Raman tweezers were used by Chen et al. [97] and Rusciano et al. [18] to investigate β-thalassemic erythrocytes. Chen et al. [97] used the ratio of bands at 1638 and 1547 cm^{-1} to distinguish oxygenated and deoxygenated erythrocytes and demonstrated that RBCs of thalassemia (HbH-CS) were easy to oxygenate, but difficult to deoxygenate [97]. Rusciano et al. [18] also observed differences in Hb oxygenation via blue shift of the band at 1588 cm^{-1} ($\nu(C_\alpha C_m)_{asym}$) in erythrocytes from healthy volunteers to 1581 cm^{-1} in β-thalassemic erythrocytes. The authors additionally investigated the differences in distribution of Hb between normal and β-thalassemic erythrocytes. Of note is the fact, that although the study included multiple RBCs per groups, it originated only from three individuals belonging to three experimental groups (healthy, α- and β-thalassemia, $n_{group} = 1$).

A broader study, including 147 patients (65 age-matched healthy and 82 with β-thalassemia: 33 major and 49 minor) was conducted by Jia et al. [95]. The authors reported an increase in β-sheet and β-turns reflecting the change in the proportion of globin chains for β-thalassemic erythrocytes (via the bands at 1546 and 1448 cm^{-1}). Simultaneously, the study did not indicate any differences in Hb among different genotypes of patients with β-thalassemia major. A similar increase in the content of β-sheet proteins and decrease of

α-helical proteins was reported by Liu et al. [94] in a study including lysed RBCs from 92 patients (35 age-matched healthy, 56 with β-thalassemia major, and 1 with Hb H disease) based on IR spectroscopy. Interestingly, the authors attempted to utilize, in addition to the fingerprint region, spectral features associated with the −SH stretching vibration present at around 2550 cm^{-1}. This band originates from thiols of cysteine residuals and, although it is generally very weak, it is present in a spectral region free from other bands from biological compounds. In Hb there are six cystidine residuals, two associated with the α-chain and four with the β-chain. These give rise to bands located at 2553 and 2589 cm^{-1}, respectively. Although the change in their relative intensities between healthy and β-thalassemic RBC lysates was noticeable, the use of only the selected range from the fingerprint (1800−1400 cm^{-1}) enabled more accurate clustering via HCA, compared to the inclusion of the -SH spectral range (2600−2500 cm^{-1}). A LDA based on six selected spectral regions (from fingerprint and high wave number regions) enabled the accuracy of the classification of 98% with only one misclassification (false negative) [94]. Chen et al. [96] also showed the ability for rapid classification of healthy and β-thalassemic patients, using twice diluted hemolytic solution samples. Moreover, a prediction of Hb and HbA_2 content with 100% sensitivity and specificity was possible.

Almost all Raman and IR studies of β-thalassemia, focused on mechanistic and diagnostic objectives, are conducted on RBCs or RBCs lysates. Perrone et al. [98] proposed a different approach, centered around detection of the body's redox state via RRS-based monitoring of carotenoid levels in skin. The level of skin carotenoids (termed skin carotenoid score, SCS) can be considered as a biomarker of the body's antioxidant status. Carotenoids are present in significant amounts in healthy tissues and can provide an enhancement of the Raman signal in the order of 10^5 (using 488-nm excitation wavelength) [99]. Regular blood transfusions and increased absorption of iron (further saturated to ferritin) in β-thalassemic patients leads to accumulation of low-molecular-weight iron, resulting in increased production of reactive oxygen species (ROS) and potentially organ damage (iron toxicity). The group previously demonstrated reduced plasma levels of antioxidants in β-thalassemic patients (via nonspectroscopic methods) [100] and subsequently demonstrated that SCS obtained by RRS is a reflection of their plasma concentrations [98]. The SCS was nearly twice as low for β-thalassemic patients, compared to healthy ones. Furthermore, the authors demonstrated a correlation between SCS and plasma alanine and aspartate transaminase—commonly used parameters for iron status and pathology development in the liver. An advantage of this work is the fact that the measurements were performed *in vivo*, from the upper 0.1-mm stratum corneum of skin of the palm and hand, in a truly rapid and noninvasive manner.

Unlike β-thalassemia, SCD is caused by a single, specific point mutation [93,101] (replacement of adenine for thymine [101]) encoding a modification in the β-globulin chain of Hb (more specifically, replacement of glutamic acid with valine at the sixth position)

[101,102]. This type of Hb is termed Hb S (HbS), in contrast to unaltered Hb, termed HbA. HbS upon deoxygenation polymerizes [101,102] and is accumulated in RBCs, leading to their loss of elasticity and characteristic sickle shape. The result is that RBCs become incapable of deformation as they travel through narrow blood capillaries, which leads to vessel occlusion, insufficient oxygen delivery, and subsequent severe consequences (e.g., organ damage) [93,101,102]. In addition, due to greater fractions in deformed RBCs, their viability is severely affected with an average life span reduced to 15−30 days, compared to 120 days for normal RBCs [93]. This obviously has a significant impact on the life expectancy of SCD patients.

Only a limited amount of literature utilizing Raman spectroscopy for SCD studies is available [101−104]. Wood and McNaughton [104] recorded the first Raman spectra and images of sickle cells. The spectra showed distinct difference between deoxygenated sickle cells and normal deoxygenated cells with strong evidence of heme aggregation in the sickle cells as revealed by a number of intense bands at 1235, 1249, 1366, and 1397 cm^{-1} known to increase in response to protein denaturation [39]. The images showed Hb distributed in different concentrations throughout the sickle cell [104]. Other studies investigated the chemical characterization of altered RBCs [101] resulting in impaired functionality [102] and promising, preliminary attempts for rapid diagnosis [101,103]. Liu et al. [102] studied SCD RBCs using laser tweezers Raman spectroscopy (LTRS). The group previously reported a new phenomenon, in which the oxygenated state of RBCs can be manipulated mechanically by changing the laser trap power. This enables one to increase the deoxygenation of RBC by increasing the laser power leading to an increase in mechanical forces applied to the cell by the laser trap. Using this approach, both SCD RBCs and normal RBCs transitioned from an oxygenated to deoxygenated state [established using Raman bands at 1224, 1371, 1566, 1639 (oxygenated), and 1211, 1547 cm^{-1} (deoxygenated)]. However, the SCD RBCs deoxygenated more rapidly compared to normal RBCs [102]. In fact, the oxygenation pattern correlated with content of HbS. Importantly, under application of no mechanical force the cells were indistinguishable. The authors attributed the difference in the oxygenation profile to a weaker binding affinity of SCD RBCs to oxygen, which is released upon mechanical perturbation much earlier than in normal RBCs. Differences in mechanical stiffness between cells may also contribute to this effect [102]. This approach is particularly unique, as it utilizes LTRS to determine the mechanochemical oxygenation properties.

Filho et al. [101] focused on a strictly chemical difference between SCD and normal RBCs. The main expected chemical difference between these lies in the minor difference in composition of Hb (replacement of glutamic acid with valine). The group was able to demonstrate differences in Raman spectral of SCD and normal RBCs attributed to these components. PCA combined with analysis of variance showed statistically significant loadings that are important for discrimination. In particular a higher intensity of bands at

882 and 1373 cm^{-1} for SCD RBCs relative to normal RBCs was observed. These bands are attributed to valine. On the other hand, normal RBCs exhibited an increase in intensity of bands at 1547 and 1622 cm^{-1}, assigned to glutamic acid [101]. Although the authors applied only an exploratory data analysis method and no discriminatory method, these results clearly showed the ability of RRS to reflect modifications of Hb composition and provide a promising basis for its further exploration as a potential rapid diagnostic tool.

13.6.4 Optical laser tweezer studies on red blood cells

LTRS is a technique combining optical tweezers (also known as optical traps) with Raman spectroscopy [18,102,105–112]. Optical (or laser) tweezers use focused laser light to create an intensity gradient over a small spatial dimension, in order to trap microscopic dielectric objects (suspended in liquid media) by the action of gradient forces [105,106]. The optical tweezers method can be coupled with various other research methods, for example, fluorescence. The relatively recent combination of laser tweezers with Raman spectroscopy attracted a lot of interest because it opens new research possibilities, including, for example, studies at the single-cell level under physiological conditions [105–108]. The most commonly adapted LTRS systems utilize near-infrared wavelength light at 1064 nm for object trapping and 785-nm light for Raman excitation, although the first systems were based on a single laser beam for both [105,106]. A broad review presenting the technological advantages of LTRS and a wide range of applications was published by Chang [105].

An advantage of LTRS is the ability to probe living cells, without the use of any fixative or cell immobilization [106]. Immobilization by chemical or physical means can affect the physiochemical microenvironment and affect the electrochemical properties across the cell membrane, disrupting cellular functions [106]. One of the first applications of LTRS showed the ability to probe a single, living RBC in saline [107]. Bankapur et al. [106] later demonstrated that by using this approach additional bands in RRS spectra of RBCs can be observed, compared to conventional measurement methods [106]. The authors attributed the bands to Hb and the ability to observe the larger optical density achieved in the experimental arrangement, contributing to higher signal sensitivity [106]. The use of different Raman excitation wavelengths (514 and 633 nm) in LTRS systems was also demonstrated [108].

The Raman spectral profile of Hb is sensitive to oxygenation and causes significant shifts enabling oxyHb, deoxyHb, and metHb to be distinguished using the LTRS approach. The LTRS approach combined with a microfluidic device was adapted to study the oxygenation cycle in single, living RBCs [109]. Later on, Rao et al. [110] demonstrated for the first time the ability of LTRS not only to detect the oxygenation status in RBCs, but also to directly induce transition between oxyHb and deoxyHb through the mechanical forces applied to

RBCs by laser tweezers. Upon stretching of RBCs by laser tweezers the cells transited to the deoxygenated state. This occurred when the cells were stretched up to an additional 40% of their original size. The authors attributed the phenomenon of mechanically induced change in oxygenation status to enhanced rates of interactions between neighboring Hb molecules and interactions of Hb with the cell membrane [110]. Similar results were achieved by Rusciano et al. [18], who additionally demonstrated that the process is reversible. Subsequently, Liu et al. [113] confirmed that RBCs trapped in single-beam optical tweezers undergo deoxygenation and reported that they become more deoxygenated with increasing laser trap power. Ahlawat et al. [111] conducted a broad study to further investigate the phenomenon of laser trap–induced deoxygenation and its dependence on power. The authors showed that the deoxygenation is significantly more pronounced in the cellular sites overlapping with the laser tweezer beam, compared to the rest of the cell. Furthermore, the authors demonstrated the same effect of power-dependent laser trap deoxygenation occurring in Hb solutions, indicating that it cannot be attributed to mechanical deformations of cells. Instead, the authors hypothesized that it results from photoinduced dissociation of oxygen from heme [111]. Of note, however, is the importance of using relatively low laser power for Raman excitation, as it was reported that optically trapped RBCs can demonstrate photodegradation of Hb (denaturation, hemichrome formation, and potential aggregation) even upon usage of 785-nm excitation wavelength with 5 mW [114]. The ability to study the transition between oxygenated and deoxygenated states of RBCs in response to trapping by laser tweezers was used to investigate oxygenation profiles (and thus oxygen affinity) in relation to various blood disorders, including β-thalassemia [18,97] and sickle cell disease [102].

LTRS was also applied to probe oxidative stress in living RBCs [115]. The oxidative stress was induced by the introduction of hydroxyl radicals generated by the Fenton reaction. The authors reported a substantial increase in the intensity of bands in the region between 650–500 cm^{-1} (665, 640, 565, 550, 519, 500 cm^{-1}) associated with stretching of Fe–O_2 (565 cm^{-1}, oxyHb) or S–S stretching. This reflects increased oxygenation of oxidative stress–induced RBCs and increased concentration of disulfide bridges, most likely resulting from the formation of protein-ss-glutathione, known to be secreted upon oxidative stress [115]. The oxidative state was also examined for RBCs infected with a malaria parasite *Plasmodium vivax* [112]. The authors reported a significant decrease in the low-spin Raman marker band at 1223 cm^{-1} (oxyHb) and an increase in the high-spin Raman marker band at 1210 cm^{-1} (deoxyHb) for iRBC compared to uninfected controls. This indicates reduced oxygen affinity to Hb in iRBCs and remains consistent with the anemic condition and acute respiratory syndrome commonly diagnosed in patients with *P. vivax* infection [112].

Raj et al. [116] applied LTRS to study the mechanochemical properties of RBCs in the context of structural changes upon cell deformation. The mechanical flexibility of RBCs is one of their crucial characteristics because the cells—typically with a diameter of

7−8.5 μm—often need to pass through narrow capillaries, smaller than 3 μm [117]. The authors stretched the cell by attaching two microparticles to the opposing ends of RBC and manipulate their position via the optical trap [116]. This enabled them to avoid direct influence of the laser trap on the cell and study specific changes occurring upon temporary cell deformation. In addition to the changes in Hb, the authors reported substantial differences in the band at 1035 cm^{-1}, which remained constant upon low deformations and continuously increased at intermediate deformation (10%−20%) and reached saturation around the deformation level of 20%. The authors attributed the spectral differences to linker proteins, spectrin, and Hb attached to the cell membrane (10^6 Hb/cell). At low deformation levels spectrin bears most of the forces and rearranges itself (without changes in primary structures). At the intermediate level the linker proteins, spectrin network, and membrane-related Hb undergo severe structural perturbation. At high deformation levels a nonlinear response was observed and attributed to a release of spectrin filaments from linker proteins and their rebonding [116].

Another application of LTRS to studies of RBC was shown by Deng et al. [118], who demonstrated a decrease in the viability of RBCs by 34.1% over a period of 25 minutes after admission of 20% alcohol solution. The viability of cells was characterized using the Raman band at 752 cm^{-1} (porphyrin breathing mode) and was shown to decrease more rapidly upon increase of the alcohol concentrations [118].

13.6.5 Detection, differentiation, and visualization of the dysfunctional Hb adducts and its metabolites

Detection and differentiation of dysfunctional Hb forms is important and carries enormous diagnostic potential related to the possibility of showing damage and changes within the RBCs. The oxidation of RBCs is expressed by oxyHb degradation and mainly by metHb (metHb) and ROS production. MetHb is unable to bind oxygen as it has high-spin iron ion at +3 oxidation state due to a bonded OH^- ion or H_2O molecule in the sixth coordination site [11]. The next step of the oxidation process of RBCs involves denaturation of metHb to hemichromes and additional ROS formation [119]. The H_2O molecule at the sixth coordination site in metHb is replaced by an amino acid (typically His) from a distal part (E7) of the globin chain. This oxidation process is accompanied by the release of a superoxide (O^{2-}) radical with conversion of iron to a low-spin Fe(III) state that has lost its paramagnetic properties [13]. It was postulated that hemichromes may precipitate in RBCs to form coccoid inclusions called Heinz bodies or Heinz−Ehrlich bodies. It was previously shown that hemichrome formation is an innate characteristic of physiologically normal Hb molecules and that Heinz body formation within normal human erythrocytes may be a potential marker of oxidative level in RBCs [120]. A new type of aggregates, an advanced Heinz body-like

aggregates (AHBA), were also studied with the application of RS and complementary techniques. AHBA were formed in RBCs in response to glutaraldehyde treatment and were characterized by a higher level of hemoglobin (Hb) degradation compared to typical Heinz bodies [121].

The detailed analysis of the heme aggregation and denaturation within human RBCs were previously carried out with the application of RRS [39,122]. The comprehensive studies of the chemical imaging of human RBCs in smears was also shown [123] and the high-resolution *in situ* distributions of the Fe^{2+}/Fe^{3+} ions of heme inside the RBCs were demonstrated. Analysis of thin smears of air-dried whole blood allows for the imaging of a thin outer layer of hemichrome in the periphery of RBCs. Such RRS studies were additionally supported with the results from AFM and scanning near-field optical microscopy, as presented in Fig. 13.15.

The studies on the dysfunctional form of Hb found inside murine isolated primary Kupffer cells and the murine macrophage cell line were also presented [124]. They have shown that RRS was found useful for studies of different stages of erythrophagocytosis of senescent RBCs. RRS allowed for detection of oxyHb, metHb, hemichrome, hematin, and biliverdin

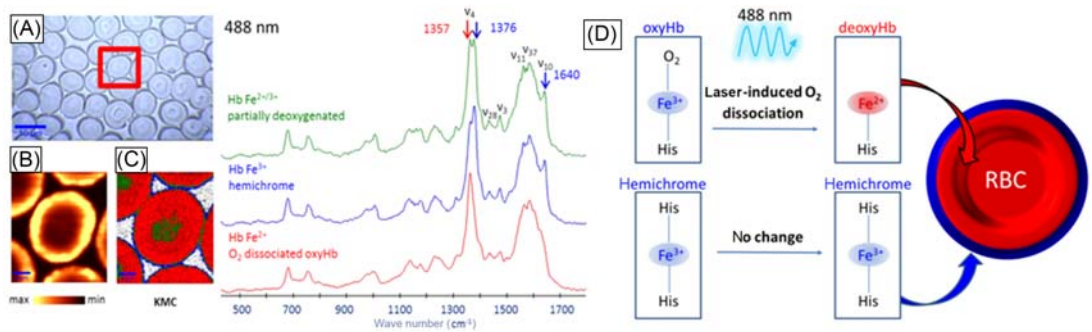


Figure 13.15

(A) Representative white-light optical image (× 60) of the air-dried human RBCs (whole blood, smear on the CaF_2 slide) with the labeled (*red*) investigated area; Raman maps obtained with 488-nm laser wavelength and with a sampling density of 0.11 μm (maximal spatial resolution equal to 0.33 μm). (B) Integration maps of a Hb's band in the spectral region of around 1350–1380 cm^{-1} (ν_4). (C) The K-means clustering (KMC) results with the three main classes for chosen RBCs. Classes include hemichrome Hb Fe^{3+} (*blue*), laser-induced O_2 dissociated oxyHb, Hb Fe^{2+} (*red*), and partially deoxyHb $Fe^{2+/3+}$ (*green*). Gray class observed between RBC's corresponds to substrate signal and is removed as outliers. The average spectra of the respective three classes of RBCs obtained from the KMC. (D) Model of changes of Hb inside air-dried RBC measured with 488-nm wavelength where laser-induced O_2 dissociation of oxyHb to deoxyHb takes place and hemichrome remains unchanged. ***Source*: *Reproduced from K.M. Marzec, et al., High-resolution Raman imaging reveals spatial location of heme oxidation sites in single red blood cells of dried smears, J. Raman Spectrosc. 46 (1) (2015) 76–83 [123] with permission from John Wiley and Sons.*

in situ in macrophages and Kupffer cells. Work also confirmed that during the erythrophagocytosis process the massive accumulation of heme derivatives is followed by further lipid accumulation in the form of lipid droplets [125,126]. Additionally, RRS was found to be a useful technique to image RBCs in tissue vessels. Spectroscopic studies of atherosclerotic plaque biomarkers as well as the possibility of diagnostics of unstable atherosclerotic plaques based on the presence of intraplaque thrombus were also reported [127–129].

13.7 Conclusion

The application of RRS to RBCs research has blossomed over the past 20 years providing insights into how Hb behaves in the highly concentrated environment of the RBCs. The fundamental observations of how different excitation wavelengths produce vastly different profiles to aqueous solutions has led to an understanding of the excitation profile of Hb within functional erythrocytes. The importance of understanding how laser power and accumulation time can result in protein denaturation of heme aggregation, along with the formation of other heme-related molecules, like metHb and hemichrome, needs to be carefully understood in order to interpret spectra of these light-sensitive molecules. When these conditions are met there are many possibilities and applications that can be derived from this highly sensitive technique, from malaria diagnosis to detecting abnormal Hbs in cells through to blood storage. The continued development of faster CCDs and combined Raman and infrared systems will see the technology find more exciting applications in the future.

Acknowledgments

BRW was supported by an Australian Research Council (ARC) Future Fellowship grant FT120100926. KMM would like to thank the support from the National Centre for Research and Development, Poland (LIDER/13/0076/L-8/16/NCBR/2017).

References

[1] M.C. Tobin, Raman spectra of crystalline lysozyme, pepsin, and alpha chymotrypsin, Science 161 (3836) (1968) 68–69.
[2] T.C. Strekas, T.G. Spiro, Hemoglobin: resonance Raman spectra, Biochim. Biophys. Acta 263 (3) (1972) 830–833.
[3] H. Brunner, A. Mayer, H. Sussner, Resonance raman scattering on the haem group of oxy- and deoxyhaemoglobin, J. Mol. Biol. 70 (1) (1972) 153–156.
[4] H. Brunner, Identification of the iron-ligand vibration of oxyhemoglobin, Naturwissenshaften 129 (1974).
[5] B. Benko, N.T. Yu, Resonance Raman studies of nitric oxide binding to ferric and ferrous hemoproteins: detection of Fe(III)--NO stretching, Fe(III)--N--O bending, and Fe(II)--N--O bending vibrations, Proc. Natl Acad. Sci. USA 80 (22) (1983) 7042–7046.

[6] K. Bajdor, H. Oshio, K. Nakamoto, Observation of two oxygen-isotope sensitive bands in the low-frequency resonance Raman spectrum of oxy (phthalocyanato) iron (II), J. Am. Chem. Soc. 106 (1984) 7273–7274.
[7] B.R. Wood, et al., Resonance Raman spectroscopy of red blood cells using near-infrared laser excitation, Anal. Bioanal. Chem. 387 (5) (2007) 1691–1703.
[8] M. Tsubaki, N.T. Yu, Resonance Raman investigation of nitric oxide bonding in nitrosylhemoglobin A and -myoglobin: detection of bound N-O stretching and Fe-NO stretching vibrations from the hexacoordinated NO-heme complex, Biochemistry 21 (6) (1982) 1140–1144.
[9] R.S. Armstrong, M.J. Irwin, P.E. Wright, Resonance Raman evidence for constrained heme structure in soybean leghemoglobin and its derivatives, Biochem. Biophys. Res. Commun. 95 (2) (1980) 682–689.
[10] T.G. Spiro, T.C. Strekas, Resonance Raman spectra of hemoglobin and cytochrome c: inverse polarization and vibronic scattering, Proc. Natl Acad. Sci. USA 69 (9) (1972) 2622–2626.
[11] T.G. Spiro, T.C. Strekas, Resonance Raman spectra of heme proteins. Effects of oxidation and spin state, J. Am. Chem. Soc. 96 (2) (1974) 338–345.
[12] T. Yamamoto, et al., The Valence and Spin State of Iron in Oxyhemoglobin as Inferred from Resonance Raman Spectroscopy, Biol. Chem. 248 (1973) 5211.
[13] J.J. Weiss, Nature of the iron-oxygen bond in oxyhaemoglobin, Nature 203 (1964) 182–183.
[14] L.D. Spaulding, et al., Resonance Raman spectra of metallooctaethylporphyrins. Structural probe of metal displacement, J. Phys. Chem. Soc. 97 (1975) 2517.
[15] S. Asher, Resonance Raman spectroscopy of hemoglobin, Methods Enzymol. 76 (1981) 371–413.
[16] T.G. Spiro, in: T.T. Theophanides (Ed.), Infrared and Raman Spectroscopy of Biological Molecules, Vol. 1, Reidel Publ. Co, Dordrecht, 1979.
[17] T.G. Spiro, G. Smulevich, C. Su, Probing protein structure and dynamics with resonance Raman spectroscopy: cytochrome c peroxidase and hemoglobin, Biochemistry 29 (19) (1990) 4497–4508.
[18] G. Rusciano, et al., Raman tweezers as a diagnostic tool of hemoglobin-related blood disorders, Sens (Basel) 8 (12) (2008) 7818–7832.
[19] D.L. Rousseau, M.R. Ondrias, Resonance Raman scattering studies of the quaternary structure transition in hemoglobin, Annu. Rev. Biophys. Bioeng. 12 (1983) 357–380.
[20] M. Nagai, et al., A role of heme side-chains of human hemoglobin in its function revealed by circular dichroism and resonance Raman spectroscopy, Biophys. Rev. 10 (2) (2018) 271–284.
[21] M.Z. Lu, et al., Application of Raman spectroscopy to the research on hemoglobin structure and function, Guang Pu Xue Yu Guang Pu Fen. Xi 34 (2) (2014) 439–444.
[22] T. Kitagawa, Raman and infrared analysis of cooperativity in hemoglobin, Tanpakushitsu Kakusan Koso 32 (6) (1987) 584–593.
[23] J.M. Friedman, Time-resolved resonance Raman spectroscopy as probe of structure, dynamics, and reactivity in hemoglobin, Methods Enzymol. 232 (1994) 205–231.
[24] D.F. Wallach, S.P. Verma, Raman and resonance-Raman scattering by erythrocyte ghosts, Biochim. Biophys. Acta 382 (4) (1975) 542–551.
[25] S.P. Verma, D.F. Wallach, Multiple thermotropic state transitions in erythrocyte membranes. A laser-Raman study of the CH-stretching and acoustical regions, Biochim. Biophys. Acta 436 (2) (1976) 307–318.
[26] S.P. Verma, D.F. Wallach, Erythrocyte membranes undergo cooperative, pH-sensitive state transitions in the physiological temperature range: evidence from Raman spectroscopy, Proc. Natl Acad. Sci. USA 73 (10) (1976) 3558–3561.
[27] S.P. Verma, D.F. Wallach, Carotenoids as a Raman-active probes of erythrocyte membrane structure, Biochim. Biophys. Acta 401 (2) (1975) 168–176.
[28] F.P. Milanovich, et al., Raman spectroscopic analysis of Dutch Belt rabbit erythrocyte ghosts, Chem. Phys. Lipids 17 (1) (1976) 79–84.

[29] R.B. Mikkelsen, S.P. Verma, D.F. Wallach, Effect of transmembrane ion gradients on Raman spectra of sealed, hemoglobin-free erythrocyte membrane vesicles, Proc. Natl Acad. Sci. USA 75 (11) (1978) 5478–5482.
[30] J.L. Lippert, L.E. Gorczyca, G. Meiklejohn, A laser Raman spectroscopic investigation of phospholipid and protein configurations in hemoglobin-free erythrocyte ghosts, Biochim. Biophys. Acta 382 (1) (1975) 51–57.
[31] B.R. Wood, D. McNaughton, Micro-Raman characterization of high- and low-spin heme moieties within single living erythrocytes, Biopolymers 67 (4-5) (2002) 259–262.
[32] S. Franzen, et al., Direct evidence for the role of haem doming as the primary event in the cooperative transition of haemoglobin, Nat. Struct. Biol. 1 (4) (1994) 230–233.
[33] X.-Y. Li, et al., Consistent porphyrin force field. 2. Nickel octaethylporphyrln skeletal and substituent mode assignments from 15N, meso-d4, and methylene-d16 Raman and infrared isotope shifts, J. Phys. Chem. 94 (1990) 47–61.
[34] K.M. Marzec, et al., Red blood cells polarize green laser light revealing hemoglobin's enhanced non-fundamental Raman modes, Chemphyschem. 15 (18) (2014) 3963–3968.
[35] G.J. Puppels, et al., Raman microscopic approach to the study of human granulocytes, Biophys. J. 60 (1991) 1046–1056.
[36] B.L. Salmaso, et al., Resonance Raman microspectroscopic characterization of eosinophil peroxidase in human eosinophilic granulocytes, Biophys. J. 67 (1994) 436–446.
[37] N.M. Sijtsema, et al., Intacellular reactions in single human granulocytes upon phorbol myristate acetate activation using confocal Raman microspectroscopy, Biophys. J. 78 (2000) 2606–2613.
[38] G.J. Puppels, et al., Develop- ment and application of Raman microspectroscopic and Raman imaging techniques for cell biological studies, J. Mol. Struct. 347 (1995) 477–483.
[39] B.R. Wood, et al., Raman microspectroscopy and imaging provides insights into heme aggregation and denaturation within human erythrocytes, J. Biomed. Opt. 10 (1) (2005) 14005.
[40] K. Ramser, C. Fant, M. Käll, Importance of substrate and photo-induced effects in Raman spectroscopy of single functional erythrocytes, J. Biomed. Opt. 82 (2003) 173–178.
[41] M. Asghari-Khiavi, et al., Correlation of atomic force microscopy and Raman micro-spectroscopy to study the effects of ex vivo treatment procedures on human red blood cells, Analyst 135 (2010) 525–530.
[42] B.R. Wood, D. McNaughton, Resonance Raman spectroscopy in malaria research, Expert. Rev. Proteom. 3 (5) (2006) 525–544.
[43] T.G. Spiro, Resonance Raman spectroscopy as a probe of heme protein structure and dynamics, Adv. Protein Chem. 37 (1985) 111–159.
[44] B.R. Wood, D. McNaughton, Resonant Raman scattering of heme molecules in cells and in the solid state, in:P. Lasch, J. Kneipp (Eds.), Biomedical Vibrational Spectroscopy, 2008, John Wiley & Sons, Inc.
[45] G. Smulevich, A. Feis, B.D. Howes, Fifteen years of Raman spectroscopy of engineered heme containing peroxidases: what have we learned? Acc. Chem. Res. 38 (5) (2005) 433–440.
[46] D. Perez-Guaita, et al., Parasites under the spotlight: applications of vibrational spectroscopy to malaria research, Chem. Rev. 118 (11) (2018) 5330–5358.
[47] R. Barer, Refractometry and interferometry of living cells, J. Opt. Soc. Am. 47 (6) (1957) 545–556.
[48] B.R. Wood, D. McNaughton, Raman excitation wavelength investigation of single red blood cells in vivo, J. Raman Spectrosc. 33 (2002) 517–523.
[49] M. Abe, T. Kitagawa, Y. Kyogoku, Resonance Raman spectra of octaethylporphyrinato–Ni(II) and meso–deuterated and 15N substituted derivatives. II. A normal coordinate analysis, J. Chem. Phys. 19 (1978) 4526–4534.
[50] S. Hu, K.M. Smith, T.G. Spiro, Assignment of protoheme resonance Raman spectrum by heme labeling in myoglobin, J. Am. Chem. Soc. 118 (1996) 12638–12646.
[51] B.R. Wood, B. Tait, D. McNaughton, Micro-Raman characterisation of the R to T state transition of haemoglobin within a single living erythrocyte, Biochim. Biophys. Acta (BBA) Mol. Cell Res. 1539 (1) (2001) 58–70.

[52] B.R. Wood, et al., Resonance Raman spectroscopy reveals new insight into the electronic structure of beta-hematin and malaria pigment, J. Am. Chem. Soc. 126 (30) (2004) 9233–9239.
[53] D. Perez-Guaita, et al., Resonance Raman and UV-visible microscopy reveals that conditioning red blood cells with repeated doses of sodium dithionite increases haemoglobin oxygen uptake, ChemistrySelect 2 (2017) 3342–3346.
[54] S. Aramaki, H. Hamaguchi, M. Tasumi, Resonance Raman-scattering with a forbidden but vibronically allowed electronic-transition - excitation profiles of combinations and overtones of copper tetraphenylporphyrin in the Q-band resonance, Chem. Phys. Lett. 96 (5) (1983) 555–559.
[55] S.A. Asher, UV resonance Raman studies of molecular structure and dynamics: applications in physical and biophysical chemistry, Annu. Rev. Phys. Chem. 39 (1988) 537–588.
[56] M. Kozicki, et al., An attenuated total reflection (ATR) and Raman spectroscopic investigation into the effects of chloroquine on *Plasmodium falciparum*-infected red blood cells, Analyst 140 (7) (2015) 2236–2246.
[57] B.R. Wood, et al., Resonance Raman microscopy in combination with partial dark-field microscopy lights up a new path in malaria diagnostics, Analyst 134 (6) (2009) 1119–1125.
[58] B.R. Wood, et al., Raman imaging of hemozoin within the food vacuole of *Plasmodium falciparum* trophozoites, FEBS Lett. 554 (3) (2003) 247–252.
[59] F. Chen, et al., Direct detection of malaria infected red blood cells by surface enhanced Raman spectroscopy, Nanomed.: Nanotechnol. Biol. Med. 12 (6) (2016) 1445–1451.
[60] A.J. Hobro, et al., Raman spectroscopic analysis of malaria disease progression via blood and plasma samples, Analyst 138 (14) (2013) 3927–3933.
[61] A.F. Slater, et al., An iron-carboxylate bond links the heme units of malaria pigment, Proc. Natl Acad. Sci. USA 88 (2) (1991) 325–329.
[62] C. Brémard, et al., Spectroscopic investigations of malaria pigment, Appl. Spectrosc. 47 (11) (1993) 1837–1842.
[63] C.W. Ong, et al., Resonance Raman microspectroscopy of normal erythrocytes and *Plasmodium berghei*-infected erythrocytes, Appl. Spectrosc. 53 (9) (1999) 1097–1101.
[64] L. Puskar, et al., Raman acoustic levitation spectroscopy of red blood cells and *Plasmodium falciparum* trophozoites, Lab. Chip 7 (9) (2007) 1125–1131.
[65] Q. Liu, et al., Surface enhanced Raman spectroscopy for malaria diagnosis and intradermal measurements, SPIE BiOS, Vol. 10509, SPIE, 2018.
[66] K. Chen, et al., Towards ultrasensitive malaria diagnosis using surface enhanced Raman spectroscopy, Sci. Rep. 6 (2016) 20177.
[67] C. Yuen, Q. Liu, Magnetic field enriched surface enhanced resonance Raman spectroscopy for early malaria diagnosis, J. Biomed. Opt. 17 (1) (2012) 017005.
[68] B.M. Cooke, N. Mohandas, R.L. Coppel, Malaria and the red blood cell membrane, Semin. Hematol. 41 (2) (2004) 173–188.
[69] F. Wunderlich, et al., Distinct lipid compositions of parasite and host cell plasma membranes from *Plasmodium chabaudi*-infected erythrocytes, Mol. Biochem. Parasitol. 44 (2) (1991) 271–277.
[70] N.L. Garrett, et al., Bio-sensing with butterfly wings: naturally occurring nano-structures for SERS-based malaria parasite detection, Phys. Chem. Chem Phys 17 (33) (2015) 21164–21168.
[71] M. Bilal, et al., Optical diagnosis of malaria infection in human plasma using Raman spectroscopy, J. Biomed. Opt. 20 (1) (2015) 017002.
[72] G.T. Webster, et al., Resonance Raman spectroscopy can detect structural changes in haemozoin (malaria pigment) following incubation with chloroquine in infected erythrocytes, FEBS Lett. 582 (7) (2008) 1087–1092.
[73] L. Frame, et al., Development of a label-free Raman imaging technique for differentiation of malaria parasite infected from non-infected tissue, Analyst 143 (1) (2017) 157–163.
[74] J.W. Kang, et al., Combined confocal Raman and quantitative phase microscopy system for biomedical diagnosis, Biomed. Opt. express 2 (9) (2011) 2484–2492.

[75] D. Perez-Guaita, et al., Multimodal vibrational imaging of cells, Vibrat. Spectrosc. 91 (2017) 46–58.
[76] E.A. Hod, et al., Transfusion of red blood cells after prolonged storage produces harmful effects that are mediated by iron and inflammation, Blood 115 (21) (2010) 4284–4292.
[77] C.G. Atkins, et al., Raman spectroscopy of stored red blood cells: evaluating clinically relevant biochemical markers in donated blood, Clinical and Biomedical Spectroscopy and Imaging IV, Optical Society of America, Munich, 2015.
[78] R. Gautam, et al., Characterization of storage-induced red blood cell hemolysis using Raman spectroscopy, Lab. Med. 49 (4) (2018) 298–310.
[79] M.J. Vandromme, G. McGwin Jr., J.A. Weinberg, Blood transfusion in the critically ill: does storage age matter? Scand. J. Trauma. Resusc. Emerg. Med. 17 (2009) 35.
[80] M.A. Chaves, M.S. Leonart, A.J. do Nascimento, Oxidative process in erythrocytes of individuals with hemoglobin S, Hematology 13 (3) (2008) 187–192.
[81] M.Z. Vardaki, et al., Raman spectroscopy of stored red blood cell concentrate within sealed transfusion blood bags, Analyst 143 (24) (2018) 6006–6013.
[82] A. D'Alessandro, F. Gevi, L. Zolla, Red blood cell metabolism under prolonged anaerobic storage, Mol. Biosyst. 9 (6) (2013) 1196–1209.
[83] J.Y. Oh, et al., Predicting storage-dependent damage to red blood cells using nitrite oxidation kinetics, peroxiredoxin-2 oxidation, and hemoglobin and free heme measurements, Transfusion 55 (12) (2015) 2967–2978.
[84] E.U. Nwose, et al., Erythrocyte oxidative stress in clinical management of diabetes and its cardiovascular complications, Br. J. Biomed. Sci. 64 (1) (2007) 35–43.
[85] N.J. Raat, C. Ince, Oxygenating the microcirculation: the perspective from blood transfusion and blood storage, Vox Sang. 93 (1) (2007) 12–18.
[86] A.G. Kriebardis, et al., RBC-derived vesicles during storage: ultrastructure, protein composition, oxidation, and signaling components, Transfusion 48 (9) (2008) 1943–1953.
[87] A.M. Pistorius, et al., A single assay for multiple storage-sensitive red blood cell characteristics by means of infrared spectroscopy, Transfusion 50 (2) (2010) 366–375.
[88] A.L. Hansen, et al., The effect of processing method on the in vitro characteristics of red blood cell products, Vox Sang. 108 (4) (2015) 350–358.
[89] J.D. Kurach, et al., Segments from red blood cell units should not be used for quality testing, Transfusion 54 (2) (2014) 451–455.
[90] W.F. Wolkers, et al., In situ assessment of erythrocyte membrane properties during cold storage, Mol. Membr. Biol. 19 (1) (2002) 59–65.
[91] H.M. Heise, et al., Multivariate determination of glucose in whole blood by attenuated total reflection infrared spectroscopy, Anal. Chem. 61 (18) (1989) 2009–2015.
[92] C.G. Atkins, et al., Using Raman spectroscopy to assess hemoglobin oxygenation in red blood cell concentrate: an objective proxy for morphological index to gauge the quality of stored blood? Analyst 142 (12) (2017) 2199–2210.
[93] V. Masilamani, et al., A novel technique of spectral discrimination of variants of sickle cell anemia, Dis. Markers 2018 (2018) 5942368.
[94] K.-Z. Liu, et al., Infrared spectroscopic identification of β-thalassemia, Clin. Chem. 49 (7) (2003) 1125–1132.
[95] W. Jia, et al., Raman characterizations of red blood cells with beta-thalassemia using laser tweezers Raman spectroscopy, Med. (Baltim.) 97 (39) (2018) e12611.
[96] J. Chen, et al., A rapid quantification method for the screening indicator for beta-thalassemia with near-infrared spectroscopy, Spectrochim. Acta A Mol. Biomol. Spectrosc. 193 (2018) 499–506.
[97] X. Chen, G. Wang, X. Yin, J. Liu, H. Yao, S. Huang, et al., Single-cell Raman spectral analysis of oxygenated and deoxygenated thalassemia erythrocytes, Acta Opt. Sin. 29 (10) (2009) 2854–2859.

[98] A. Perrone, L. Tesoriere, A.M. Pintaudi, A. Attanzio, P. Rigano, A. Maggio, et al., Raman spectroscopy technology to monitor the carotenoids in skin of thalassemia patients: a novel non-invasive tool relating oxidative stress with iron burden, Thalassemia Rep. 4 (2) (2014) 38–42.
[99] I.V. Ermakov, et al., Resonance Raman detection of carotenoid antioxidants in living human tissue, J. Biomed. Opt. 10 (6) (2005) 064028.
[100] M.A. Livrea, et al., Oxidative stress and antioxidant status in beta-thalassemia major: iron overload and depletion of lipid-soluble antioxidants, Blood 88 (9) (1996) 3608–3614.
[101] A.C. Filho, et al., Raman spectroscopy for a rapid diagnosis of sickle cell disease in human blood samples: a preliminary study, Lasers Med. Sci. 30 (1) (2015) 247–253.
[102] R. Liu, et al., Novel single-cell functional analysis of red blood cells using laser tweezers Raman spectroscopy: application for sickle cell disease, Exp. Hematol. 41 (7) (2013) 656–661. e1.
[103] M. Nahavandi, et al., Near-infrared spectra absorbance of blood from sickle cell patients and normal individuals, Hematology 14 (1) (2009) 46–48.
[104] B.R. Wood, D. McNaughton, Resonant Raman studies on functional erythorcytes, in: P.D. O'Malley (Ed.), New Developments in Sickle Cell Disease, Nove Science Publishers Inc, New York, 2006, pp. 63–119.
[105] J.W. Chan, Recent advances in laser tweezers Raman spectroscopy (LTRS) for label-free analysis of single cells, J. Biophoton. 6 (1) (2013) 36–48.
[106] A. Bankapur, et al., Raman tweezers spectroscopy of live, single red and white blood cells, PLoS One 5 (4) (2010) e10427.
[107] C. Xie, M.A. Dinno, Y.-Q. Li, Near-infrared Raman spectroscopy of single optically trapped biological cells, Opt. Lett. 27 (4) (2002) 249–251.
[108] R. Gessner, et al., Identification of biotic and abiotic particles by using a combination of optical tweezers and in situ Raman spectroscopy, Chemphyschem 5 (8) (2004) 1159–1170.
[109] K. Ramser, et al., A microfluidic system enabling Raman measurements of the oxygenation cycle in single optically trapped red blood cells, Lab. Chip 5 (4) (2005) 431–436.
[110] S. Rao, et al., Raman study of mechanically induced oxygenation state transition of red blood cells using optical tweezers, Biophys. J. 96 (1) (2009) 209–216.
[111] S. Ahlawat, et al., Raman spectroscopic investigations on optical trap induced deoxygenation of red blood cells, Appl. Phys. Lett. 103 (18) (2013) 183704.
[112] R. Dasgupta, et al., Studies on erythrocytes in malaria infected blood sample with Raman optical tweezers, J. Biomed. Opt. 16 (7) (2011) 077009.
[113] R. Liu, et al., Power dependent oxygenation state transition of red blood cells in a single beam optical trap, Appl. Phys. Lett. 99 (4) (2011) 043702.
[114] R. Dasgupta, et al., Hemoglobin degradation in human erythrocytes with long-duration near-infrared laser exposure in Raman optical tweezers, J. Biomed. Opt. 15 (5) (2010) 055009.
[115] E. Zachariah, et al., Probing oxidative stress in single erythrocytes with Raman tweezers, J. Photochem. Photobiol. B 100 (3) (2010) 113–116.
[116] S. Raj, et al., Mechanochemistry of single red blood cells monitored using Raman tweezers, Biomed. Opt. Express 3 (4) (2012) 753–763.
[117] Y. Li, et al., Mechanical property analysis of stored red blood cell using optical tweezers, Colloids Surf. B Biointerf. 70 (2) (2009) 169–173.
[118] J.L. Deng, et al., Study of the effect of alcohol on single human red blood cells using near-infrared laser tweezers Raman spectroscopy, J. Raman Spectrosc. 36 (3) (2005) 257–261.
[119] J.M. Gutteridge, B. Halliwell, Iron toxicity and oxygen radicals, Baillieres Clin. Haematol. 2 (2) (1989) 195–256.
[120] Y. Sugawara, et al., Molecular biosensing mechanisms in the spleen for the removal of aged and damaged red cells from the blood circulation, Sens. (Basel) 10 (8) (2010) 7099–7121.

[121] K. Bulat, et al., Multimodal detection and analysis of a new type of advanced Heinz body-like aggregate (AHBA) and cytoskeleton deformation in human RBCs, Analyst 145 (10) (2020) 1749–1758. https://doi.org/10.1039/C9AN01707K.
[122] U. Neugebauer, et al., Spectroscopic detection and quantification of heme and heme degradation products, Anal. Bioanal. Chem. 404 (10) (2012) 2819–2829.
[123] K.M. Marzec, et al., High-resolution Raman imaging reveals spatial location of heme oxidation sites in single red blood cells of dried smears, J. Raman Spectrosc. 46 (1) (2015) 76–83.
[124] J. Dybas, et al., Raman imaging of heme metabolism in situ in macrophages and Kupffer cells, Analyst 143 (14) (2018) 3489–3498.
[125] K. Kochan, et al., IR and Raman imaging of murine brains from control and ApoE/LDLR(-/-) mice with advanced atherosclerosis, Analyst 141 (18) (2016) 5329–5338.
[126] K.M. Marzec, et al., Raman microimaging of murine lungs: insight into the vitamin A content, Analyst 140 (7) (2015) 2171–2177.
[127] K.M. Marzec, et al., Visualization of the biochemical markers of atherosclerotic plaque with the use of Raman, IR and AFM, J. Biophoton. 7 (9) (2014) 744–756.
[128] T.P. Wrobel, et al., Effects of low carbohydrate high protein (LCHP) diet on atherosclerotic plaque phenotype in ApoE/LDLR-/- mice: FT-IR and Raman imaging, Sci. Rep. 5 (2015) 14002.
[129] N.M. Htun, et al., Near-infrared autofluorescence induced by intraplaque hemorrhage and heme degradation as marker for high-risk atherosclerotic plaques, Nat. Commun. 8 (1) (2017) 75.

CHAPTER 14

Ultraviolet Raman spectroscopy for understanding structure and formation mechanism of amyloid fibrils*

Joseph Handen and Igor K. Lednev

Department of Chemistry, University at Albany, State University of New York, Albany, NY, United States

14.1 Introduction

Raman spectroscopy has benefited from improvements in technology making ultraviolet (UV) Raman an accessible technique with many applications. Through the utilization of UV excitation, specific and significant enhancement of signal due to increased efficiency of Raman scattering, can be achieved. In particular, deep UV ($\leq$210 nm) avoids interference from fluorescence, and selectively enhances the peptide backbone, making it suitable for the study of biological studies [1].

Protein fibrillation is often associated with human diseases such as Alzheimer's, for example. The fact that numerous proteins and peptides fibrillize suggests that this conformation lies in the general structural repertory of proteins. The association of fibrillation with human disease motivates investigations of the structure of these fibers and their mechanism of formation [2]. Unfortunately, the two classic tools of structural biology, X-ray crystallography and conventional NMR, are ill-suited for studying noncrystalline and insoluble full-length fibrils [3]. As shown below, deep ultra violet resonance Raman (DUVRR) is uniquely able to probe protein structure at all phases of fibrillation [2–12]. The use of chemometrics [13–19] and two-dimensional correlation spectroscopy (2DCoS) to these UVRR studies dramatically increases the information content, as shown below.

* Adapted in part with permission from Jakubek et al. TrAC, 103 (2018) 223. Copyright 2017 Elsevier; and Oladepo et al. Chem. Rev. 112 (2012) 2604. Copyright 2011 American Chemical Society.

Vibrational Spectroscopy in Protein Research.
DOI: https://doi.org/10.1016/B978-0-12-818610-7.00014-1

14.2 Two-dimensional correlation deep ultraviolet resonance Raman spectroscopy

2DCoS has recently found increasing use in vibrational spectroscopy [20]. Vibrational 2DCoS is particularly beneficial for protein studies because it can resolve the often heavily overlapped bands in vibrational spectra into their separate components that arise from the various secondary structure motifs, amino acids, and side-chain groups. Additionally, the sequential temporal information regarding conformational changes provided by 2DCoS enables determination of structural transitions and kinetic reaction mechanisms [7].

14.2.1 Fibril nucleus formation

Lednev's group demonstrated that DUVRR combined with 2DCoS is a powerful tool for studying nucleus formation during fibrillation. Hen egg white lysozyme (HEWL), a protein chosen for its well-studied structural properties and folding/unfolding behaviors [3], was used as a model for amyloidogenic proteins. Solutions of HEWL at pH 2.0 were incubated at 65°C for various times. UVRR was measured during incubation and the resulting kinetic spectral data were analyzed using 2DCoS [6].

Only random coil and β-sheet conformations are known to give rise to the C_α−H bending region intensities observed in the 1380−1400 cm^{-1} region. Lednev et al. correlated C_α−H region band intensities with other UVRR bands [7]. Synchronous $\Phi(\nu_1,\nu_2)$ and asynchronous Ψ (ν_1,ν_2) 2D-Raman spectra were calculated using Noda's approach:

$$\phi(\nu 1,\nu 2)+i\psi(\nu 1,\nu 2)=\frac{1}{\pi(T_{\max}-T_{\min})}\times\int_0^\infty \tilde{Y}_1(\omega)\tilde{Y}_2^*(\omega)d\omega \tag{14.1}$$

where $\tilde{Y}_1(\omega)$ and $\tilde{Y}_2^*(\omega)$ are the forward and conjugate Fourier transforms, respectively, of the experimental spectral intensities $\tilde{y}(\nu,t)$ for all frequencies ν from ν_1 to ν_2, and incubation times t, during the observation period from $T_{\min}$ to $T_{\max}$ [6]. In Fig. 14.1, only the C_α−H region and the α-helix subband of the amide I region are shown for clarity.

The circled cross-peaks A, B, and C in Fig. 14.1 represent the peaks relating β-sheet to disordered structure, α-helix to β-sheet, and α-helix to disordered structure, respectively. Since the synchronous cross-peak A is positive and cross-peaks B and C are negative, β-sheet and disordered structure are positively correlated. Conversely, the negative peaks B and C imply negative correlations between α-helix and β-sheet, and α-helix and disordered structure. The asynchronous 2D-Raman correlation map (Fig. 14.1B) shows a positive peak at A, suggesting a change in the disordered structure spectral region happens first. Considering the negative synchronous peak, and the negative asynchronous peak at B, it can be deduced that spectral changes in the disordered structure spectral region happen

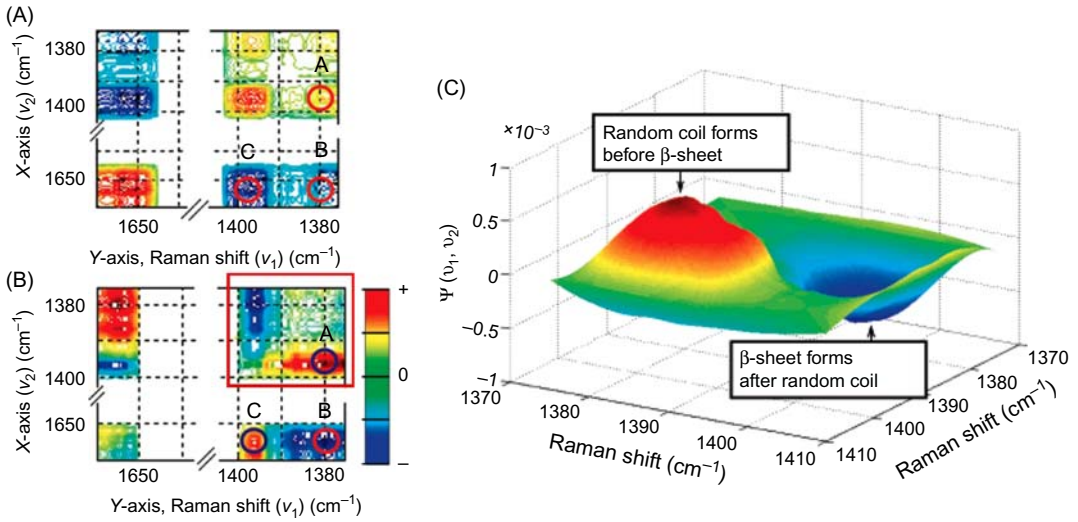


Figure 14.1

2D Correlation spectroscopy was used to establish sequential order of events during the fibrillation of hen egg white lysozyme. The synchronous (A) and asynchronous (B) 2D-deep ultraviolet resonance Raman (DUVRR) correlation maps of the C_α—H bending region and α-helix subband of the amide I region suggest the order of structural change to be: random coil→α-helix→β-sheet. The 3D perspective plot (C) of the asynchronous C_α—H bending region [*red rectangle* in (B)] illustrates the implication that a random coil forms before the β-sheet.

Source: *(A and B) Adapted with permission from V.A. Shashilov, I.K. Lednev, 2D correlation deep UV resonance Raman spectroscopy of early events of lysozyme fibrillation: kinetic mechanism and potential interpretation pitfalls, J. Am. Chem. Soc. 130 (2008) 309—317 [7]. Copyright 2008 American Chemical Society. (C) Adapted with permission from V. Shashilov, M. Xu, V.V. Ermolenkov, L. Fredriksen, I.K. Lednev, Probing a fibrillation nucleus directly by deep ultraviolet Raman spectroscopy, J. Am. Chem. Soc. 129 (2007) 6972—6973 [6]. Copyright 2007 American Chemical Society.*

before changes in the α-helix region. Similarly, the negative synchronous peak, and the positive asynchronous peak at C implies that β-sheet spectral changes occur after α-helix changes [7].

14.2.2 Apparent inverse order

It is notable that at first glance, this interpretation of the 2D correlation maps suggests an anomalous sequence of events. Logically, α-helix melting needs to occur before any disordered structure can form. This apparent paradox is resolved by noting that the asynchronous 2DCoS probes both the sequential order of events in a chemical reaction and the average rate of concentration change of each species [7]. The latter is an important fact that, if overlooked, can lead to serious misinterpretation of spectroscopic data. This phenomenon can readily be explained by considering what contributes to an asynchronous

2D map. In any chemical reaction with the formation of an intermediate, the formation of the product will be in-phase with the decrease in the reactant, and thus the intermediate should change slightly ahead to compensate for the delay in product formation [7].

14.2.3 Extraction of characteristic times for structural changes

The characteristic times for the evolution of secondary and tertiary structures were extracted from the 2D-UVRR data by means of *kv* correlation analysis. Following the approach of Shanmukh et al. [21], the asynchronous correlation of the experimental decaying intensities with a reference set of exponentially decaying intensities was calculated by:

$$\psi(\tau,\nu) = \frac{1}{N-1}\sum_{j=0}^{N-1} y(\nu,n_j)\sum_{k=0}^{N-1} N_{jk} \times \exp\left(\frac{-t}{\tau} + R\right) \tag{14.2}$$

where $\Psi(\tau,\nu)$ is the correlation intensity at a point (τ,ν); N is the total number of spectra; n_j is the order number of the spectrum; τ is the characteristic time of the exponential decay; N_{jk} is the Hilbert–Noda transformation matrix; R is the constant matrix, and $y(\nu,n)$ are the spectral intensities. After performing this analysis, the characteristic times were determined to be disordered structure formation ($\tau \sim 22$ hours), tertiary structure and α-helix melting ($\tau \sim 30$ hours), and nucleus β-sheet formation ($\tau \sim 35$ hours) [7].

14.3 Elucidating the kinetic mechanism of early events of hen egg white lysozyme fibrillation

The asynchronous 2D-Raman correlation map suggests that the mechanism is stepwise, rather than parallel; otherwise, the formation of random coil and β-sheet would be completely correlated [6]. This implies that one of the four possible two-step mechanisms below is occurring.

mechanism (1) $A \rightarrow B \rightarrow P$
mechanism (2) $A \Leftrightarrow B \rightarrow P$
mechanism (3) $A \rightarrow B \Leftrightarrow P$
mechanism (4) $A \Leftrightarrow B \Leftrightarrow P$

It has been shown that the formation of a partially unfolded intermediate from native HEWL is an irreversible transition [22]. This eliminates mechanisms (2) and (4) as possibilities. An irreversible conversion of B to P [mechanisms (1) and (2)] would appear in the 2D-Raman correlation map as a negative correlation between A and B at the beginning of the reaction, followed by a positive correlation between A and B for the remainder of the reaction. However, a reversible transition from B to P [mechanisms (3) and (4)] would lead

to a negative correlation between A and B at all stages of the reaction, which was the case for HEWL. Considering these details, only mechanism (3) remains, suggesting that the actual reaction mechanism is:

$$\alpha\text{-}helix \xrightarrow{k_1} \text{disordered structure} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \beta\text{-sheet} \qquad \text{(model 14.1)}$$

The rate constants k_1, k_2, and k_{-2} were determined by using the experimentally determined [6] characteristic times for the mechanism:

$$\text{native} \xrightarrow{K_1} \text{PU} \underset{K_{-2}}{\overset{K_2}{\rightleftharpoons}} \text{nucleus} \qquad \text{(model 14.2)}$$

Relating the integrated rate laws for these two mechanisms yielded the following relations between the rate constants k_1, k_2, and k_{-2} and K_1, K_2, and K_{-2}:

$$\begin{aligned} k_1 &= K_1 \\ k_2 &= \frac{K_2}{1-R} \\ k_{-2} &= K_{-2} - \frac{R}{1-R} \times K_2 \end{aligned} \qquad (14.3)$$

where R is the fraction of disordered structure in native HEWL (60% [23] in this case). Taking K_i to be the reciprocal of the characteristic time τ_{Ki}, one obtains $K_1 = 0.033\ \text{h}^{-1}$, $K_2 = 0.0125\ \text{h}^{-1}$, $K_{-2} = 0.067\ \text{h}^{-1}$, and thus τ_{k1}, τ_{k2}, and τ_{k-2} are 30, 32, and 21 hours, respectively [7].

A simulation-aided approach was utilized for model testing and the results proved to be in close agreement with experimental results. The rate constants k_1, k_2, and k_{-2} of model (14.1) were extracted by simulating different sets of rate constants k_i to calculate the evolution profiles of pure secondary structure. Using spectral profiles of the pure secondary structures, modeled as mixed Gaussian−Lorentzian shapes, and the simulated evolution profiles, *kv* correlation maps were constructed as described above, and compared to the experimental *kv* correlation maps. When the characteristic times were extracted from the best-matched *kv* correlation map, as described above, and converted to rate constants K_i, the results were identical to the experimentally derived rate constants [7].

These data are consistent with the step-by-step mechanism proposed by Dobson and coworkers [24]:

$$\text{Native} \rightarrow \text{PU} \Leftrightarrow \beta\text{-sheet} \qquad \text{(model 14.3)}$$

Seeding experiments were performed to further test this interpretation. The supernatant of a HEWL sample that had been incubated for 48 hours was used to seed the fibrillation of

fresh HEWL. The effective elimination of the fibrillation lag-phase observed is in complete agreement with the proposed mechanism [6].

14.4 Structural characterization of fibrillar proteins

14.4.1 Hydrogen deuterium exchange

Lendev and coauthors have combined hydrogen−deuterium exchange (HX) with UVRR to probe the secondary structure of the fibril cross-β core. HX is such a useful technique for characterizing protein structure because in an amino acid residue, only the main chain NH group and O, N, and S bound protons can be exchanged easily, while the carbon-bound protons generally do not exchange [25]. This, combined with the fact that strongly hydrogen-bonded secondary structures greatly reduce the rate of HX [26], allows for a particularly clever use of HX to probe the fibril core. It was hypothesized that the tightly packed hydrophobic core of amyloid fibrils would experience a drastically reduced HX rate compared to the amide N−H protons in the unordered fragments [10]. Mikhonin and Asher [27] have shown that HX will cause a downshift of the amide II DUVRR band from $\sim$1555 to $\sim$1450 cm^{-1} (amide II′) along with the virtual disappearance of the amide III band in an unordered protein [12]. Fig. 14.2 illustrates these spectral changes in DUVRR spectra of fibrillar and unordered lysozyme upon deuteration.

14.4.2 Bayesian source separation

To deduce structural information about the fibril core, the overlapping spectral bands of the deuterated unordered fractions and the still-protonated cross-β core need to be resolved. The Bayesian approach [28] was utilized, instead of a blind source separation algorithm such as ICA or pure variable methods, because it is uniquely suited to solve extremely ill-suited MCR and prior-dominated problems.

14.4.3 Structure determination of the lysozyme fibril core

Lednev et al. combined DUVRR spectroscopy and HX exchange to probe the structure of HEWL fibrils [12]. The Bayesian signal dictionary approach was used to incorporate a priori information about characteristic spectral bands. In this work, the reference band library contained mixed Gaussian and Lorentzian shapes of certain widths and spectral positions obtained by fitting the DUVRR spectra of fibrils with Gaussian/Lorentzian peaks having a predefined width range. Prior information regarding the experimental spectra of deuterated lysozyme and reference spectrum for the cross-β-sheet were incorporated by augmenting the matrices with an additional row. The final step is accomplished by using the floating point genetic algorithm to determine the remaining parameters [10].

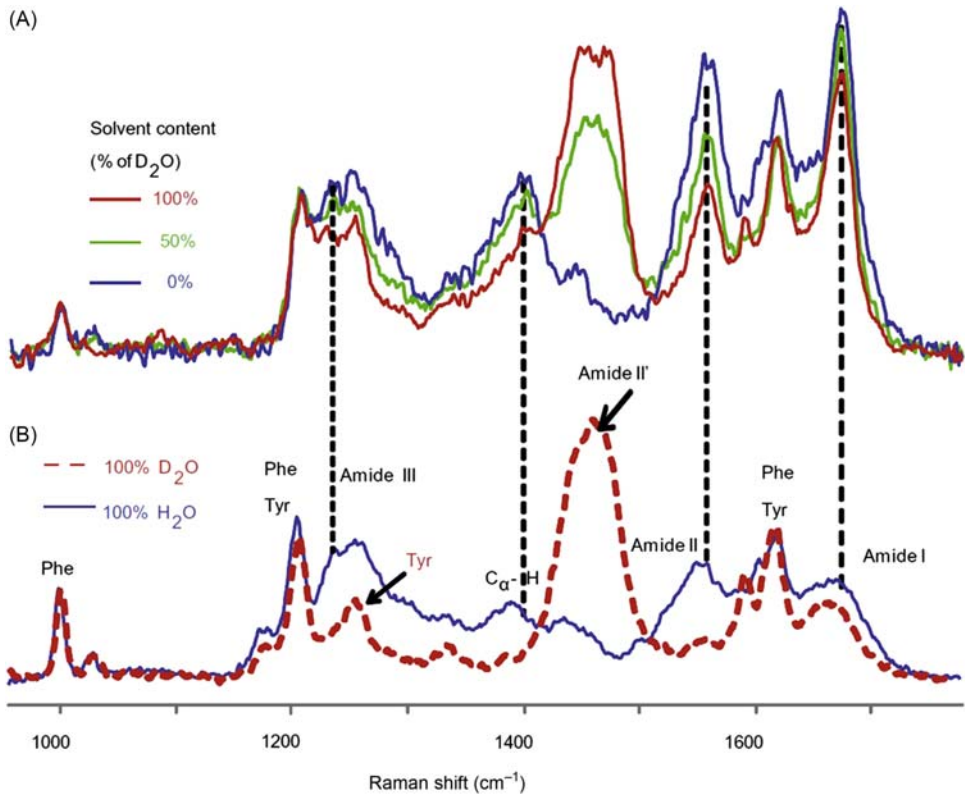


Figure 14.2

Deep ultraviolet resonance Raman (DUVRR) spectral changes in fibrillar and unordered hen egg white lysozyme after hydrogen–deuterium exchange provide evidence of a still-protonated core. (A) DUVRR spectra of hen egg white lysozyme (HEWL) fibrils in H_2O (*blue*), D_2O (*red*), and 50/50% H_2O/D_2O mixture (*green*). (B) DUVRR spectrum of unordered lysozyme in H_2O (*blue*), and D_2O (*red*). Source: *Reprinted with permission from V.A. Shashilov, V. Sikirzhytski, L.A. Popova, I.K. Lednev, Quantitative methods for structural characterization of proteins based on deep UV resonance Raman spectroscopy, Methods (San. Diego, Calif.) 52 (2010) 23–37 [10]. Copyright 2010 Elsevier.*

By extracting the DUVRR spectrum of the fibril core, the exact position of the amide III_3 (β) band could be determined, and Asher group's method was used to calculate the Ψ dihedral angle distribution [12]. The spectrum of the cross-β-core of HEWL fibrils extracted by this method is shown in Fig. 14.3. The distribution of the Ψ dihedral angle obtained from the Raman signature of the fibril core provides an important characteristic of the polypeptide conformation and should allow for differentiating different types of cross-β core in amyloid fibrils.

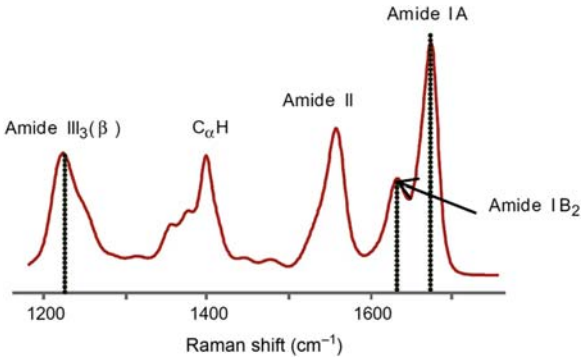


Figure 14.3
Deep ultraviolet resonance Raman spectrum of the cross-β core of hen egg white lysozyme extracted using Bayesian source separation. Source: *Reprinted with permission from V.A. Shashilov, V. Sikirzhytski, L.A. Popova, I.K. Lednev, Quantitative methods for structural characterization of proteins based on deep UV resonance Raman spectroscopy, Methods (San. Diego, Calif.) 52 (2010) 23–37 [10]. Copyright 2010 Elsevier.*

14.4.4 Structural variations in the cross-β core

14.4.4.1 genetically engineered polypeptides

Lednev and coworkers utilized relatively simple polypeptides to mimic the fibril formation of globular proteins. Several synthetic polypeptides exhibiting typical fibrillogenic properties have been designed and synthesized by using genetic engineering, allowing for varying the polypeptide sequence to probe the sequence specific mechanism of fibrillation [29–33]. These large polypeptides are soluble under close to physiological conditions, but form fibrils under acidic pH. They exhibit a lag phase and show accelerated fibrillation on seeding or agitation.

A 687-residue polypeptide consisting of 32 amino acid repeats, $GH_6[(GA)_3GY(GA)_3GE(GA)_3GH(GA)_3GK]_{21}GAH_6$ (YEHK21) was found to form fibrils at room temperature [33]. The formation of these fibrils was thermoreversible. DUVRR spectra were recorded at various temperatures ranging from 25°C to 90°C, after reaching equilibrium. The most significant spectral changes occurred at temperatures greater than 80°C. Fig. 14.4 shows the Raman spectra of YEHK21 in solution measured at 25°C and 90°C. The spectra demonstrate changes consistent with the melting of a β-sheet structure into an unordered conformation. Spectra measured at various times after heating were virtually the same, indicating a slow recovery process [33].

The analysis of polypeptide secondary structure content was complicated by the temperature dependence of the basis spectra. This was overcome by exploiting the slow

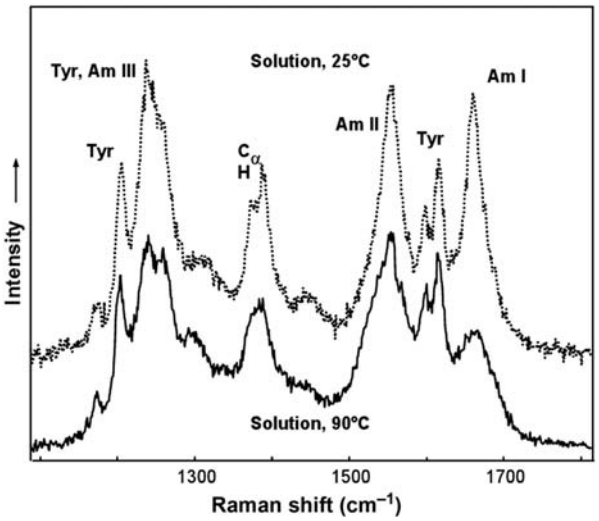


Figure 14.4

Deep ultraviolet resonance Raman (DUVRR) spectra of a genetically engineered polypeptide (YEHK21) heated to various temperatures were recorded to detect spectral changes resulting from β-sheet melting. DUVRR spectra of YEHK21 in solution at 25°C and 90°C, illustrating spectral differences characteristic of β-sheet melting. Source*: Reprinted with permission from I.K. Lednev, V.V. Ermolenkov, S. Higashiya, L.A. Popova, N.I. Topilina, J.T. Welch, Reversible thermal denaturation of a 60-kDa genetically engineered beta-sheet polypeptide, Biophys. J. 91 (2006) 3805–3818 [33]. Copyright 2006 Elsevier.*

recovery process. DUVRR spectra were recorded after 5 minutes of heating to various temperatures ranging from 25°C to 125°C, and subsequent rapid cooling to room temperature. In this way, all measurements were taken at the same temperature, ensuring any spectral changes were a result of temperature-induced conformational changes. This spectral dataset was used for understanding YEHK21 polypeptide melting [33].

A melting curve was constructed using a basis set of three DUVRR spectra: the YEHK21 backbone Raman spectra (initial sample and sample treated at 125°C) and the DUVRR spectrum of tyrosine. The first two spectra represent completely folded YEHK21 (mostly β-sheet and turn conformation), and the unordered YEHK21 (mainly polyproline II conformation), while tyrosine was included for use as an internal standard. These three basis spectra were used to fit the Raman spectra of YEHK21 at various temperatures. This showed that the melting process could be considered as a two-state process with each spectrum described with different proportions of β-sheet and turn/polyproline II conformations. The percentage of folded contribution was plotted versus temperature to create the melting curve of Fig. 14.5 [33].

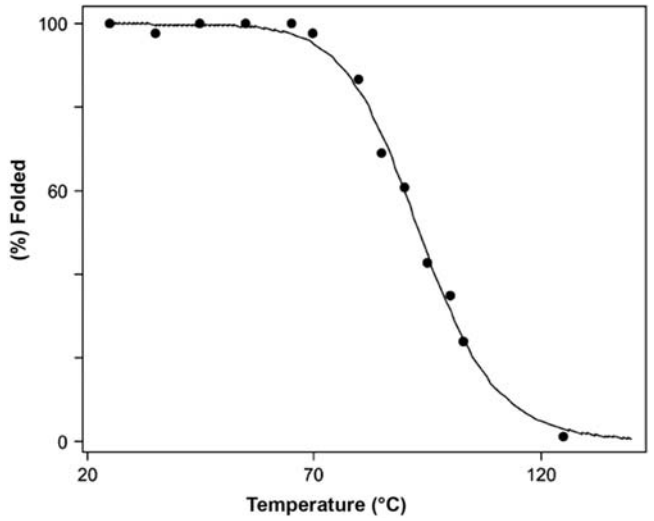


Figure 14.5

A YEHK21 melting curve was constructed by plotting the fraction of folded peptide against temperature (*solid circles*). The solid line represents the best fit for Eq. (14.4). Source*: Reprinted with permission from I.K. Lednev, V.V. Ermolenkov, S. Higashiya, L.A. Popova, N.I. Topilina, J.T. Welch, Reversible thermal denaturation of a 60-kDa genetically engineered beta-sheet polypeptide, Biophys. J. 91 (2006) 3805–3818 [33]. Copyright 2006 Elsevier.*

Thermodynamic parameters were estimated by fitting the data to:

$$f_{\mathrm{F}} = \frac{1}{1 + e^{-\left(\frac{\Delta H}{RT} - \frac{\Delta S}{R}\right)}} \tag{14.4}$$

where f_{F} is the fraction of folded polypeptide after treatment at temperature T and R is the gas constant. The parameters ΔH and ΔS were determined to be 135 ± 14 kJ/mol and 370 ± 40 J/(mol K), where the standard error was represented by $\pm 2\sigma$. This represents a first approximation because Eq. (14.4) is typically applied to two-state processes where ΔCp, the heat capacity change, has negligible effect on the enthalpy of folding over the narrow temperature range of the transition [34].

Temperature jump experiments were performed to analyze the kinetics of YEKH21 β-sheet melting. The DUVRR spectra of YEHK21 were collected every 20 seconds after a temperature jump from 25°C to 85°C. The results of the temperature jump experiments were similar to the steady-state measurements. It was also found that recovery of the YEHK21 Raman spectrum was dependent on the temperature to which it was heated before being cooled to room temperature. Complete recovery of the YEHK21 Raman spectrum was observed after approximately 6 hours of standing at room temperature after being briefly heated to 85°C. On the other hand, only 80% recovery was achieved after a

5-minute heating to 100°C before sitting at room temperature for 4 hours. The kinetics for the recovery of the DUVRR spectrum after heating was determined to be roughly monoexponential with a characteristic time of approximately 1 hour [33].

A study of three similarly engineered polypeptides that form amyloid-like fibrils provided the first direct evidence that the structure of the cross-β core is sequence-dependent. The three polypeptides consisted of seven repeats of 32, 40, or 48 amino acid repeats, $(GA)_nGY(GA)_nGE(GA)_nGH(GA)_nGK$ ($n = 3$, 4, or 5, respectively, 32YEHK7, 40YEHK7, or 48YEHK7). These β-sheet-forming repetitive polypeptide blocks and selected amino acids at turn positions facilitated the formation of the targeted amyloid-like fibrils [11]. These polypeptides were designed to have the same turn structure so any pair-wise subtraction would result in the pure resonance Raman spectra of the fibrillar β-sheet core [10]. The validity of this hypothesis is verified by inspecting the difference spectrum (Fig. 14.6), where they are virtually identical within experimental error [11].

The studies above are the first direct measurements of the β-turn spectroscopic signature. Pure variable methods [11] were used to analyze the spectra of the YEHK fibrils to resolve the contributions from the β-sheet and β-turn portions. The spectra of the two components were calculated and the experimental spectra of the YEHK fibrils were able to be fit with a linear combination of the two components (Fig. 14.6). Using the Raman signature of the cross-β core, the dominant amide III peak was found to be 1241 cm^{-1}, corresponding to a Ψ angle of 150 degrees and an antiparallel β-sheet conformation. The narrow

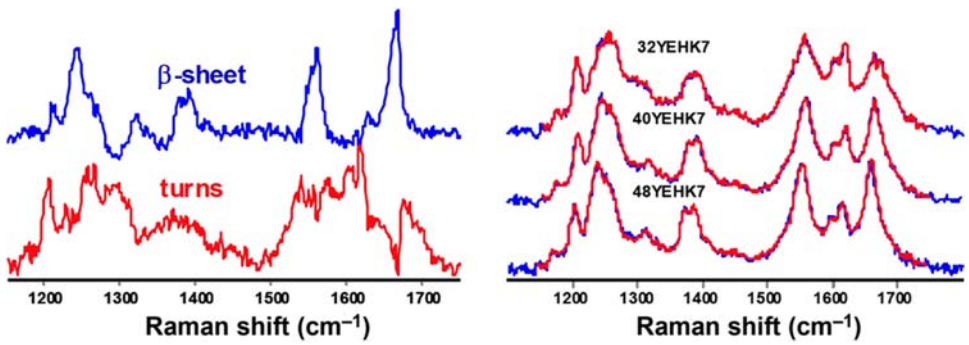


Figure 14.6

The deep ultraviolet resonance Raman (DUVRR) spectral signatures of YEHK polypeptide β-sheet and turns fit the experimental spectra with close agreement, indicating that these peptides are composed primarily of these two structures. DUVRR spectra of β-sheet and turns (left), and the DUVRR spectra of YEHK fibrils fitted with a linear combination of these two components (right). Source: *Adapted with permission from V. Sikirzhytski, N.I. Topilina, S. Higashiya, J.T. Welch, I.K. Lednev, Genetic engineering combined with deep UV resonance Raman spectroscopy for structural characterization of amyloid-like fibrils, J. Am. Chem. Soc. 130 (2008) 5852–5853 [11]. Copyright 2008 American Chemical Society.*

Ramachandran Ψ angular distribution of YEHK fibrils suggests a highly ordered β-sheet [11], consistent with the findings reported on lysozyme fibrils [12]. Notably, however, the UVRR spectra of the cross-β core of YEHK and lysozyme were different, indicating a sequence dependence of the core β-structure [11].

14.4.4.2 Amyloid β

The application of the HX-DUVRR method to amyloid β (Aβ) fibrils revealed that parallel and antiparallel β-sheets have different distributions of Ψ dihedral angles. Full-length unlabeled $A\beta_{1-40}$ peptide and $A\beta_{32-42}$ fragment fibrils were subjected to postfibrillation HX, UVRR spectra of the fibrillar core extracted, and the Ψ dihedral angle distributions calculated. It was shown that $A\beta_{1-40}$ peptide exhibited a parallel β-sheet conformation, atypical for globular proteins, and that $A\beta_{32-42}$ fragment fibrils possessed the more typical antiparallel conformation [5]. The availability of the three-dimensional structure of the $A\beta_{1-40}$ fibrils obtained by solid-state NMR [35] was used to verify the applicability of the Asher group's approach [36] for the quantitative characterization of peptide conformation in the fibril core using DUVRR spectroscopy.

Using this structural information, the amide III region of the DUVRR spectra of the $A\beta_{1-40}$ fibril core was calculated by using the following method, and compared to experimental results. The contributions of each peptide residue in the β-strand that was resistant to HX were calculated using a set of residue-specific Ψ angles reported for the $A\beta_{1-40}$ fibril core [35] as Gaussian curves with a half-width at half-maximum of 13 cm^{-1}. As shown in Fig. 14.7B, the calculated spectrum (blue) is in excellent agreement with the experimental spectrum (black). It is notable that there is a significant difference in the spectral profile of the $A\beta_{1-40}$ fibril core (Fig. 14.7B), and those of the $A\beta_{32-42}$ (Fig. 14.7A), hamster prion protein, and YEHK peptide (Fig. 14.7C). The major difference between the calculated amide III band region and the experimentally measured one is the appearance of two bands at 1280 and 1250 cm^{-1}. These bands, however, are also present in the spectrum of $A\beta_{32-42}$, and are outside the β-sheet region; indicating that these peaks are most likely not part of the amide III_3 vibration and do not interfere with the analysis.

14.4.5 Aromatic side chain as a reporter of local environment

Xu et al [3]. assessed phenylalanine's ability to act as an intrinsic DUVRR spectroscopic probe of local environment. A 1.17-mM *N*-acetyl-L-phenylalanine ethyl ester (ac-phe-ee) solution in acetonitrile was diluted in a series of acetonitrile-water solutions with a final acetonitrile content ranging from 5% to 100% and DUVRR spectra were recorded. As seen in Fig. 14.8, the spectra comprised two strong peaks corresponding to acetonitrile, and several prominent phenylalanine bands. The band at 1000 cm^{-1}, derived from phenyl ring stretching, exhibited the strongest change in Raman cross section as the concentration of

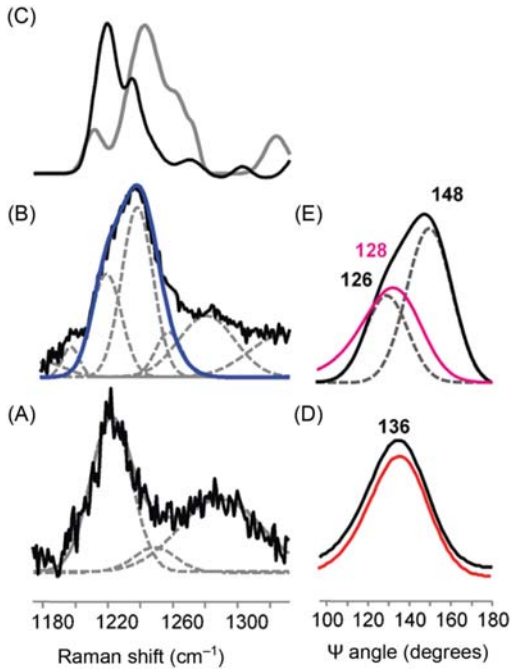


Figure 14.7

Solid-state NMR data of $A\beta_{1-40}$ fibrils were used to verify the applicability of the Asher group's Ψ distribution method to quantitatively characterize the fibril core conformation of peptides using deep ultraviolet resonance Raman (DUVRR) spectroscopy. (A) DUVRR spectra of the $A\beta_{34-42}$ fibril core (*solid line*) and the best fit with Gaussian peaks (*dotted lines*); (B) DUVRR spectra of the $A\beta_{1-40}$ fibril core (*solid line*) and the best fit with Gaussian peaks (*dotted lines*) and the calculated spectrum based on SSNMR data (*blue line*); (C) hamster prion protein (*black*) and YEHK peptide (*gray*); (D) Ψ angle distribution in the $A\beta_{34-42}$ fibril core (*black*) and average antiparallel β-sheets in globular proteins; (E) calculated Ψ angle distribution in the $A\beta_{1-40}$ fibril core (*black*), the best fit with Gaussian curves (*dashed lines*), and the Ψ angle distribution in average parallel β-sheets in globular proteins (*pink*). Source*: Adapted with permission from L.A. Popova, R. Kodali, R. Wetzel, I.K. Lednev, Structural variations in the cross-beta core of amyloid beta fibrils revealed by deep UV resonance Raman spectroscopy, J. Am. Chem. Soc. 132 (2010) 6324–6328 [5]. Copyright 2010 American Chemical Society.*

acetonitrile varied. The inset of Fig. 14.8 shows that the Raman cross section of this band increases monotonically, reaching a plateau at approximately 50% acetonitrile. Practically, this indicates that the 1000 cm^{-1} band of phenylalanine is sensitive to water exposure, making it a probe of protein local environment and tertiary structure.

In the study of HEWL, DUVRR spectra were recorded at various stages of fibrillation, analyzing the supernatant and gelatinous phase separately (Fig. 14.9). The spectra were dominated by contributions of the amide chromophore and phenylalanine. There was a

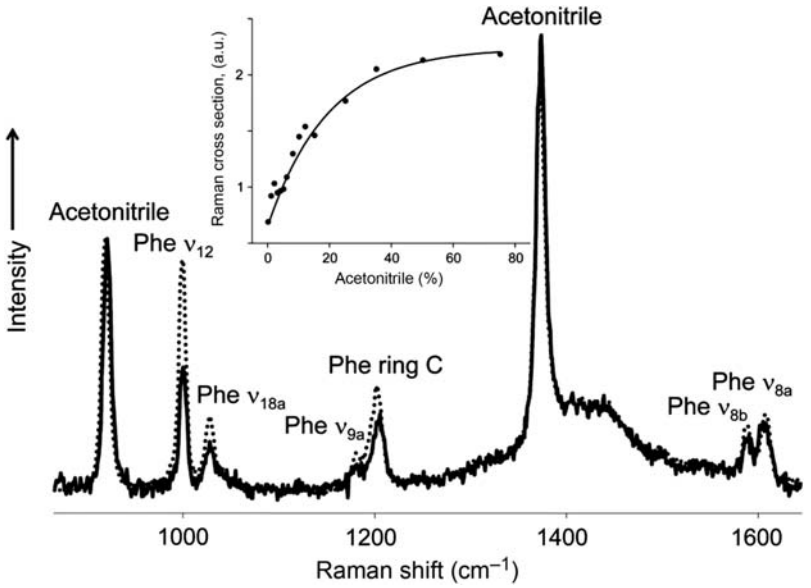


Figure 14.8

The Raman cross-section of the 1000 cm^{-1} band phenylalanine band was shown to vary as a function of water exposure, making it a useful probe of local environment. Deep ultraviolet resonance Raman spectra of *N*-acetyl-L-phenylalanine ethyl ester in water and 50% acetonitrile (*dotted line*) and 5% acetonitrile (*solid line*). Inset: Raman cross section of the 1000 cm^{-1} band as a function of acetonitrile concentration. Source*: Reprinted with permission from V.A. Shashilov, I.K. Lednev, 2D correlation deep UV resonance Raman spectroscopy of early events of lysozyme fibrillation: kinetic mechanism and potential interpretation pitfalls, J. Am. Chem. Soc. 130 (2008) 309–317 [7]. Copyright 2008 John Wiley and Sons.*

dramatic reduction in the 1000 cm^{-1} phenylalanine band intensity with increased incubation time. Fig. 14.9 inset shows the monoexponential decrease with a characteristic time of 29 ± 2 hours. This indicates that phenylalanine is increasingly exposed to water as fibrillation progresses [7].

JiJi and coworkers also used DUVRR spectroscopy to study amyloid-β (Aβ) peptide fibrillation [37]. They observed that the intensities of phe and tyr bands increase upon addition of myricetin which inhibits Aβ peptide fibrillation. They concluded that myricetin interacts with these aromatic amino acid side chains and prohibits fibrillation of Aβ peptide.

14.5 Conformation changes of amyloid fibrils

Amyloid fibrils prepared from full-length proteins are typically very stable and considered to be the most thermodynamically stable form of proteins because of the high resistance to thermal and solvent-induced denaturation [38,39]. Very harsh denaturating conditions such

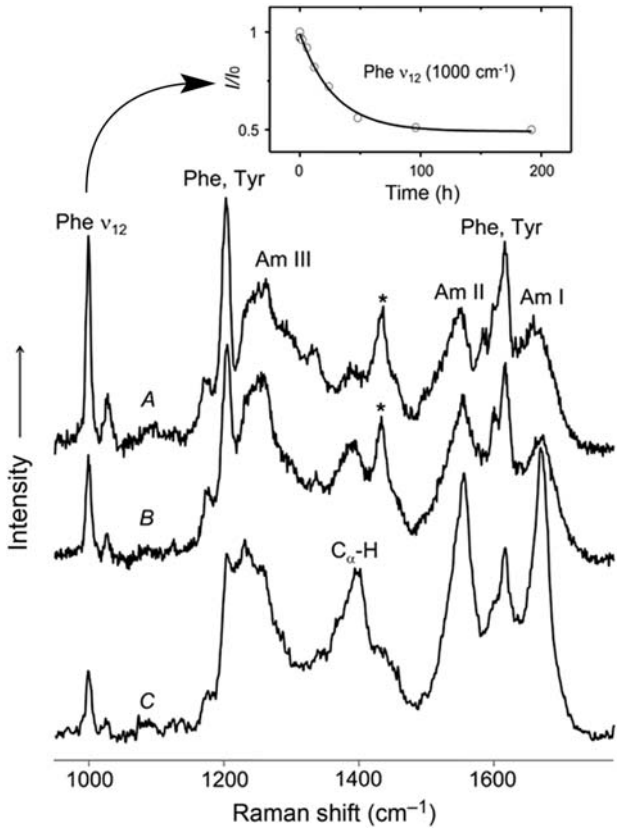


Figure 14.9

Deep ultraviolet resonance Raman spectra of hen egg white lysozyme were recorded at various times of incubation. Not incubated (A), soluble phase after 2 days of incubation (B), and the gelatinous phase (C). * is the internal standard peak. Source*: Adapted with permission from V.A. Shashilov, I.K. Lednev, 2D correlation deep UV resonance Raman spectroscopy of early events of lysozyme fibrillation: kinetic mechanism and potential interpretation pitfalls, J. Am. Chem. Soc. 130 (2008) 309–317 [7]. Copyright 2008 John Wiley and Sons.*

as high pH, low temperature, and high pressure are required to disintegrate mature fibrils [40,41], and no newly formed fibrils are expected to appear under such conditions. The ability of mature fibrils to change their structure and morphology has not been questioned, to the best of our knowledge, because no conventional methods would allow its detection without special sample preparation.

14.5.1 Spontaneous refolding of amyloid fibrils from one polymorph to another

Kurouski et al. [42] utilized DUVRR spectroscopy to monitor the stability of apo-α-lactalbumin fibrils during the preparation of samples for atomic force microscopic (AFM)

characterization. UVRR spectroscopy was chosen since it does not require any sample preparation. Surprisingly, melting and refolding of β-sheet was evident in the Raman spectra measured for mature apo-α-lactalbumin fibrils after the temperature was dropped from 37°C to 25°C and the solution was desalted.

Later, DUVRR spectroscopy was also used to characterize the initial (Polymorph I) and final (Polymorph II) polymorphs [42]. The HX combined with DUVRR spectroscopy (see above for details) indicated that the amide III_3 region of the Polymorph I spectrum shows two strong peaks at 1226 and 1253 cm^{-1}, corresponding to Ψ dihedral angles of 134.5 and 163.5 degrees (Fig. 14.10). Polymorph II, however, only exhibited a peak at 1253 cm^{-1} suggesting only a single β-sheet form is present, clearly indicating a difference in the structure of the fibril cores [43]. This result showed that changes in the fibril cross-β core correlate with noticeable changes in fibril morphology, indicating that apo-α-lactalbumin fibril polymorphism is correlated by a specific organization of the fibril core. The application of DUVRR spectroscopy provided a unique opportunity for detecting the spontaneous transition between the two fibril polymorphs, observations that were not easily evident by using other conventional biophysical methods [42]. Spontaneous refolding of mature fibrils from one polymorph to another turned out to be a general phenomenon and was demonstrated for insulin [44] and HET-s (218 − 289) prion fibrils [45] triggered by the pH change

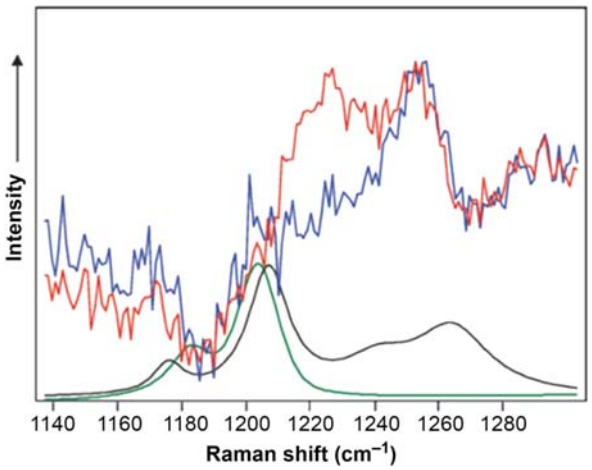


Figure 14.10

Amide III Raman band region of the protonated fibril core of Polymorphs I (*red*) and II (*blue*) in D_2O. The deuterated part of the protein does not contribute to this spectral region. The contribution of aromatic amino acids is quantitatively removed by subtracting the spectra of phenylalanine (*green*) and tyrosine (*black*). Source: *Adapted with permission from D. Kurouski, W. Lauro, I.K. Lednev, Amyloid fibrils are "alive": spontaneous refolding from one polymorph to another, Chem. Commun. 46 (2010) 4249–4251 [42]. Copyright 2010 Royal Society of Chemistry.*

14.5.2 Purple fibrils and a new protein chromophore

Lednev et al. studied two polymorphs of HEWL fibrils formed simultaneously in a solution of 20% acetic acid and 0.2 M NaCl solution that naturally separate into two colored phases—purple and white—after 48 hours of incubation at pH 2.0. Each layer was found to contain fibrils, which when examined by AFM, showed different morphologies. When subjected to a shift in pH, the fibrils from the purple layer (PF) disintegrated, but when the resulting aggregates were centrifuged and resuspended in the initial conditions, the PF reformed. This was shown to be repeatable and led to the hypothesis that a unique fibril polymorphic structural arrangement of HEWL is responsible for the appearance of the purple chromophore [46].

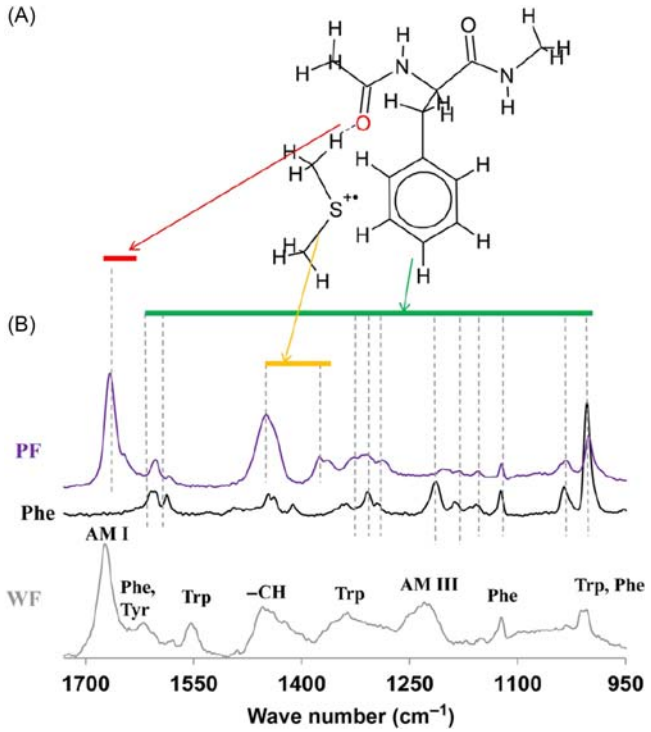


Figure 14.11

(A) Schematic representation of the S:./pi bond model. (B) Resonance Raman spectrum of purple fibrils (PF) and nonresonance Raman spectrum of white fibrils (WF). Source: *(A) Adapted from X. Chen, Y. Tao, J. Li, H. Dai, W. Sun, X. Huang, et al., Aromatic residues regulating electron relay ability of S-Containing amino acids by formations of S∴ π multicenter three-electron bonds in proteins, J. Phys. Chem. C 116 (2012) 19682−19688 [48]. (B) Adapted with permission from T. Quiñones-Ruiz, M.F. Rosario-Alomar, K. Ruiz-Esteves, M. Shanmugasundaram, V. Grigoryants, C. Scholes, et al., Purple fibrils: a new type of protein chromophore, J. Am. Chem. Soc. 139 (2017) 9755−9758 [46]. Copyright 2010 American Chemical Society.*

When investigating these fibrils using DUVRR spectroscopy, an excitation of 514.5 nm was used because it fell within the 550-nm absorption band of the PF suspension, providing a resonance enhancement of the Raman-scattering chromophore. While these spectra were all made up of the Am I, Am II, and the aromatic amino acid sidechains, the resonance-enhanced spectrum of PF was found to be different from the off-resonance spectra (785 nm excitation) of the PF and the white fibrils. The enhanced spectrum of the PF contained peaks assigned to Phe vibrational modes [46]. Additionally, there were bands at 1451 and 1377 cm^{-1} not assignable to Phe, which were tentatively assigned to CH_2 bending modes of Met [47]. The Am I band, which primarily results from C = O stretching, of the resonance spectrum has a different position from the off-resonance spectrum, indicating that the carbonyls contribute to the resonance enhancement (Fig. 14.11). Together, this evidence suggests that Phe and Met side chains with a carbonyl group comprise the purple chromophore, which is in good agreement with electron paramagnetic resonance spectroscopic results [46].

14.6 Conclusions

UVRR is a powerful tool for the study of biomolecular structure and dynamics including the secondary structure conformational analysis at the level of individual peptide bonds in large proteins. DUVRR is uniquely suitable for structural characterization of proteins at all stages of amyloid fibril formation. The combination of HX with DUVRR spectroscopy allows for probing the variations in cross-β core structure of fibril polymorphs. The application of advanced statistics, chemometrics, and 2DCoS in particular for DUVRR spectra analysis significantly increases the quality and amount of information on protein structure and dynamics. This approach is especially useful for establishing the kinetic mechanisms of amyloid fibrillation.

References

[1] I. Lednev, Vibrational Spectroscopy: Biological Applications of Ultraviolet Raman Spectroscopy, Nova Science Publishers, Inc, New York, 2007, pp. 1–26.
[2] I.K. Lednev, V. Shashilov, M. Xu, Ultraviolet Raman spectroscopy is uniquely suitable for studying amyloid diseases, Curr. Sci. 97 (2009) 180–185.
[3] M. Xu, V.V. Ermolenkov, V.N. Uversky, I.K. Lednev, Hen egg white lysozyme fibrillation: a deep-UV resonance Raman spectroscopic study, J. Biophoton. 1 (2008) 215–229.
[4] I. Lednev, D. Kurouski, W. Lauro, Amyloid fibrils are “alive”: spontaneous refolding from one polymorph to another, Chem. Comm. 46 (2010) 4249–4251.
[5] L.A. Popova, R. Kodali, R. Wetzel, I.K. Lednev, Structural variations in the cross-beta core of amyloid beta fibrils revealed by deep UV resonance Raman spectroscopy, J. Am. Chem. Soc. 132 (2010) 6324–6328.
[6] V. Shashilov, M. Xu, V.V. Ermolenkov, L. Fredriksen, I.K. Lednev, Probing a fibrillation nucleus directly by deep ultraviolet Raman spectroscopy, J. Am. Chem. Soc. 129 (2007) 6972–6973.

[7] V.A. Shashilov, I.K. Lednev, 2D correlation deep UV resonance Raman spectroscopy of early events of lysozyme fibrillation: kinetic mechanism and potential interpretation pitfalls, J. Am. Chem. Soc. 130 (2008) 309–317.
[8] V.A. Shashilov, I.K. Lednev, Two-dimensional correlation Raman spectroscopy for characterizing protein structure and dynamics, J. Raman Spectrosc. 40 (2009) 1749–1758.
[9] V.A. Shashilov, I.K. Lednev, Advanced statistical and numerical methods for spectroscopic characterization of protein structural evolution, Chem. Rev. 110 (2010) 5692–5713.
[10] V.A. Shashilov, V. Sikirzhytski, L.A. Popova, I.K. Lednev, Quantitative methods for structural characterization of proteins based on deep UV resonance Raman spectroscopy, Methods (San. Diego, Calif.) 52 (2010) 23–37.
[11] V. Sikirzhytski, N.I. Topilina, S. Higashiya, J.T. Welch, I.K. Lednev, Genetic engineering combined with deep UV resonance Raman spectroscopy for structural characterization of amyloid-like fibrils, J. Am. Chem. Soc. 130 (2008) 5852–5853.
[12] M. Xu, V. Shashilov, I.K. Lednev, Probing the cross-b core structure of amyloid fibrils by hydrogen-deuterium exchange deep ultraviolet resonance Raman spectroscopy, J. Am. Chem. Soc. 129 (2007) 11002–11003.
[13] R.G. Brereton, Chemometrics: Data Analysis for the Laboratory and Chemical Plant, Wiley, Chinchester (England), 2003.
[14] A. Hyvarinen, J. Karhunen, E. Oja, What Is Independent Component Analysis? John Wiley & Sons Inc, 2002, pp. 145–164.
[15] A. Hyvärinen, P.O. Hoyer, M. Inki, Topographic inpedendent component analysis: visualizing the dependence structure, Neural Comput. 13 (2001) 1527–1558.
[16] X. Shao, G. Wang, S. Wang, Q. Su, Extraction of mass spectra and chromatographic profiles from overlapping GC/MS signal with background, Anal. Chem. 76 (2004) 5143–5148.
[17] A. Cichocki, S.-i Amari, Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications, J. Wiley & Sons, Chichester, 2002.
[18] A.J. Bell, T.J. Sejnowski, An information-maximization approach to blind separation and blind deconvolution, Neural Comput. 7 (1995) 1129–1159.
[19] A. Hyvärinen, E. Oja, Independent component analysis: algorithms and applications, Neural Netw. 13 (2000) 411–430.
[20] I. Noda, Y. Ozaki, Two-Dimensional Correlation Spectroscopy: Applications in Vibrational and Optical Spectroscopy, John Wiley & Sons, Chichester; Hoboken, NJ, 2004.
[21] S. Shanmukh, R.A. Dluhy, kv Correlation analysis. A quantitative two-dimensional IR correlation method for analysis of rate processes with exponential functions, J. Phys. Chem. A 108 (2004) 5625–5634.
[22] M. Xu, V.A. Shashilov, V.V. Ermolenkov, L. Fredriksen, D. Zagorevski, I.K. Lednev, The first step of hen egg white lysozyme fibrillation, irreversible partial unfolding, is a two-state transition, Protein Sci. 16 (2007) 815–832.
[23] P.J. Artymiuk, C.C.F. Blake, D.W. Rice, K.S. Wilson, The structures of the monoclinic and orthorhombic forms of hen egg-white lysozyme At 6-A resolution, Acta Crystallogr. Sect. B-Struct. Sci. 38 (1982) 778–783.
[24] D.R. Booth, M. Sunde, V. Bellotti, C.V. Robinson, W.L. Hutchinson, P.E. Fraser, et al., Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrillogenesis, Nature 385 (1997) 787–793.
[25] S.W. Englander, T.R. Sosnick, J.J. Englander, L. Mayne, Mechanisms and uses of hydrogen exchange, Curr. Opin. Struct. Biol. 6 (1996) 18–23.
[26] L.P. DeFlores, A. Tokmakoff, Water penetration into protein secondary structure revealed by hydrogen-deuterium exchange two-dimensional infrared spectroscopy, J. Am. Chem. Soc. 128 (2006) 16520–16521.
[27] A.V. Mikhonin, S.A. Asher, Uncoupled peptide bond vibrations in alpha-helical and polyproline II conformations of polyalanine peptides, J. Phys. Chem. B 109 (2005) 3047–3052.

[28] K.H. Knuth, Bayesian Source Separation and Localization, 1998, pp. 147–158.
[29] N.I. Topilina, V.V. Ermolenkov, S. Higashiya, J.T. Welch, I.K. Lednev, Beta-sheet folding of 11-kDa fibrillogenic polypeptide is completely aggregation driven, Biopolymers 86 (2007) 261–264.
[30] N.I. Topilina, V.V. Ermolenkov, V. Sikirzhytski, S. Higashiya, I.K. Lednev, J.T. Welch, et al., Designed 11 kDa polypeptide: model for amyloidogenic intrinsically disordered proteins, Biopolymers 93 (2010) 607–618.
[31] N.I. Topilina, S. Higashiya, N. Rana, V.V. Ermolenkov, C. Kossow, A. Carlsen, et al., Bilayer fibril formation by genetically engineered polypeptides: preparation and characterization, Biomacromolecules 7 (2006) 1104–1111.
[32] N.I. Topilina, V. Sikirzhytsky, S. Higashiya, V.V. Ermolenkov, I.K. Lednev, J.T. Welch, Charge distribution and amyloid fibril formation: insights from genetically engineered model systems, Biomacromolecules 11 (2010) 1721–1726.
[33] I.K. Lednev, V.V. Ermolenkov, S. Higashiya, L.A. Popova, N.I. Topilina, J.T. Welch, Reversible thermal denaturation of a 60-kDa genetically engineered beta-sheet polypeptide, Biophys. J. 91 (2006) 3805–3818.
[34] L. Swint, A.D. Robertson, Thermodynamics of unfolding for turkey ovomucoid 3rd domain - thermal and chemical denaturation, Protein Sci. 2 (1993) 2037–2049.
[35] A.T. Petkova, Y. Ishii, J.J. Balbach, O.N. Antzutkin, R.D. Leapman, F. Delaglio, et al., A structural model for Alzheimer's beta-amyloid fibrils based on experimental constraints from solid state NMR, Proc. Natl Acad. Sci. USA 99 (2002) 16742–16747.
[36] A.V. Mikhonin, S.V. Bykov, N.S. Myshakina, S.A. Asher, Peptide secondary structure folding reaction coordinate: correlation between UV Raman amide III frequency, Psi Ramachandran angle, and hydrogen bonding, J. Phys. Chem. B 110 (2006) 1928–1943.
[37] M. Wang, R.D. JiJi, Spectroscopic detection of β-sheet structure in nascent Aβ oligomers, J. Biophoton. 4 (2011) 637–644.
[38] I.V. Baskakov, Thermodynamics and protein folding, in: J.D. Sipe (Ed.), Amyloid Proteins: The Beta Sheet Conformation and Disease, WILEY-VCH Verlag Gmbh & Co. KgaA, 2005, p. 53.
[39] M.G. Wolf, J. Gestel, S.W. de Leeuw, Modeling amyloid fibril formation: a free-energy approach, Methods Mol. Biol. 474 (2008) 153–179.
[40] J. Brange, L. Andersen, E.D. Laursen, G. Meyn, E. Rasmussen, Toward understanding insulin fibrillation, J. Pharm. Sci. 86 (1997) 517–525.
[41] H.Y. Kim, M.K. Cho, D. Riedel, C.O. Fernandez, M. Zweckstetter, Dissociation of amyloid fibrils of alpha-synuclein in supercooled water, Angew. Chem. Int. Ed. 47 (2008) 5046–5048.
[42] D. Kurouski, W. Lauro, I.K. Lednev, Amyloid fibrils are "alive": spontaneous refolding from one polymorph to another, Chem. Commun. 46 (2010) 4249–4251.
[43] I. Lednev, D. Kurouski, W. Lauro, Amyloid fibrils are "alive": spontaneous refolding from one polymorph to another, Chem. Commun. (2010) 172.
[44] D. Kurouski, R.K. Dukor, X. Lu, L.A. Nafie, I.K. Lednev, Spontaneous inter-conversion of insulin fibril chirality, Chem. Commun. 48 (2012) 2837–2839.
[45] M. Shanmugasundaram, D. Kurouski, W. Wan, G. Stubbs, R.K. Dukor, L.A. Nafie, et al., Rapid filament supramolecular chirality reversal of HET-s (218–289) prion fibrils driven by pH elevation, J. Phys. Chem. B 119 (2015) 8521–8525.
[46] T. Quiñones-Ruiz, M.F. Rosario-Alomar, K. Ruiz-Esteves, M. Shanmugasundaram, V. Grigoryants, C. Scholes, et al., Purple fibrils: a new type of protein chromophore, J. Am. Chem. Soc. 139 (2017) 9755–9758.
[47] M. Graff, J. Bukowska, Surface-enhanced Raman scattering (SERS) spectroscopy of enantiomeric and racemic methionine on a silver electrode-evidence for chiral discrimination in interactions between adsorbed molecules, Chem. Phys. Lett. 509 (2011) 58–61.
[48] X. Chen, Y. Tao, J. Li, H. Dai, W. Sun, X. Huang, et al., Aromatic residues regulating electron relay ability of S-Containing amino acids by formations of S∴ π multicenter three-electron bonds in proteins, J. Phys. Chem. C 116 (2012) 19682–19688.

CHAPTER 15

Recent advances in Raman spectroscopy of proteins for disease diagnosis

Keita Iwasaki[1], Hemanth Noothalapati[2,3] and Tatsuyuki Yamamoto[2,4]

[1]The United Graduate School of Agricultural Sciences, Tottori University, Tottori, Japan, [2]Raman Project Center for Medical and Biological Applications, Shimane University, Matsue, Japan, [3]Shimane University Administration Office, Shimane University, Matsue, Japan, [4]Faculty of Life and Environmental Science, Shimane University, Matsue, Japan

15.1 Introduction

Proteins are large and complex molecules that perform a vast array of biological functions in living beings. They play crucial roles in keeping the structure and function, and are also important in the regulation of processes in cells, tissues, and organs. There are tens of thousands of proteins in humans and each set has specific functions which include carrying out metabolic reactions, nucleic acid replication, providing structural integrity, transporting molecules, etc. Also, proteins do not function in isolation. Almost all proteins either interact with other proteins or bind to other molecules and such interactions enable them to perform functions as signal receptors, motors, catalysts, switches, etc. Such specificity is possible only due to their three-dimensional structure acquired by the precise sequence of amino acids. By definition, all disease states ultimately result in fundamental changes in chemical composition in cells and/or tissues. Hence, any change including deficiency, imbalance (under or over expression), or structural variation (misfolding) of proteins leads to disease. Therefore proteins serve as important biomarkers of disease and have been at the forefront in disease diagnosis. Depending on the purpose, many techniques are available to detect, isolate, purify, and characterize the structure of proteins. Raman spectroscopy (RS) is one of the emerging techniques for nonbioptic detection/quantification of proteins in cells and tissues which is especially useful in diagnosis and prognosis.

When atoms in a molecule are in periodic motion, molecular vibration occurs and the frequency associated with the motion is called vibrational frequency. RS is a powerful spectroscopic technique that can be used to observe vibrational modes in a molecule and differentiate molecular structures based on the inelastic scattering of monochromatic light. A Raman spectrum which is otherwise called a molecular fingerprint provides rich chemical

Vibrational Spectroscopy in Protein Research.
DOI: https://doi.org/10.1016/B978-0-12-818610-7.00015-3

information with high specificity and can be applied in both qualitative and quantitative analytical applications. Advantages of RS over other spectroscopic methods include (1) it is applicable to biological cells and tissues without the worry about interference from water; (2) it does not require external dye probes or genetic manipulations, which translates to measuring the sample as is with virtually no sample pretreatment; (3) it is a low or noninvasive method, and hence suitable for in vivo applications (nonbioptic); and (4) RS, when coupled with a microscope, achieves sub-μm spatial resolution and provides detailed space-resolved molecular information. Because of its ability to detect biochemical changes at the molecular level, RS has gained much attention as a useful bioanalytical tool in recent years, and has been successfully applied in a myriad of different problems in biological and medical sciences, which include structural identification and label-free molecular imaging of living cells [1], probing cellular metabolic pathways at the single cell level [2], cellular and tissue discrimination [3–5], disease diagnosis and prognosis [6–8].

RS is generally considered very weak and hence many variations such as surface-enhanced Raman spectroscopy (SERS), resonance Raman spectroscopy (RRS), surface-enhanced resonance Raman spectroscopy, tip-enhanced Raman spectroscopy, etc. have been developed and applied to enhance Raman signals. Moreover, other nonlinear modalities, such as coherent anti-Stokes Raman spectroscopy, stimulated Raman spectroscopy, etc., have been useful for real-time detection. Also, different methods employ different optical configurations for Raman spectral acquisition. Detailed discussion on each modality or instrument configuration is beyond the scope of this work. But, irrespective of the modality employed, Raman spectrum contains rich chemical information and the way we analyze spectra is of the utmost importance. Many data analytical procedures including both univariate and multivariate statistics have been developed to extract as much information as possible to make an informed decision. The aim of this chapter is to discuss recent advancements in normal spontaneous Raman spectroscopic or RRS detection of proteins and the most commonly utilized data analytical procedures for disease diagnosis, especially focusing on three major noncommunicable diseases (NCD): cancer, diabetes, and cardiovascular disease (CVD) [9–11].

15.2 Data analysis

15.2.1 Preprocessing

The Raman spectrum is also called a molecular fingerprint and in principle gives both qualitative and quantitative information. However, after measuring the Raman spectra, it is imperative to handle them carefully, keeping the design of the experiment, instrument performance, and acquisition parameters in mind to obtain accurate information in a reliable way. First, it is crucial to preprocess the data to remove any noise or unwanted signals, such as cosmic rays, detector noise, fluorescence, biological media, glass substrate, etc.,

that are mixed up with the actual information. Among many, the following steps are usually followed in order to analyze and interpret the data easily. We sometimes observe sharp spikes in our spectrum affecting only a few pixels. These artefacts are due to cosmic rays (high energy particles from outer space) reaching the charge-coupled device (CCD) detector. Since such events are totally random and unwanted, the likelihood increases with long exposure times. Therefore they are removed first to make sure the data is artefact-free. Also, data obtained from an instrument at one go with the same experimental conditions can be preprocessed together. However, in order to combine data sets obtained on different days, it is vital to first calibrate the wave number axis precisely before further preprocessing can be performed. Next, it is important to understand that a CCD detector will accumulate charge in its pixels proportional to exposure time even when there is no light entering the CCD. Therefore a dark spectrum is measured for any given acquisition condition and subtracted to eliminate its contribution from the raw data. Further, the performance of CCD detection is not uniform and its quantum efficiency depends on wavelength. Since the Raman spectrum includes a range of wavelengths, it is important to perform intensity correction. It is especially important when one is looking for semiquantitative or quantitative information. In addition to these, there is another complexity when dealing with biological samples, that is, intrinsic fluorescence, which not only competes but is also overwhelmingly intense when compared to the Raman signal. Therefore fluorescence background subtraction is required in order to extract chemical information from Raman spectra. These help in eliminating broad backgrounds but do not address high frequency noise. Depending on the quality of the raw data, smoothing can be employed to remove such noise. The aforementioned preprocessing is routinely performed but the data may need further processing. Since the Raman signal intensity depends on instrumental conditions (e.g., exposure time, laser power) and nature of sample (e.g., thickness), spectral normalization may be required in many situations. Moreover, a chosen analytical method may need extra processing to make it easy for data visualization, etc. [e.g., outlier removal and data centering in principal component analysis (PCA)].

15.2.2 Spectral analysis

Two major approaches are followed: univariate and multivariate.

15.2.2.1 Univariate analysis

Univariate analysis is the simplest form of data analysis. As the name suggests, it deals with only one variable at a time and can either be inferential or descriptive. It helps in describing the data using one particular variable but does not throw any light on causes or relationships, that is, univariate analysis explores each variable in a data set separately. Some examples include mean, median, mode, standard deviation, and variance. In the context of Raman spectral data, a single variable means wave number (individual Raman

band) and calculating intensities of each band by curve fitting or obtaining ratios of such calculated quantities can be thought of as a simple univariate analysis.

15.2.2.2 Multivariate analysis

Although univariate analysis is useful, it is severely limited in understanding the relationships between variables (wave numbers). Generally, a Raman spectrum contains thousands of such variables and in most cases molecular information is not completely represented by one or two Raman bands. Another important thing to keep in mind is the large volume (thousands of spectra) of data recorded by Raman hyperspectral imaging of biological samples (cells and tissues). Moreover, the differences among spectra are not always obvious but subtle. Therefore multivariate analysis (MA) is necessary where multiple variables are considered simultaneously to extract complete chemical information and to make the best use of Raman spectrum. Many multivariate analytical methods have been developed and applied to Raman spectroscopic data, especially to find cause and relationships. Roughly, MA can be classified into two major categories: multivariate classification (MC) and multivariate regression (MR). MC tries to finds patterns in the given data and helps in identifying to which set of categories the new observation belongs. It can be further divided into either unsupervised or supervised learning. Unsupervised MC methods [e.g., singular value decomposition (SVD), PCA, cluster analysis (CA), independent component analysis, multivariate curve resolution-alternating least squares (MCR-ALS) analysis, etc.] are exploratory in nature and needs no a priori information. Therefore it is very useful in biomedical applications of vibrational spectroscopy as it can separate disease-related samples into a different class even when the relationship is not initially evident. In contrast, supervised MC methods (e.g., linear discriminant analysis (LDA), partial least squares discriminant analysis (PLS-DA), support vector machines (SVM), neural networks (NN), etc.) requires a priori knowledge, such as class labels identifying the sample group. MR analyses such as linear regression, principle component regression, partial least squares regression (PLSR), etc. model the data by training from data sets with known information in order to estimate relationships among variables. It will eventually be applied to predict the unknown once a model is developed with sufficient accuracy. Although detailed explanation of each of the abovementioned methods is beyond the scope of this work, MCR-ALS analysis needs special mention.

15.2.2.3 Multivariate curve resolution analysis

MCR-ALS analysis is a powerful matrix factorization method with nonnegative constraints to decompose highly complex Raman spectroscopic data into a small number of physically meaningful pure spectra along with their concentration profiles. We can arrange experimentally acquired spectra into a $\boldsymbol{m} \times \boldsymbol{n}$ data matrix $\mathbf{A}$ in which $\boldsymbol{m}$ represents wave number axis (number of pixels in each spectrum) and $\boldsymbol{n}$ represents number of spectra

acquired at a different time or position or chemical composition, etc. In MCR-ALS, a low-rank approximation of matrix **A** is obtained by solving the following equation:

$$A = \mathrm{WH}, \tag{15.1}$$

where **W** is a $\boldsymbol{m \times k}$ matrix whose columns represent pure component spectra and **H** is a $\boldsymbol{k \times n}$ matrix whose rows represent intensity profiles of corresponding pure spectral components. The parameter $\boldsymbol{k}$ defines the number of components in the constructed model, and is determined a priori by the user.

In general, **W** and **H** are iteratively refined using alternating least squares so as to minimize the Frobenius norm $\|\mathbf{A}-\mathbf{WH}\|^2$. Practically, neither a Raman spectrum nor its concentration profile can take negative values and so nonnegative constraints on both **W** and **H** matrices ($\mathbf{W} \geq \mathbf{0}$; $\mathbf{H} \geq \mathbf{0}$) are also imposed. Therefore MCR-ALS analysis, which only requires nonnegative constraints, has huge advantages over other factorization methods, such as SVD or PCA in spectral analysis and provides physically meaningful results that can be interpreted in a straightforward manner.

Additionally, to improve the prediction accuracy and interpretability of the model, Lasso regression (L1-norm; α) can be applied as follows:

$$(\mathbf{W^T}\mathrm{W} + \alpha^2 E)\mathrm{H} = \mathbf{W^T}A, \tag{15.2}$$

$$(H\mathbf{H^T} + \alpha^2 E)\mathrm{W} = \mathrm{H}\mathbf{A^T}, \tag{15.3}$$

where **E** is a $\boldsymbol{k \times k}$ matrix whose elements are all unity. To obtain a sparser solution, another penalty such as ridge regression (L2-norm; β) can also be used in the following way:

$$(\mathbf{W^T}W + \beta^2 I)H = \mathbf{W^T}A, \tag{15.4}$$

$$(H\mathbf{H^T} + \beta^2 I)\mathrm{W} = \mathrm{H}\mathbf{A^T}, \tag{15.5}$$

where **I** is a $\boldsymbol{k \times k}$ identity matrix.

15.3 Raman spectroscopy for disease diagnosis

15.3.1 Cancer

Cancer is one of the leading causes of death globally. According to global cancer statistics, the International Agency for Research on Cancer estimated 18.1 million new cases and about 9.6 million cancer deaths in 2018 [9]. Among all cancers, lung cancer is the most commonly occurring and the leading cause of cancer deaths in both sexes combined. In males, lung cancer leads the number of deaths while breast cancer tops the chart among

women. Undoubtedly, cancer acts as the most important barrier to increasing life expectancy worldwide. However, cancer mortality can be reduced if the condition is detected and treated early. RS has been suggested as a potential tool in the early diagnosis of various cancers.

15.3.1.1 Breast cancer

As mentioned earlier, breast cancer is the most commonly found cancer in women with over 2 million new cases and over 600,000 deaths in 2018 [9]. It is the second most common cancer overall. RS has been considered for breast for a long time. Frank et al. carried out some of the early studies on the application of RS for discriminating breast cancer tissue [12,13]. The authors compared 784-nm excited Raman spectra of histologically normal human breast biopsy samples with both benign (fibrocystic change) and malignant tumors (infiltrating ductal carcinoma—IDC) from 12 patients. The study concluded that Raman spectra from normal samples were more lipid-like and the diseased condition were similar to proteins, especially to that of human collagen. Haka et al. then advanced the concept by measuring 130 Raman spectra from human breast tissues that included normal, fibrocystic change, fibroadenoma, and IDC from 58 patients. Raman spectra were analyzed by nine components in a linear combination model that well represents breast features. Out of the nine, the relative abundance of fat and collagen proved to be a good indicator of disease condition and could distinguish cancer from normal and benign tissues with high sensitivity and specificity, as shown in Fig. 15.1 [14].

Martin and coworkers showed an increase in the main amino acid contents in collagen (proline, glycine, tyrosine, phenyl alanine, etc.) was clear in IDC using FT-RS [15], while others have demonstrated significant differences in proteins' secondary structures as a marker for discriminating cancerous breast tissue [16,17]. Also, a specific human epidermal growth factor receptor protein (HER2) helps to control breast cell proliferation and survival and its overexpression is known to be associated with certain breast cancers and its recurrence. Bi et al. were even able to discriminate different HER2 expression groups with high sensitivity and specificity by applying multivariate statistical analysis to Raman spectra. In addition to HER2 amplification status, the study also investigated acquired drug resistance, especially from the treatment of lapatinib in breast cancer cells [18]. In a different approach, researchers have investigated the differences in serum samples from breast cancer patients and found clear differences in their amino acid/protein content [19]. Serum-based RS has also been applied to study pre- and postbreast adenocarcinoma resection in rat models where Murali Krishna and group found marked differences in amino acid content [20]. With such progress in discriminating breast cancer cells in vitro, there is now a more concerted effort to translate it to clinics for in vivo applications. One of the areas where RS can be immediately applied is in detection of tumor margins during surgery, as unexcised positive margins are associated with high recurrence rates. Haka et al. further used their

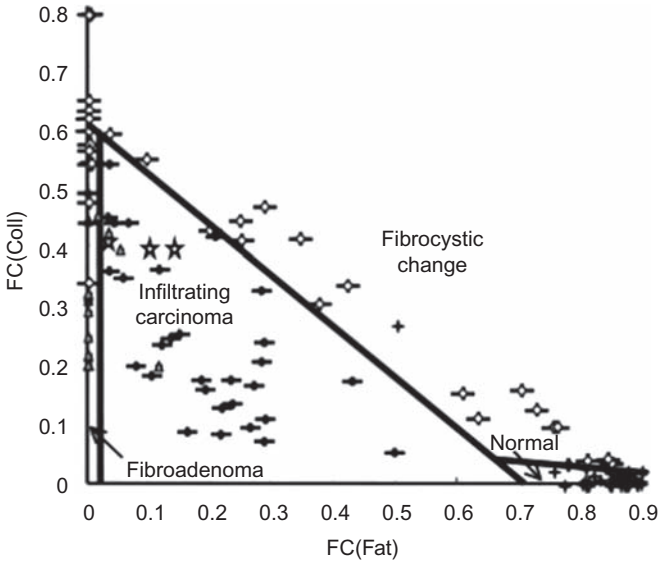


Figure 15.1: Breast cancer.

Scatter plot of fit coefficients (FC) displaying fat and collagen content for all pathologies encountered in Haka et al.'s study. Decision lines separate samples according to diagnoses very well. Gray stars, normal; diamonds, fibrocystic change; triangles, fibroadenoma; squares, infiltrating carcinoma; white stars, ductal epithelial hyperplasia. Source*: Adapted with permission from A.S. Haka, et al., Diagnosing breast cancer by using Raman spectroscopy, Proc. Natl Acad. Sci. U.S.A. 102 (35) (2005) 12371–12376. Copyright (2005) National Academy of Sciences, United States.*

developed indicators to demonstrate the feasibility of RS for real-time margin assessment during surgery by collecting spectra from breast tissue in vivo [21]. Notingher and coworkers proposed selective-sampling RS to detect positive margins during breast conserving surgery and achieved an automated intraoperative evaluation of tumor excision with high sensitivity and specificity based on the concentration of biochemical components such as collagen, nucleic acid, and fats [22,23]. Mahadevan-Jansen and coworkers attempted to develop a three-dimensional scanner based on RS that will help as an automated guidance tool during intraoperative surgery to ensure complete tumor removal in a single surgical procedure [24].

15.3.1.2 Lung cancer

Lung cancer, which includes both small cell and nonsmall cell lung cancer (NSCLC) types, is the most common cancer in both genders and is the leading cause of cancer deaths. In 2018 alone, over 2.1 million new cases and about 1.8 million deaths were predicted due to lung cancer worldwide [9]. It is common for lung cancer to metastasize before it is diagnosed, most often in metastatic stage IV, as often no symptoms develop when lung cancer is present. Hence early detection followed by appropriate treatment is necessary.

RS was proposed to be a good alternative for the diagnosis of lung cancers by the Hamaguchi group as early as 2001 using 1064-nm excitation [25]. The Zeng group utilized 785-nm excited RS to distinguish tumor from normal human bronchial tissues in vitro [26]. Protein secondary structure was identified to be a major discriminator. For instance, amide I and amide III band positions of proteins were found to be in the α-helix conformation while the bands were shifted in malignant tissues indicating an increase in relative amounts of proteins in β-sheets or random coils. However, interference from a very high fluorescence background hindered the measurement of fresh lung tissues. Hamaguchi and coworkers developed a near-infrared multichannel Raman spectrometer with 1064-nm excitation and successfully obtained high signal-to-noise ratio Raman spectra from fresh human lung tissues [27]. The authors found the amide I band to be very sensitive to cancer development. It is clear from Fig. 15.2 that amide I is markedly different in cancers, which was particularly attributed to tissues becoming dense due to an increase in a fibrous component containing collagen-like proteins.

In order to translate the method to the real clinical setting, Short et al. developed an endoscopic Raman probe to collect the first real-time in vivo human lung spectra using a reusable fiber catheter passed through a bronchoscope channel, but fluorescence interference was still high [28]. Since 1064-nm excited Raman is rare and expensive, Magee et al. developed a shifted subtracted RS equipped with a mini Raman probe with

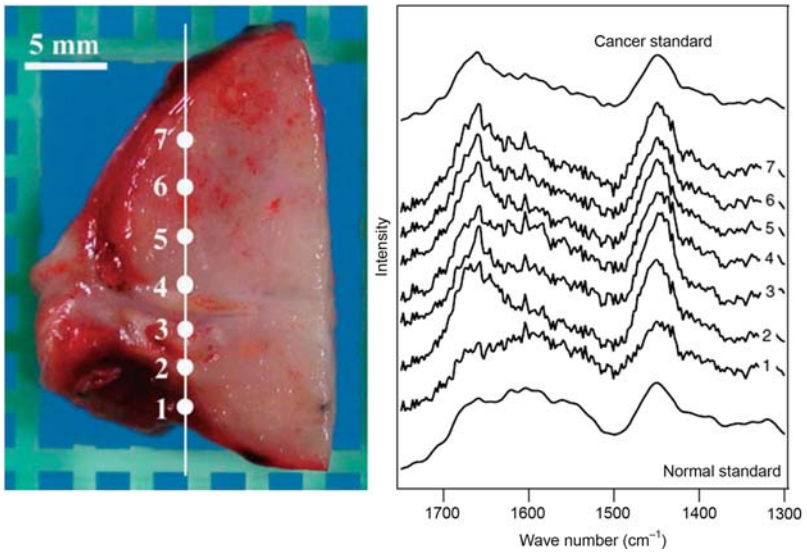


Figure 15.2: Lung cancer.
Photograph (*left*) and near-infrared Raman spectra (*right*) of fresh cancerous lung tissue measured along the line in picture. Both normal and cancer standard spectra are also included for comparison. Source: *Adapted with permission from Y.K. Min, et al., 1064 nm near-infrared multichannel Raman spectroscopy of fresh human lung tissues, J. Raman Spectrosc. 36 (1) (2005) 73–76.*

785-nm excitation to avoid the strong fluorescence background usually observed in lung tissues under visible excitation [29]. The authors combined the previous data analysis approach with PCA and successfully demonstrated ex vivo diagnosis of lung cancers and concluded that the used minifiber optical probe is suitable for insertion into the working channel of a bronchoscope. The Zeng group further developed and tested a miniature Raman probe for the early detection of cancer and reported the first in vivo Raman spectrum from peripheral lung [30]. The developed probe was just 1.35 mm in diameter with a minimum bend radius of 13 mm. They successfully measured a reasonably good fingerprint and high wave number regions of the Raman spectrum from cancerous nodule, normal tissue, and whole blood with clear observations of the changes of protein signatures.

A considerable amount of research has also been done on human lung cell lines. Sato and coworkers investigated the normal lung cell along with four types of pathologically different lung cancer cells with a visible laser emitting 532-nm light as the excitation source [31]. The measured Raman spectra had strong bands assignable to cytochrome-*c* due to a resonance effect with the excitation wavelength. The authors concluded that proteins in general and cytochrome-*c* in particular is useful to discriminate cancer cells from normal cells by PCA. In addition to PCA, the authors employed stepwise LDA for successful classification of histological types of these cells by malignancy type and strain with accuracy close to 100%. Harder et al. employed RS combined with PCA to identify radiation-induced biochemical changes in human NSCLC xenografts. The authors investigated the effect of single fractions of 0, 5, and 15 Gy radiation to the tumor harvested from mice, which was developed by injecting human NSCLC cells, and concluded there were marked differences in proteins and glycogen [32]. Byrne and coworkers performed detailed studies on short- and long-time exposure effects of the chemotherapeutic agent doxorubicin (DOX) on human lung adenocarcinoma cells with type II phenotype—A549. By employing a range of multivariate statistics consisting of PCA, PLSR, and ICA, the authors elucidated the spectroscopic signatures including proteins associated with DOX uptake and cellular interaction [33]. Based on the results, they tested the possibility of differentiating two lung cancer cell lines, A549 and human lung epidermoid carcinoma cell line (Calu-1), in response to DOX and also correlated the results with the expression levels of an important antiapoptotic protein bcl-2 [34,35]. Further, they also studied the interaction of the drug actinomycin D with A549 and Calu-1 in vitro. The authors concluded that the latter exhibits a higher sensitivity compared to the former based on subsequent cellular responses, especially those of protein profiles [36]. Recently, many researchers have attempted to employ SERS for lung cancer diagnosis but this is outside the scope of this chapter.

15.3.1.3 Oral cancer

Oral cancers are one of the most common cancers in the head and neck region with over 350,000 new cases and about 170,000 deaths in 2018 (survival rate is only $\sim$50%) [9]. It is especially common among developing South Asian countries such as India and poses a

significant health burden. Such high incidence rates are particularly attributed to extensive tobacco and alcohol consumption. Murali Krishna and coworkers have long been studying the application of RS for oral cancer diagnosis. Protein contribution to the Raman spectrum, as discussed earlier, has been the main discriminator for classifying normal, inflammatory, premalignant, and malignant oral ex vivo tissues [37]. In a study on the normal oral buccal mucosa tissues, they recorded Raman spectra from superior and inferior surfaces of intact tissues/incised biopsies and found that all normal oral tissues can be classified successfully against tumor [38]. They also succeeded in the identification of premalignant lesions and malignancy associated changes/cancer field effects responsible for secondary tumors in oral buccal mucosa in vivo [39]. Sahu et al. used a serum Raman spectroscopic approach aided with multivariate statistics for oral cancer screening [40], detection [41], and recurrence [42], in which proteins played an important role. The authors further advanced their study for the early detection of potential second primary/recurrences in oral cancer [43]. Recently they have applied similar discriminating factors via Raman exfoliate cytology for precancer diagnosis and prognosis prediction in oral cancers [6,44]. As in other cancers, identifying a positive tumor resection margin is necessary during oral cancer surgery. Barroso et al. attempted to discriminate healthy and cancerous tissue in oral squamous cell carcinoma (OSCC) from 14 patients undergoing tongue resection based on water content by measuring the high wave number region of the Raman spectrum. Even though the water content in squamous cell carcinoma was significantly higher than the surrounding healthy tissues and served as a discriminator, protein content played an important role as a standard in obtaining accurate values for water [45,46]. Typically, most of the studies successfully utilized either the ratio of groups of molecules using univariate analysis (protein/lipid, etc.) or multivariate approaches (PCA, LDA, CA, etc.) for diagnostic purposes without obtaining any pure biomolecular information that can aid in objective discrimination. To this end, Chen et al. demonstrated the applicability of RS coupled with MCR analysis for an objective and automated diagnosis of oral cancer. Interestingly, the study led to the extraction of a pure keratin Raman spectrum, an established protein marker for OSCC, from cancer tissues, which served for objective diagnosis as shown in Fig. 15.3 [47].

Following a similar approach using nonnegative least squares algorithm, Cals et al. made a detailed analytical study of the molecular composition of OSCC and healthy tongue tissue. Raman spectra were recorded from 44 tongue samples from 21 patients that included 142 OSCC, 202 surface squamous epithelium, 61 muscle, 65 adipose tissue, 581 connective tissue (CT), 26 gland, and 10 nerve regions (total of 1087 spectra) [48]. Each tissue spectrum was fit with 55 pure compound reference spectra and grouped based on their molecular class, such as proteins, lipids, and carbohydrates. A large number of proteins were included in the reference to simulate the measured spectrum. While the contribution of proteins was the largest (Fig. 15.4), the relative concentrations of proteins to other classes served as a good discriminator with 92% specificity and 98% sensitivity.

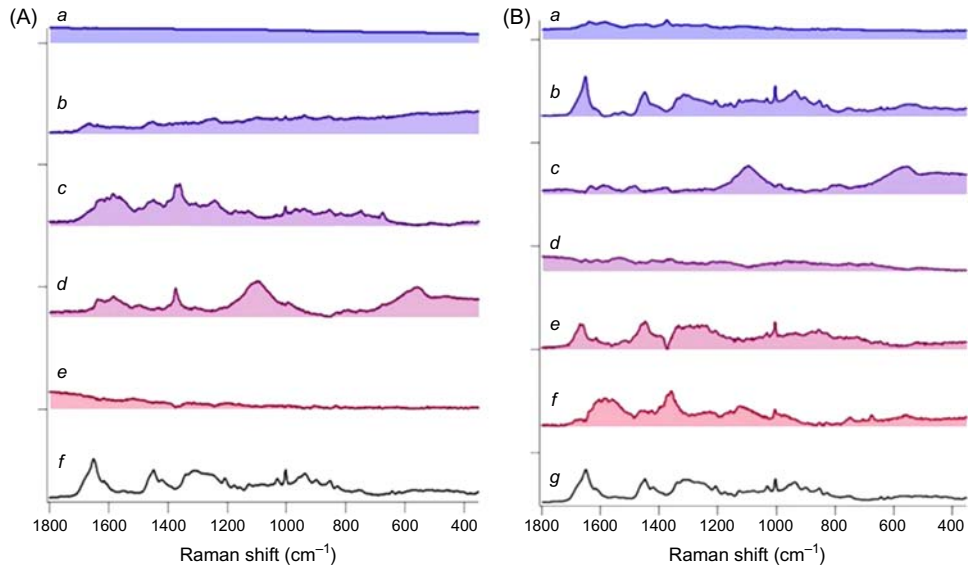


Figure 15.3: Oral cancer I.
MCR-ALS analysis results from normal and cancerous oral tissues. Intrinsic spectral components (A) normal tissue sample (*a*–*e*) compared with standard keratin spectrum (*f*) from a patient and (B) oral squamous cell carcinoma tissue from the same patient (*a*–*f*) compared with standard keratin spectrum (*g*). Source: *Adapted with changes from P.H. Chen, et al., Automatic and objective oral cancer diagnosis by Raman spectroscopic detection of keratin with multivariate curve resolution analysis, Sci. Rep. 6 (2016), 20097.*

15.3.1.4 Ovarian cancer

Ovarian cancer (OC) ranks first in deaths due to cancers of the female reproductive system and fifth in all cancer deaths among women. According to 2019 key statistical estimates for OC, 22,530 new cases and 13,980 deaths among women are predicted in the United States alone [49]. Effective screening for early diagnosis of OC can contribute toward timely treatment, thereby bringing down the mortality rate. Since there are no reliable screening methods available, RS has been proposed to be a good alternative.

In one of the first studies, Murali Krishna and coworkers evaluated the suitability of ex vivo handled ovarian tissues for optical diagnosis by 833-nm excited RS. The authors studied formalin-fixed, paraffin-embedded, and deparaffinized sections of the same ovarian normal and malignant tissues and found formalin-fixation to be the most suitable method with clear differences in spectral features between normal and malignant tissues [50]. Using PCA, they concluded that normal tissues contain high protein content, especially in helical and disordered structures, while less proteins are detected in malignant tissues. The authors then extended their study to discriminate normal, benign, and malignant tissues using Raman and

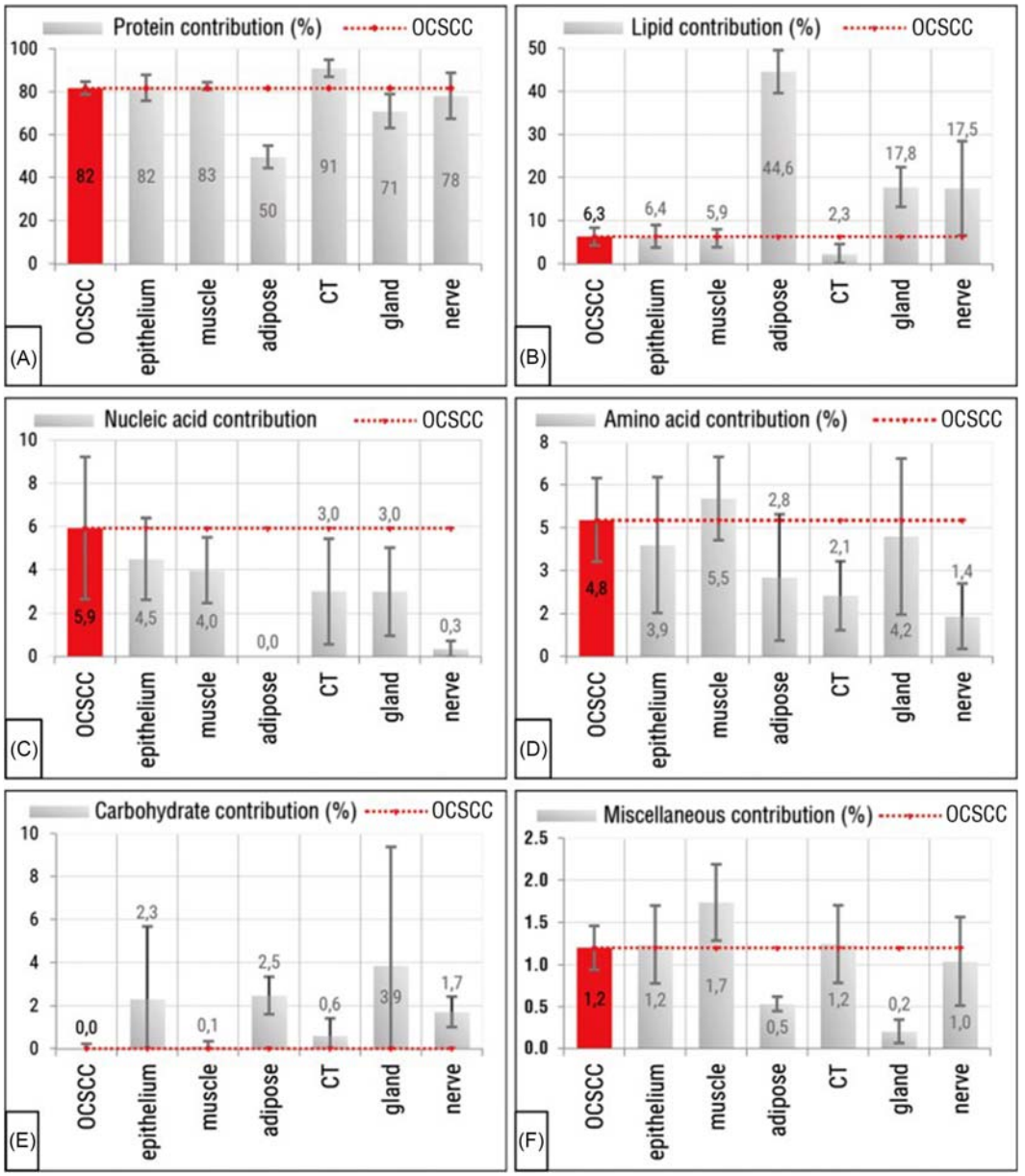


Figure 15.4: Oral cancer II.

Discriminating differences in the molecular composition of fitted individual oral cavity squamous cell carcinoma (OCSCC) and healthy tissue-type spectra, showing the contributions of (A) proteins; (B) lipids; (C) nucleic acids; (D) amino acids; (E) carbohydrates; (F) miscellaneous. *CT*, Connective tissue. Source*: Adapted with permission from F.L.J. Cals, et al., Raman spectroscopic analysis of the molecular composition of oral cavity squamous cell carcinoma and healthy tongue tissue, Analyst 143 (17) (2018) 4090–4102.*

FTIR spectroscopies by successfully employing CA [51]. Following the previous work, they tested 785-nm excited RS combined with multiple statistical approaches, such as PCA, Mahalanobis distance, spectral residuals, and limit test methodologies, to distinguish normal from malignant ovarian tissues [52]. They concluded that the malignant spectrum showed a broad amide I band and an intense amide III band compared to normal tissues in which bands at 855 and 940 cm^{-1} were more pronounced indicating protein structural changes as the main discriminator. One of the problems with OC is its recurrence even after platinum-based chemotherapy. Patients with recurrent OCs are categorized into either platinum-resistant or platinum-sensitive disease. However, there is a need for the development of better diagnostic procedures for platinum-resistant disease. To this end, Moradi et al. applied 785-nm excited RS to discriminate cisplatin-resistant (A2780cp) and cisplatin-sensitive (A2780s) human OC cells [53]. By employing PCA-LDA, relative increases in proteins and glutathione in resistant cells were found to be the main discriminators from the sensitive type and demonstrated the potential of RS in identifying chemotherapy-resistant cancers prior to treatment.

Given the grave situation associated with OCs, it is desirable to have reliable screening tools in order to diagnose the disease in its early stages. Current screening methods which include testing for overexpression of cancer antigen (CA-125) and transpelvic ultrasound imaging have poor sensitivity and specificity. Therefore Ullah et al. investigated the capability of RS of blood serum samples as a simple yet reliable alternative [54]. The main differences observed in the mean spectrum of healthy and OC subjects were attributed to differential expression of various proteins and their possible conformational changes. Moreover, since the spectra were very similar, the authors utilized SVM for computer-assisted screening and achieved high specificity and sensitivity. Following such studies, Martin and coworkers explored the possibility of screening OC by measuring Raman spectra of blood plasma from patients with the malignant condition and benign gynecological conditions. Additionally, they also compared the performance of simple spontaneous RS and SERS in their abilities to discriminate in the given situation and found much superior performance by normal RS using SVM analysis despite taking into account the differences in the CA-125 protein as shown in Fig. 15.5 [55].

The authors went on to then develop a three-dimensional PCA approach for the identification of OCs from healthy subjects through observed differences in amino acids and proteins among other biomolecules [56].

15.3.1.5 Prostate cancer

Prostate cancer is the second most common cancer and the fifth most common cause of cancer death among men. According to the latest key statistics available, 174,650 new prostate cancer cases and about 31,620 deaths are expected in 2019 in the United States alone [49]. Though there are several types of prostate cancer, the most common is

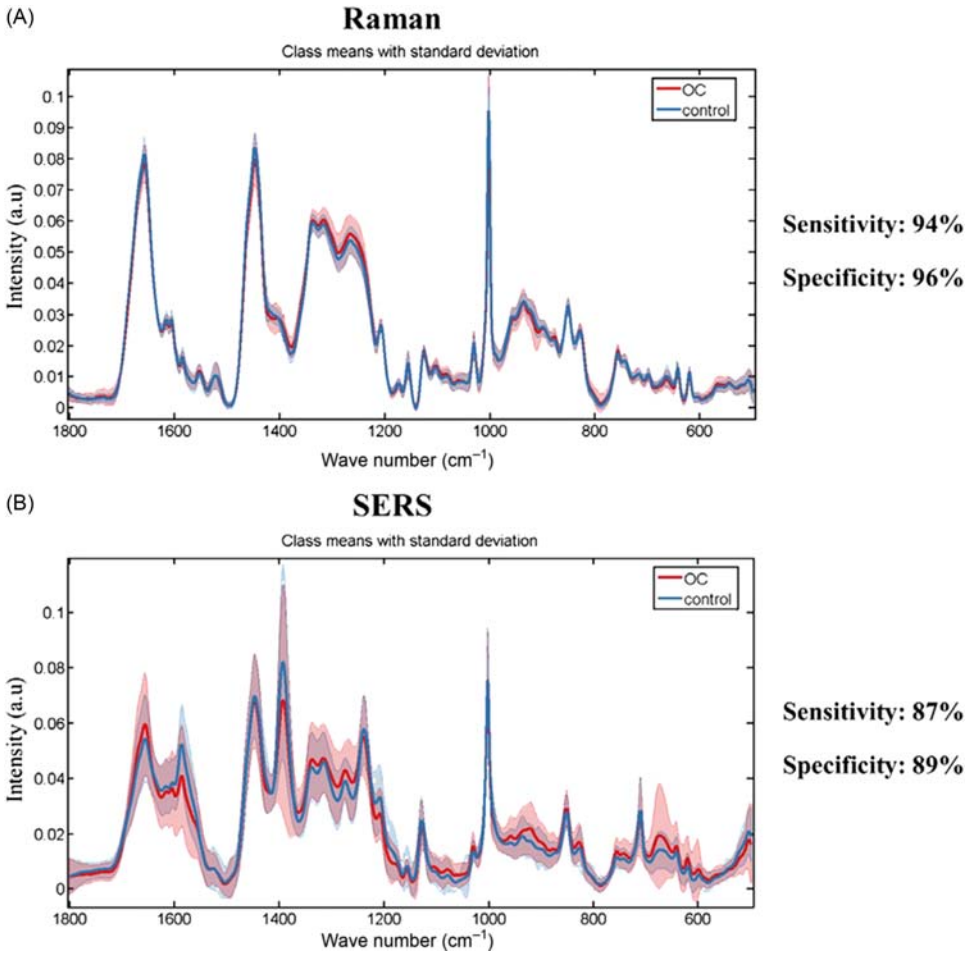


Figure 15.5: Ovarian cancer.
Diagnostic segregation of ovarian cancer (OC) with (A) Raman spectroscopy (RS) and (B) surface-enhanced Raman spectroscopy (SERS). Source*: Reprinted from Paraskevaidi et al., Raman spectroscopic techniques to detect ovarian cancer biomarkers in blood plasma, Talanta 189 (2018) 281–288, with permission from Elsevier.*

adenocarcinoma which is a glandular malignancy. In many cases, prostate cancer is asymptomatic and without screening one will not know that he has the disease. Since good treatment is available, the onus is on early detection. However, there are downsides to the current standard screening method based on the histopathological Gleason grading system and more straightforward techniques are required to assess the situation and accurately identify the stage of the disease. RS, as a potential nonbioptic technique, has been proposed as a good alternative.

Crow et al. applied 832-nm excited RS to identify and grade prostatic adenocarcinoma among tissue biopsies in vitro [57]. They included benign samples from 27 different patients including prostatic hyperplasia (a noncancerous increase in size of the prostate gland) and three risk groups of adenocarcinomas based on Gleason score (GS), viz., GS < 7 (low risk), GS $= 7$ (intermediate risk), and GS > 7 (high risk). By employing PCA-LDA, the authors were successful in discriminating different grades of adenocarcinoma with high specificity and sensitivity. Crow and coworkers then extended the study by testing the applicability of a suitable in vivo fiberoptic Raman system for discriminating benign and malignant bladder and prostate cancer. Taking advantage of the previously employed multivariate analytical methods, the authors were successful in distinguishing both groups with high accuracy [58]. Human prostate can be divided into three zones and they show zone-specific susceptibility to malignancy. While the central zone (CZ) is relatively immune to any pathology, benign prostatic hypertrophy is exclusively observed in the transition zone (TZ) and prostate adenocarcinoma is usually observed in the peripheral zone (PZ). Therefore Martin and coworkers successfully utilized RS followed by PCA to discriminate zone-specific spectral signatures in normal human prostates [59]. The authors concluded DNA to be the main discriminating factor for PZ from CZ and TZ while proteins played a key role in distinguishing CZ from TZ. They further extended their work to segregate human prostate tissues based on risk classification for adenocarcinoma. The study involved samples from high-risk (United Kingdom) and low-risk (India) populations. By employing PCA-LDA, the authors were successful in segregation where protein secondary structure variations served as biomolecular markers [60]. Wang et al. investigated the feasibility of RS in the diagnosis and prognosis of castration-resistant prostate cancer using PCA and SVM analysis [7]. Aubertin et al. attempted mesoscopic characterization of prostate cancer by not only diagnosing but grading prostate cancer. They measured fresh and nonprocessed postprostatectomy samples categorized into benign and malignant with cancer grade 1−5 using a macroscopic handheld RS probe. They found a higher intensity of peaks directly associated with proteins and a further application of a NN led them to achieve discrimination to each category with high accuracy [61]. They further combined high wave number and fingerprint RS for the detection of prostate cancer during radical prostatectomy based on supervised classification using SVM [62]. Magalhaes et al. performed a pilot study on the 1064-nm excited RS for prostate cancer diagnosis based on SVM and predicted GSs with 95% accuracy [63].

Matthews et al. studied biochemical signatures of in vitro radiation response in prostate tumor cells by applying single-cell RS and PCA analysis. Their work revealed changes in protein conformations and degradation of structured proteins to be the main factors. They also segregated cells into radiosensitive and radioresistant types, thereby demonstrating the potential of RS to investigate, monitor, and predict radiation response [64]. Harvey et al. explored the applicability of Raman optical tweezers to study urological cell lines with an

aim of identifying prostate cancer cells in patient urine and peripheral fluids. They used both single-beam forward scattering using a 514.5-nm laser and dual-beam backward scattering geometry using 1064-nm (trapping laser) and 514.5-nm (probe laser) [65]. Since both prostate and bladder cells are expected to be present in urine, the authors trapped and analyzed live prostate cancer cells (PC-3) and bladder cells (MGH-U1) using PCA and successfully distinguished them based on the differences in their biochemical components, especially those of their proteins. Interestingly they investigated the effect of the size of cells and concluded that the separation is based predominantly on cell type than size. Medipally et al. investigated different instrumental and sample preparation parameters to develop high-throughput blood plasma-based RS methods for rapid screening of prostate cancer. They concluded that while accuracy of 785 nm increases with exposure time until 40 seconds, the performance of 532-nm excitation doesn't increase after 1 second exposure [66].

15.3.1.6 Other cancers

In addition to the abovementioned cancers, RS of proteins has been useful in the diagnosis and prognosis of several other cancers including cervical cancer [67], colorectal cancer [68,69], esophageal cancer [70,71], gastric cancer [72,73], kidney cancer [74], and skin cancers [75,76], among others.

15.3.2 Diabetes

Diabetes is a chronic disease with an estimated 450 million people living with the disease and 5 million deaths related to diabetes worldwide in 2017 [10]. Diabetes causes many serious health problems if left untreated or uncontrolled. Since there is no cure for diabetes, regular monitoring of the condition is important. Traditionally, blood glucose level has been the main indicator and indeed RS has also been applied to noninvasively monitor glucose levels. Since this chapter is concerned with RS of proteins, let us see how they help in the diagnosis and prognosis of diabetes. Collagen is one of the most abundant proteins in the human body and is also a major target for glycation due to its low turnover. It is known that increased glycation of type I collagen is related to diabetes [77,78]. Glycation is different from glycosylation in that the former is a nonenzymatic and opportunistic addition of reducing sugars to a protein side chain, while the latter is an enzymatically regulated intracellular event. Particularly, accelerated accumulation of advanced glycation end-products (AGEs) are observed under hyperglycemic conditions and is also closely related to chronic stress in diabetes [79]. In fact, inhibitors of AGEs formation and AGE-protein cross-link breakers have been shown to reduce complications in experimental models [80,81]. Therefore researchers are interested in utilizing AGEs as a marker for the noninvasive screening of diabetes. Assessment of advanced glycation is especially useful to

predict associated long-term risks in diabetic patients [82]. Skin fluorescence has been proposed as a clinical tool for such purposes [83–85]. However, among the AGEs such as pentosidine, carboxymethyllysine, and glucosepane (cross-links between arginine and lysine residues in collagen), only pentosidine is fluorescent and the other two are not. So RS, being a vibrational technique, is supposed to effectively aid in detecting glycation of proteins. Martin et al. applied 785-nm excited in vivo confocal RS to measure human skin (average spectrum from skin surface to 120-μm depth) on 30 women volunteers divided into 3 groups of 10 each; elderly women with diabetes (65–80 years old), healthy elderly (65–80 years old) and healthy young women (20–33 years old). The authors reported that they identified AGEs peaks (especially those of pentosidine and glucosepane) and associated changes in collagen to quantify the glycation process in human skin [86]. Further, in order to get a holistic view and increase accuracy in a clinical setting, Paolillo et al. tested the possibility of applying both fluorescence and RS together to investigate the glycation of skin collagen in human subjects [87]. The authors studied 94 volunteers grouped into five categories of age, health condition, and Fitzpatrick skin types. The first four groups—adult/nondiabetic, middle-aged/nondiabetic, elderly/nondiabetic, and insulin resistance/diabetic—included skin types I–IV, and the fifth group—dark skin/nondiabetic—was composed of volunteers with skin type V and VI. Analysis of Raman spectra from the forearm revealed the degradation of type I collagen and increased glycation in diabetics. However, the study also pointed to the possibility of diagnostic error in dark-skinned people.

Although RS-based AGEs detection is in the beginning phase, successful development will have far-reaching significance as AGEs are thought to play a crucial role in the onset and progression of CT disease, affecting both biological and biomechanical functions.

15.3.3 Cardiovascular diseases

CVD which include myocardial infarction (commonly known as heart attack) and stroke occur when blood flow to either the heart or brain gets blocked by blood clots. CVD are the world's leading cause of deaths due to NCD with over 17.5 million deaths in 2012 [88]. These diseases are preventable provided regular screening is performed to identify worsening health conditions such as high blood pressure, high low density lipoprotein cholesterol, obesity, etc. RS has been useful in fundamental studies related to myocardial infarction that could lead to new diagnostic procedures in the future. Pascut et al. demonstrated label-free Raman-activated cell sorting of cardiomyocytes derived from human embryonic stem cells [89]. Brazhe et al. applied RS to map the redox states of reduced cytochromes *c*, *c*1, and *b* of complexes II and III in mitochondria of live cardiomyocytes [90]. Takamatsu and coworkers demonstrated high-speed Raman imaging

of heart tissue based on RRS of heme proteins using a 532-nm laser [91]. They identified individual cardiomyocytes, blood vessels, and fibrotic tissues mainly from reduced *b*- and *c*-type cytochromes, oxy and deoxy hemoglobin., and collagen, respectively. The authors also employed RS for label-free evaluation of myocardial infarction and its repair in rats [92]. Raman spectra were successfully measured from unfixed frozen tissues of normal and infarcted heart tissues using 532-nm excitation on days 2, 5, and 21 after coronary ligation. They identified spectral markers for cytochrome *c* and b_5 and found decreased reduced cytochrome *c* in both necrotic and granulation stages than in normal cardiac tissue. Moreover, they observed the mature collagen contribution to be dominant in the fibrotic tissue spectrum and hence were able to discriminate cardiac samples accurately. Based on these findings, the authors constructed a model by applying PLS-DA for myocardial viability evaluating the border infarct zone in patients undergoing surgical ventricular restoration [93]. Further, they were able to identify infarcted/noninfarcted tissues with high accuracy and proposed RS for the detection of myocardial viability during surgery as shown in Fig. 15.6 [93].

Ohira et al. tested the applicability of RS for evaluating myocardial ischemia, especially during the early ischemic phase, by focusing on cytochromes in subepicardial myocardium in Langendorff-perfused rat heart during global ischemia [94]. Induction of stopped blood flow indicated the progressive reduction in mitochondrial respiration, thereby suggesting the potential of RS in evaluating acute ischemic hearts.

15.3.4 Other diseases

In addition to the abovementioned diseases, RS has been proved to be very useful in a myriad of other conditions. We have developed a RRS-based approach for eosinophilic esophagitis, which is a rare pathological condition in which the esophageal epithelial layer is densely infiltrated by eosinophils. Eosinophil peroxidase protein served as a marker for the presence of eosinophils [95]. Endoscopic observations of the esophagus can diagnose the disease in the presence of clear longitudinal furrows and multiple concentric rings which are seen only in the later stages, and therefore early diagnostics is the need of the hour. In our study, inflammation induced by the administration of interleukin-33 in mouse clearly showed the presence of eosinophils in the esophagus while none were observed in the control which can seen from Fig. 15.7 [95].

By measuring collagen, keratin, and/or other protein signatures, RS has recently been applied successfully in the diagnosis of osteoarthritis [96], osteoporosis [97], chronic fatigue syndrome [98], neurodegenerative diseases [99], Marfan syndrome [100], and male infertility [101].

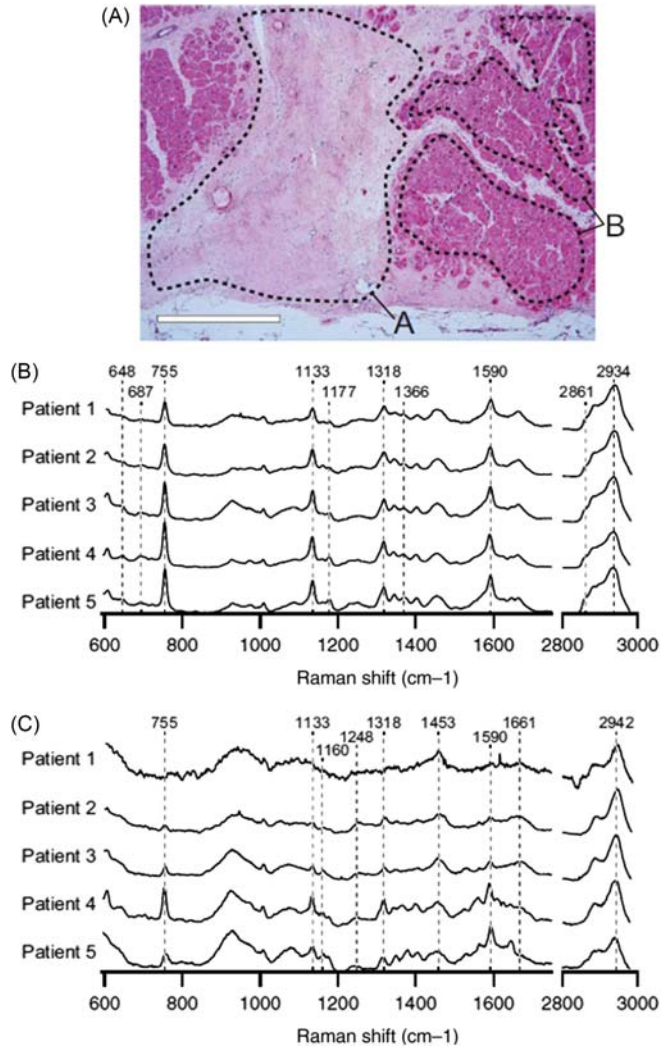


Figure 15.6: Myocardial infarction.

Raman spectra of the border zone myocardium obtained from five patients. (A) A representative HE stained image of the boundary between the infarcted region (*dashed area* A) and noninfarcted region (*dashed area* B) of the heart tissue excised from patient 3. Scale bar represents 1 mm. Representative Raman spectra of noninfarcted myocardium (B) and infarcted myocardium (C) of each patient. Each spectrum was averaged to obtain sufficient signal-to-noise (SNR) ratio for the detail identification of Raman bands, and normalized by the integrated intensity of each Raman spectrum. Source*: Adapted from T. Yamamoto, et al., Label-free evaluation of myocardial infarct in surgically excised ventricular myocardium by Raman spectroscopy, Sci. Rep. 8 (1) (2018) 14671.*

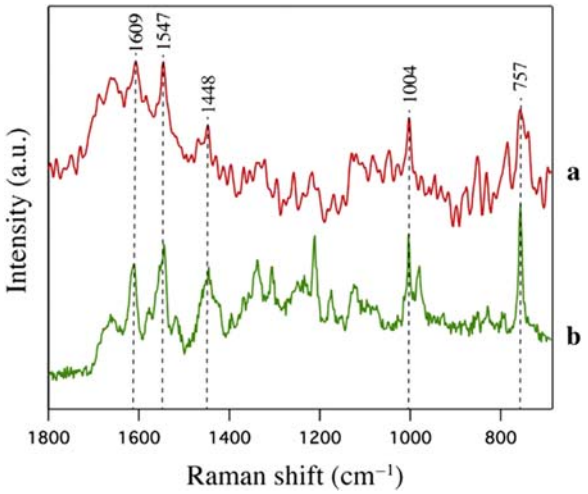


Figure 15.7: Eosinophilic esophagitis.
Comparison of Raman spectra of (A) IL-33 administered mouse esophagus and (B) single eosinophil extracted from whole blood. Excellent overlap of bands is observed indicating the molecular component in (A) is from protein eosinophil peroxidase. Source*: Reprinted from N. Hemanth, et al., Towards the development of a non-bioptic diagnostic technique for eosinophilic esophagitis using Raman spectroscopy, Vib. Spectrosc. 85 (2016) 7–10, with permission from Elsevier.*

15.4 Conclusions

Diseases affect people physically, emotionally, financially, and socially and are a significant source of psychological stress. In order to save time, cost, and to reduce stress, it is important to be able to diagnose the condition as early as possible. Undoubtedly, mortality can be reduced if diseases are detected and treated early. In addition to early diagnosis, development of effective screening methodologies which aim to identify individuals with abnormalities but who have not yet developed any symptoms is also equally important. Hence there is an absolute need for the development of new nonbioptic techniques that can achieve rapid (real-time detection) yet objective diagnosis. In this regard, RS has immense potential, particularly because of the wealth of molecular information contained in its spectra, and RS-based diagnostic methods can positively address these issues. As we have seen in this chapter, RS has been applied successfully both in vitro and in vivo to cells, tissues, and organs. Thanks to significant development in fiber probe technology, RS has even been successfully demonstrated in clinical settings during intraoperative procedures. Considerable research is also underway by RS researchers to identify spectral markers for disease diagnoses using body fluids such as blood, plasma, serum, urine, etc. On the data analysis front, PCA and PCA-LDA have traditionally been the most sought-after multivariate statistical approaches and machine learning methods, such as SVM and NN, have been picking up the pace recently. Most of these methods

discriminate using subtle differences in spectral response but offer little insight into the specific chemical changes responsible for discrimination. One of the major concerns in the field has been the extraction of pure chemical information from such complex biological samples which contain hundreds of different biomolecules (if not thousands) at any measured point with inherent molecular heterogeneity. Though it has partially been addressed by MCR-ALS analysis, there still are some limiting factors especially for real-time applications. Further development in MCR-ALS to decide the optimal number of components and appropriate penalty terms, etc. can help in automation and can eventually be applied in real-time diagnosis. It is straightforward to observe that RS of proteins has been useful in early diagnosis, prognosis, and therapy. In particular, amide I and amide III bands, protein/lipid ratios, collagen and its modifications, keratin and cytochromes, etc. have been the pillars for diagnostic purposes. From recent developments of RS of proteins, it can be understood that the technique is promising as a pathological tool to not only diagnose but also to grade disease stages. With many successful examples of the application of RS for margin detection, it might ultimately become a surgeon's go-to tool for assistance in the real-time assessment of margin status at the time of surgery in the near future.

References

[1] H. Noothalapati, et al., Label-free chemical imaging of fungal spore walls by raman microscopy and multivariate curve resolution analysis, Sci. Rep. 6 (2016) 27789.
[2] N. Hemanth, S. Shinsuke, Exploring metabolic pathways in vivo by a combined approach of mixed stable isotope-labeled raman microspectroscopy and multivariate curve resolution analysis, Anal. Chem. 86 (15) (2014) 7828−7834.
[3] G.R. Lloyd, et al., Discrimination between benign, primary and secondary malignancies in lymph nodes from the head and neck utilising Raman spectroscopy and multivariate analysis, Analyst 138 (14) (2013) 3900−3908.
[4] T. Tolstik, et al., Discrimination and classification of liver cancer cells and proliferation states by Raman spectroscopic imaging, Analyst 139 (22) (2014) 6036−6043.
[5] F.M. Lyng, et al., Discrimination of breast cancer from benign tumours using Raman spectroscopy, PLoS One 14 (2) (2019) e0212376.
[6] A. Sahu, et al., Raman exfoliative cytology for prognosis prediction in oral cancers: a proof of concept study, J. Biophotonics (2019) e201800334.
[7] L. Wang, et al., Raman spectroscopy, a potential tool in diagnosis and prognosis of castration-resistant prostate cancer, J. Biomed. Opt. 18 (8) (2013) 87001.
[8] N.D. Magee, et al., Raman microscopy in the diagnosis and prognosis of surgically resected nonsmall cell lung cancer, J. Biomed. Opt. 15 (2) (2010).
[9] F. Bray, et al., Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J. Clin. 68 (6) (2018) 394−424.
[10] N.H. Cho, et al., IDF Diabetes Atlas: global estimates of diabetes prevalence for 2017 and projections for 2045, Diabetes Res. Clin. Pract. 138 (2018) 271−281.
[11] A. Timmis, et al., European Society of Cardiology: Cardiovascular Disease Statistics 2017, Eur. Heart J. 39 (7) (2018) 508−579.
[12] C.J. Frank, R.L. McCreery, D.C. Redd, Raman spectroscopy of normal and diseased human breast tissues, Anal. Chem. 67 (5) (1995) 777−783.

[13] C.J. Frank, et al., Characterization of human breast biopsy specimens with near-IR Raman spectroscopy, Anal. Chem. 66 (3) (1994) 319–326.
[14] A.S. Haka, et al., Diagnosing breast cancer by using Raman spectroscopy, Proc. Natl Acad. Sci. U.S.A. 102 (35) (2005) 12371–12376.
[15] R.A. Bitar, et al., Biochemical analysis of human breast tissues using Fourier-transform Raman spectroscopy, J. Biomed. Opt. 11 (5) (2006) 054001.
[16] B. Brozek-Pluska, et al., Raman spectroscopy and imaging: applications in human breast cancer diagnosis, Analyst 137 (16) (2012) 3773–3780.
[17] C.H. Liu, et al., Resonance Raman and Raman spectroscopy for breast cancer detection, Technol. Cancer Res. Treat. 12 (4) (2013) 371–382.
[18] X. Bi, et al., Evaluating HER2 amplification status and acquired drug resistance in breast cancer cells using Raman spectroscopy, J. Biomed. Opt. 19 (2) (2014) 025001.
[19] J.L. Pichardo-Molina, et al., Raman spectroscopy and multivariate analysis of serum samples from breast cancer patients, Lasers Med. Sci. 22 (4) (2007) 229–236.
[20] T. Bhattacharjee, et al., Raman spectroscopy of serum: a study on 'pre' and 'post' breast adenocarcinoma resection in rat models, J. Biophotonics 8 (7) (2015) 575–583.
[21] A.S. Haka, et al., In vivo margin assessment during partial mastectomy breast surgery using raman spectroscopy, Cancer Res. 66 (6) (2006) 3317–3322.
[22] K. Kong, et al., Towards intra-operative diagnosis of tumours during breast conserving surgery by selective-sampling Raman micro-spectroscopy, Phys. Med. Biol. 59 (20) (2014) 6141–6152.
[23] D.W. Shipp, et al., Intra-operative spectroscopic assessment of surgical margins during breast conserving surgery, Breast Cancer Res. 20 (1) (2018) 69.
[24] G. Thomas, et al., Evaluating feasibility of an automated 3-dimensional scanner using Raman spectroscopy for intraoperative breast margin assessment, Sci. Rep. 7 (1) (2017) 13548.
[25] S. Kaminaka, et al., Near-infrared Raman spectroscopy of human lung tissues: possibility of molecular-level cancer diagnosis, J. Raman Spectrosc. 32 (2) (2001) 139–141.
[26] Z. Huang, et al., Near-infrared Raman spectroscopy for optical diagnosis of lung cancer, Int. J. Cancer 107 (6) (2003) 1047–1052.
[27] Y.K. Min, et al., 1064 nm near-infrared multichannel Raman spectroscopy of fresh human lung tissues, J. Raman Spectrosc. 36 (1) (2005) 73–76.
[28] M.A. Short, et al., Development and preliminary results of an endoscopic Raman probe for potential in vivo diagnosis of lung cancers, Opt. Lett. 33 (7) (2008) 711–713.
[29] N.D. Magee, et al., Ex vivo diagnosis of lung cancer using a Raman miniprobe, J. Phys. Chem. B 113 (23) (2009) 8137–8141.
[30] H.C. McGregor, et al., Development and in vivo test of a miniature Raman probe for early cancer detection in the peripheral lung, J. Biophotonics 11 (2018) 11.
[31] Y. Oshima, et al., Discrimination analysis of human lung cancer cells associated with histological type and malignancy using Raman spectroscopy, J. Biomed. Opt. 15 (1) (2010) 017009.
[32] S.J. Harder, et al., Raman spectroscopy identifies radiation response in human non-small cell lung cancer xenografts, Sci. Rep. 6 (2016) 21006.
[33] Z. Farhane, F. Bonnier, H.J. Byrne, Monitoring doxorubicin cellular uptake and trafficking using in vitro Raman microspectroscopy: short and long time exposure effects on lung cancer cell lines, Anal. Bioanal. Chem. 409 (5) (2017) 1333–1346.
[34] Z. Farhane, et al., Differentiating responses of lung cancer cell lines to Doxorubicin exposure: in vitro Raman micro spectroscopy, oxidative stress and bcl-2 protein expression, J. Biophotonics 10 (1) (2017) 151–165.
[35] Z. Farhane, et al., Doxorubicin kinetics and effects on lung cancer cell lines using in vitro Raman micro-spectroscopy: binding signatures, drug resistance and DNA repair, J. Biophotonics 11 (1) (2018).

[36] Z. Farhane, F. Bonnier, H.J. Byrne, An in vitro study of the interaction of the chemotherapeutic drug actinomycin D with lung cancer cell lines using Raman micro-spectroscopy, J. Biophotonics 11 (1) (2018).
[37] R. Malini, et al., Discrimination of normal, inflammatory, premalignant, and malignant oral tissue: a Raman spectroscopy study, Biopolymers 81 (3) (2006) 179–193.
[38] A. Deshmukh, et al., Raman spectroscopy of normal oral buccal mucosa tissues: study on intact and incised biopsies, J. Biomed. Opt. 16 (12) (2011) 127004.
[39] S.P. Singh, et al., In vivo Raman spectroscopy of oral buccal mucosa: a study on malignancy associated changes (MAC)/cancer field effects (CFE), Analyst 138 (14) (2013) 4175–4182.
[40] A.K. Sahu, et al., Oral cancer screening: serum Raman spectroscopic approach, J. Biomed. Opt. 20 (11) (2015) 115006.
[41] A. Sahu, et al., Raman spectroscopy of serum: an exploratory study for detection of oral cancers, Analyst 138 (14) (2013) 4161–4174.
[42] A. Sahu, et al., Recurrence prediction in oral cancers: a serum Raman spectroscopy study, Analyst 140 (7) (2015) 2294–2301.
[43] A. Malik, et al., In vivo Raman spectroscopy-assisted early identification of potential second primary/recurrences in oral cancers: an exploratory study, Head Neck 39 (11) (2017) 2216–2223.
[44] A. Sahu, et al., Raman exfoliative cytology for oral precancer diagnosis, J. Biomed. Opt. 22 (11) (2017) 1–12.
[45] E.M. Barroso, et al., Discrimination between oral cancer and healthy tissue based on water content determined by Raman spectroscopy, Anal. Chem. 87 (4) (2015) 2419–2426.
[46] E.M. Barroso, et al., Water concentration analysis by raman spectroscopy to determine the location of the tumor border in oral cancer surgery, Cancer Res. 76 (20) (2016) 5945–5953.
[47] P.H. Chen, et al., Automatic and objective oral cancer diagnosis by Raman spectroscopic detection of keratin with multivariate curve resolution analysis, Sci. Rep. 6 (2016) 20097.
[48] F.L.J. Cals, et al., Raman spectroscopic analysis of the molecular composition of oral cavity squamous cell carcinoma and healthy tongue tissue, Analyst 143 (17) (2018) 4090–4102.
[49] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2019, CA Cancer J. Clin. 69 (1) (2019) 7–34.
[50] C.M. Krishna, et al., Evaluation of the suitability of ex vivo handled ovarian tissues for optical diagnosis by Raman microspectroscopy, Biopolymers 79 (5) (2005) 269–276.
[51] C.M. Krishna, et al., FTIR and Raman microspectroscopy of normal, benign, and malignant formalin-fixed ovarian tissues, Anal. Bioanal. Chem. 387 (5) (2007) 1649–1656.
[52] K. Maheedhar, et al., Diagnosis of ovarian cancer by Raman spectroscopy: a pilot study, Photomed. Laser Surg. 26 (2) (2008) 83–90.
[53] H. Moradi, et al., Raman micro-spectroscopy applied to treatment resistant and sensitive human ovarian cancer cells, J. Biophotonics 10 (10) (2017) 1327–1334.
[54] I. Ullah, et al., Computer assisted optical screening of human ovarian cancer using Raman spectroscopy, Photodiagn. Photodyn. Ther. 15 (2016) 94–99.
[55] M. Paraskevaidi, et al., Raman spectroscopic techniques to detect ovarian cancer biomarkers in blood plasma, Talanta 189 (2018) 281–288.
[56] C.L.M. Morais, P.L. Martin-Hirsch, F.L. Martin, A three-dimensional principal component analysis approach for exploratory analysis of hyperspectral data: identification of ovarian cancer samples based on Raman microspectroscopy imaging of blood plasma, Analyst 144 (7) (2019) 2312–2319.
[57] P. Crow, et al., The use of Raman spectroscopy to identify and grade prostatic adenocarcinoma in vitro, Br. J. Cancer 89 (1) (2003) 106–108.
[58] P. Crow, et al., Assessment of fiberoptic near-infrared raman spectroscopy for diagnosis of bladder and prostate cancer, Urology 65 (6) (2005) 1126–1130.
[59] I.I. Patel, F.L. Martin, Discrimination of zone-specific spectral signatures in normal human prostate using Raman spectroscopy, Analyst 135 (12) (2010) 3060–3069.

[60] I.I. Patel, et al., Segregation of human prostate tissues classified high-risk (UK) versus low-risk (India) for adenocarcinoma using Fourier-transform infrared or Raman microspectroscopy coupled with discriminant analysis, Anal. Bioanal. Chem. 401 (3) (2011) 969–982.
[61] K. Aubertin, et al., Mesoscopic characterization of prostate cancer using Raman spectroscopy: potential for diagnostics and therapeutics, BJU Int. 122 (2) (2018) 326–336.
[62] K. Aubertin, et al., Combining high wavenumber and fingerprint Raman spectroscopy for the detection of prostate cancer during radical prostatectomy, Biomed. Opt. Express 9 (9) (2018) 4294–4305.
[63] F.L. Magalhaes, et al., Raman spectroscopy with a 1064-nm wavelength laser as a potential molecular tool for prostate cancer diagnosis: a pilot study, J. Biomed. Opt. 23 (12) (2018) 1–6.
[64] Q. Matthews, et al., Biochemical signatures of in vitro radiation response in human lung, breast and prostate tumour cells observed with Raman spectroscopy, Phys. Med. Biol. 56 (21) (2011) 6839–6855.
[65] T.J. Harvey, et al., Spectral discrimination of live prostate and bladder cancer cell lines using Raman optical tweezers, J. Biomed. Opt. 13 (6) (2008) 064004.
[66] D.K. Medipally, et al., Development of a high throughput (HT) Raman spectroscopy method for rapid screening of liquid blood plasma from prostate cancer patients, Analyst 142 (8) (2017) 1216–1226.
[67] E.M. Kanter, et al., Multiclass discrimination of cervical precancers using Raman spectroscopy, J. Raman Spectrosc. 40 (2) (2009) 205–211.
[68] D. Petersen, et al., Raman fiber-optical method for colon cancer detection: cross-validation and outlier identification approach, Spectrochim. Acta Part AMol. Biomol. Spectrosc. 181 (2017) 270–275.
[69] S.X. Li, et al., Identification and characterization of colorectal cancer using Raman spectroscopy and feature selection techniques, Opt. Express 22 (21) (2014) 25895–25908.
[70] J.F. Wang, et al., Simultaneous fingerprint and high-wavenumber fiber-optic Raman spectroscopy improves in vivo diagnosis of esophageal squamous cell carcinoma at endoscopy, Sci. Rep. 5 (2015).
[71] G. Shetty, et al., Raman spectroscopy: elucidation of biochemical changes in carcinogenesis of oesophagus, Br. J. Cancer 94 (10) (2006) 1460–1464.
[72] T. Kawabata, et al., Optical diagnosis of gastric cancer using near-infrared multichannel Raman spectroscopy with a 1064-nm excitation wavelength, J. Gastroenterol. 43 (4) (2008) 283–290.
[73] S. Duraipandian, et al., Real-time Raman spectroscopy for in vivo, online gastric cancer diagnosis during clinical endoscopic examination, J. Biomed. Opt. 17 (8) (2012).
[74] H. Wills, et al., Diagnosis of Wilms' tumor using near-infrared Raman spectroscopy, J. Pediatr. Surg. 44 (6) (2009) 1152–1158.
[75] X. Feng, et al., Raman biophysical markers in skin cancer diagnosis, J. Biomed. Opt. 23 (5) (2018) 1–10.
[76] P. Piredda, et al., Subcellular raman microspectroscopy imaging of nucleic acids and tryptophan for distinction of normal human skin cells and tumorigenic keratinocytes, Anal. Chem. 87 (13) (2015) 6778–6785.
[77] D.M. Hudson, et al., Glycation of type I collagen selectively targets the same helical domain lysine sites as lysyl oxidase-mediated cross-linking, J. Biol. Chem. 293 (40) (2018) 15620–15627.
[78] J.G. Snedeker, A. Gautieri, The role of collagen crosslinks in ageing and diabetes—the good, the bad, and the ugly, Muscles Ligaments Tendons J. 4 (3) (2014) 303–308.
[79] K. Nowotny, et al., Advanced glycation end products and oxidative stress in type 2 diabetes mellitus, Biomolecules 5 (1) (2015) 194–222.
[80] J.L. Wilkinson-Berka, et al., ALT-946 and aminoguanidine, inhibitors of advanced glycation, improve severe nephropathy in the diabetic transgenic (mREN-2)27 rat, Diabetes 51 (11) (2002) 3283–3289.
[81] V. Thallas-Bonke, et al., Attenuation of extracellular matrix accumulation in diabetic nephropathy by the advanced glycation end product cross-link breaker ALT-711 via a protein kinase C-alpha-dependent pathway, Diabetes 53 (11) (2004) 2921–2930.
[82] S. Genuth, et al., Glycation and carboxymethyllysine levels in skin collagen predict the risk of future 10-year progression of diabetic retinopathy and nephropathy in the diabetes control and complications trial and epidemiology of diabetes interventions and complications participants with type 1 diabetes, Diabetes 54 (11) (2005) 3103–3111.

[83] B.T. Fokkens, A.J. Smit, Skin fluorescence as a clinical tool for non-invasive assessment of advanced glycation and long-term complications of diabetes, Glycoconj. J. 33 (4) (2016) 527–535.
[84] E.G. Gerrits, et al., Skin autofluorescence: a tool to identify type 2 diabetic patients at risk for developing microvascular complications, Diabetes Care 31 (3) (2008) 517–521.
[85] M. Yamanaka, et al., Non-invasive measurement of skin autofluorescence to evaluate diabetic complications, J. Clin. Biochem. Nutr. 58 (2) (2016) 135–140.
[86] A.A. Martin, et al., Detection of advanced glycation end products (AGEs) on human skin by in vivo confocal Raman spectroscopy, in: A. Mahadevan-Jansen, W. Petrich (Eds.), Biomedical Vibrational Spectroscopy 2016: Advances in Research and Industry, San Francisco, United States, 2016, p. 97040S.
[87] F.R. Paolillo, et al., Noninvasive assessments of skin glycated proteins by fluorescence and Raman techniques in diabetics and nondiabetics, J. Biophotonics 12 (1) (2019) e201800162.
[88] S. Mendis, S. Davis, B. Norrving, Organizational update: the world health organization global status report on noncommunicable diseases 2014; one more landmark step in the combat against stroke and vascular disease, Stroke 46 (5) (2015) e121–e122.
[89] F.C. Pascut, et al., Toward label-free Raman-activated cell sorting of cardiomyocytes derived from human embryonic stem cells, J. Biomed. Opt. 16 (4) (2011) 045002.
[90] N.A. Brazhe, et al., Mapping of redox state of mitochondrial cytochromes in live cardiomyocytes using Raman microspectroscopy, PLoS One 7 (9) (2012) e41990.
[91] M. Ogawa, et al., Label-free biochemical imaging of heart tissue with high-speed spontaneous Raman microscopy, Biochem. Biophys. Res. Commun. 382 (2) (2009) 370–374.
[92] N. Nishiki-Muranishi, et al., Label-free evaluation of myocardial infarction and its repair by spontaneous Raman spectroscopy, Anal. Chem. 86 (14) (2014) 6903–6910.
[93] T. Yamamoto, et al., Label-free evaluation of myocardial infarct in surgically excised ventricular myocardium by Raman spectroscopy, Sci. Rep. 8 (1) (2018) 14671.
[94] S. Ohira, et al., Label-free detection of myocardial ischaemia in the perfused rat heart by spontaneous Raman spectroscopy, Sci. Rep. 7 (2017) 42401.
[95] N. Hemanth, et al., Towards the development of a non-bioptic diagnostic technique for eosinophilic esophagitis using Raman spectroscopy, Vib. Spectrosc. 85 (2016) 7–10.
[96] Y. Takahashi, et al., Raman spectroscopy investigation of load-assisted microstructural alterations in human knee cartilage: Preliminary study into diagnostic potential for osteoarthritis, J. Mech. Behav. Biomed. Mater. 31 (2014) 77–85.
[97] M.C. Caraher, et al., Raman spectroscopy predicts the link between claw keratin and bone collagen structure in a rodent model of oestrogen deficiency, Biochim. Biophys. Acta Mol. Basis Dis. 1864 (2) (2018) 398–406.
[98] J.B. Xu, et al., A new approach to find biomarkers in chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) by single-cell Raman micro-spectroscopy, Analyst 144 (3) (2019) 913–920.
[99] G. Devitt, et al., Raman spectroscopy: an emerging tool in neurodegenerative disease research and diagnosis, ACS Chem. Neurosci. 9 (3) (2018) 404–420.
[100] E. Brauchle, et al., Raman microspectroscopy as a diagnostic tool for the non-invasive analysis of fibrillin-1 deficiency in the skin and in the in vitro skin models, Acta Biomater. 52 (2017) 41–48.
[101] T. Huser, et al., Raman spectroscopy of DNA packaging in individual human sperm cells distinguishes normal from abnormal cells, J. Biophotonics 2 (5) (2009) 322–332.

This page intentionally left blank

CHAPTER 16

Dynamics and allostery of human hemoglobin as elucidated by time-resolved resonance Raman spectroscopy

Yasuhisa Mizutani

Department of Chemistry, Graduate School of Science, Osaka University, Toyonaka, Osaka, Japan

16.1 Introduction

Proteins change their structure, therefore they function. When they function, the protein structural changes induced by a reaction at a specific site spatially extend to global structural changes, widely known as allostery [1]. The molecular mechanism of cooperativity in oxygen binding of hemoglobin (Hb) is one of the paradigmatic subjects of allostery.

Human adult hemoglobin A (HbA) has a tetramer structure, exhibiting positive cooperativity in oxygen binding. Fig. 16.1A shows a three-dimensional structure of HbA, which is composed of two α- and two β-subunits ($\alpha_2\beta_2$). Each subunit contains an iron−protoporphyrin IX as a heme group bound to a histidine residue (Fig. 16.1B). The heme iron binds diatomic molecules, such as O_2 and CO, at the opposite side of the proximal histidine. The heme pocket in HbA involves another histidine residue, which is referred as distal histidine. The cooperative oxygen binding of HbA has been explained in terms of a reversible transition between the two quaternary states upon partial ligation/deligation of four hemes [3,4]. X-ray crystallographic studies have demonstrated the presence of two distinct quaternary structures that correspond to the low-affinity (T or tense) and high-affinity (R or relaxed) states [5]. The completely unligated (deoxy) structure typically adopts the T state, while the fully ligated form adopts the R state. The dissociation of the diatomic ligands from heme in HbA, which is a highly localized perturbation, however, initiates a sequence of propagating structural changes in the subunits and intersubunit contacts, proceeding from the R state to the T state. Thus the protein dynamics of the R−T allosteric transition need to be elucidated to fully understand the allosteric mechanism of HbA, helping to establish general principles related to protein allostery.

Vibrational Spectroscopy in Protein Research.
DOI: https://doi.org/10.1016/B978-0-12-818610-7.00016-5

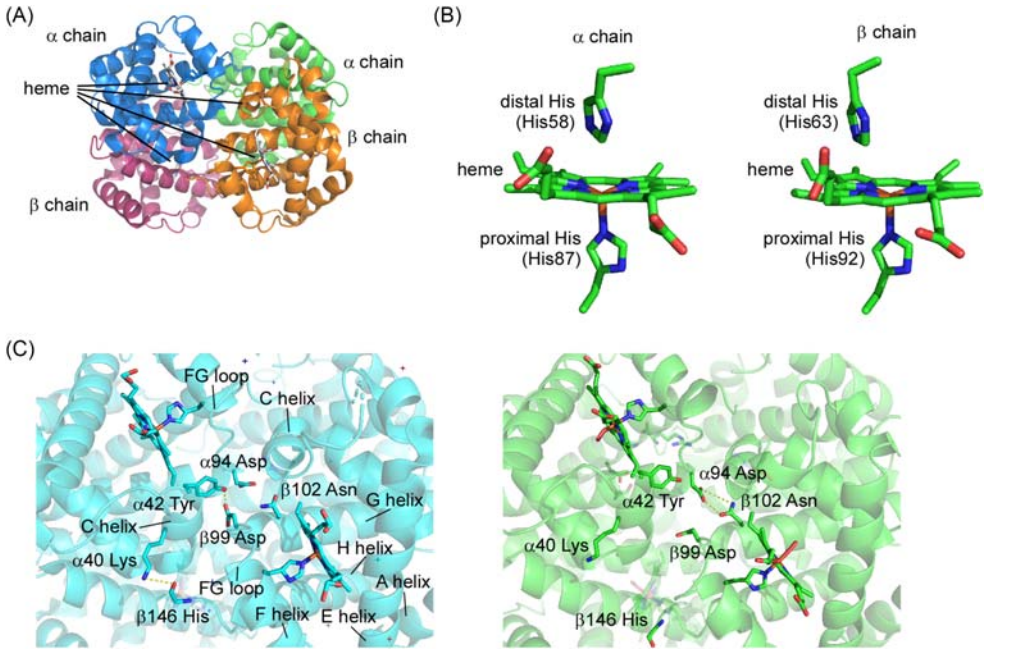


Figure 16.1

Structure of human adult hemoglobin A (HbA). (A) Three-dimensional structure of the deoxy form of HbA based on protein databank structure (PDB, 2HHB) [2]. (B) Heme structure of the α (*left*) and β subunits (*right*). (C) Intersubunit interactions in the T (*left*) and R structures (*right*).

In this chapter, we describe our work on the protein dynamics of HbA elucidated by time-resolved resonance Raman (RR) spectroscopy, summarizing results and discussion published elsewhere [6–10]. RR spectroscopy provides vibrational spectra as a sensitive structure probe for proteins. One advantage of RR spectroscopy is an enhancement of the scattering due to the resonance effect. When the excitation wavelength is matched to an electric-dipole-allowed electronic transition of the target molecule, the intensity of the associated Raman scattering is greatly enhanced. This effect arises from the electron–nuclear coupling of the electronic and vibrational transitions [11]. The resonance enhancement is particularly important in studying protein structure because sample concentrations can be as low as 10^{-6} M and scattering from the nonresonant portions of protein and buffer will not complicate the vibrational spectra. Vibrational modes that show enhancement are localized in the group of atoms that give rise to the electronic transition. We can thus observe specific parts of macromolecules by tuning excitation wavelength to match the electronic transition wavelength. Its unique sensitivity and selectivity make it possible to characterize the structure in various parts in protein, advancing the understanding of how allostery is achieved. For example, RR spectroscopy using ~400-nm light helped elucidate the molecular structures and reaction mechanisms of heme, which has strong electronic transitions in the ~400 nm region [12–14]. Pulsed lasers, which

extend the ability to investigate time domains, now allow time-resolved RR measurements to elucidate the temporal evolution of protein structures. The time resolution of vibrational spectroscopy is determined, in principle, by the vibrational dephasing time, that is $\sim$1 ps in condensed phases. With the development of ultrafast laser technology and photodetectors, recent methodological advances in RR spectrometers allow the observation of protein dynamics with unprecedented detail, revealing intriguing aspects of the allosteric mechanism.

16.2 Time-resolved resonance Raman spectrometers

Pump and probe pulses are required in time-resolved RR measurements. The wavelength of the pump pulse should fall within the electronic absorption band of the prosthetic group to be photoexcited. At the same time, the wavelength of the probe pulse needs to be close to the wavelength of the electronic transitions of the particular portion of the protein to be investigated, so that resonance enhancement of the Raman band can occur. In addition, the spectral width of the probe pulse must be narrow enough to obtain well-resolved vibrational bands. Therefore two independently tunable light sources for the pump and probe beams are necessary to obtain time-resolved RR spectra of a wide variety of molecules with high signal-to-noise ratios (SNRs) within a reasonable measurement time. Furthermore, repetition rates in the kilohertz range are desirable for practical purposes.

We developed spectrometers containing widely tunable light sources that fulfill the requirements for small timing jitter, appropriate repetition rates, and wavelength tunability of pulses applicable to time-resolved RR spectroscopy with picosecond resolution [15,16]. In order to obtain time-resolved RR spectra of a wide variety of molecules with high SNRs within a reasonable measurement time, two independently tunable light sources with a high repetition rate for the pump and probe beams were developed using a 1-kHz picosecond Ti: sapphire laser/regenerative amplifier system. The repetition rate of 1 kHz made it possible to measure spectra with low peak power. Details of the spectrometers were described elsewhere [10,17,18].

The longest delay time of the abovementioned spectrometer is 10 ns, which is limited by the length of the optical delay stage. For the delay times longer than 10 ns, we constructed a time-resolved RR spectrometer by combining two nanosecond Q-switched lasers operating at 1 kHz [19]. The timing between the pump and probe pulses was adjusted with a computer-controlled pulse generator via a general purpose interface bus (GPIB) interface.

The amount of purified protein sample solution available for spectroscopic measurements is often limited. Therefore the laser system must have greater stability and the detection system needs to possess higher sensitivity to obtain high signal-to-noise spectra. We also constructed devices for sample circulation for a limited amount of protein sample [20,21].

16.3 Primary structural dynamics of hemoglobin A following the ligand dissociation

16.3.1 Ligand photodissociation in hemoglobin A

Ultrafast absorption studies of HbA revealed that the excited-state potential energy surface of the ligated heme is dissociative regarding a ligand, so as to be able to photoinitiate a structural change from the ligated form to the deligated form. The dissociated heme is formed within 50 fs [22–24]. Protein reorganization in response to changes at the heme is expected to occur more slowly. By using CO as the ligand rather than O_2, it is possible to avoid the geminate recombination complicating the reaction dynamics in the picosecond regime [25]. These features enabled us to trigger optically the structural changes of a protein and to follow the accompanied structural relaxation using optical probes that are sensitive to different aspects of the protein's motion without complicated side reactions. For this reason, an enormous body of work has been accumulated on the relaxation processes of HbA.

16.3.2 Structural dynamics of heme in hemoglobin A upon the photodissociation

Structural dynamics of HbA following photodissociation were examined using RR spectroscopy. We first focus on the structural changes of heme [6] because heme is the site in which the photodissociation occurs. Fig. 16.2A and B show time-resolved RR spectra of CO-photodissociated HbA in the 180–850 cm^{-1} and 1030–1700 cm^{-1} regions, respectively, for delay times from −5 to 1000 ps of the probe pulse with respect to the pump pulse. The contribution from unphotodissociated species has been eliminated in the time-resolved spectra. The 1300–1700 cm^{-1} region of the spectra contains only the RR bands due to the heme in-plane vibrations. In the spectrum of the deoxy form (deoxyHb), they are ν_4 band at 1357 cm^{-1}, ν_3 band at 1472 cm^{-1}, ν_2 band at 1566 cm^{-1}, and ν_{10} band at 1623 cm^{-1}. The mode assignments made by Abe were adopted [26]. RR bands in the 200–800 cm^{-1} region contain out-of-plane modes of porphyrin ring, substituent modes, and iron–ligand modes as well as in-plane modes. The band at 215 cm^{-1} in the spectrum of the deoxy form is assigned to the stretching mode of the covalent bond between the heme iron and the N_ε of the proximal histidine, ν(Fe–His) [27]. Others are γ_7 bands at 300 cm^{-1}, ν_8 band at 341 cm^{-1}, and $\delta(C_\beta C_c C_d)$ band at 365 cm^{-1} [28]. The band at 673 cm^{-1} is due to a breathing-like mode of the porphyrin inner ring, ν_7.

The time-resolved RR spectra at 4 ps resemble those of the deoxy form. It is known that the heme adopts a planar form in the CO-bound state and becomes domed in the deoxy state. The frequencies of ν_2, ν_3, and ν_{10} modes are known to be sensitive to the heme planarity. Therefore the observed spectra indicate that the photodissociated heme has mostly relaxed to the domed structure as in the deoxy form within the instrument response time ($\sim$2 ps).

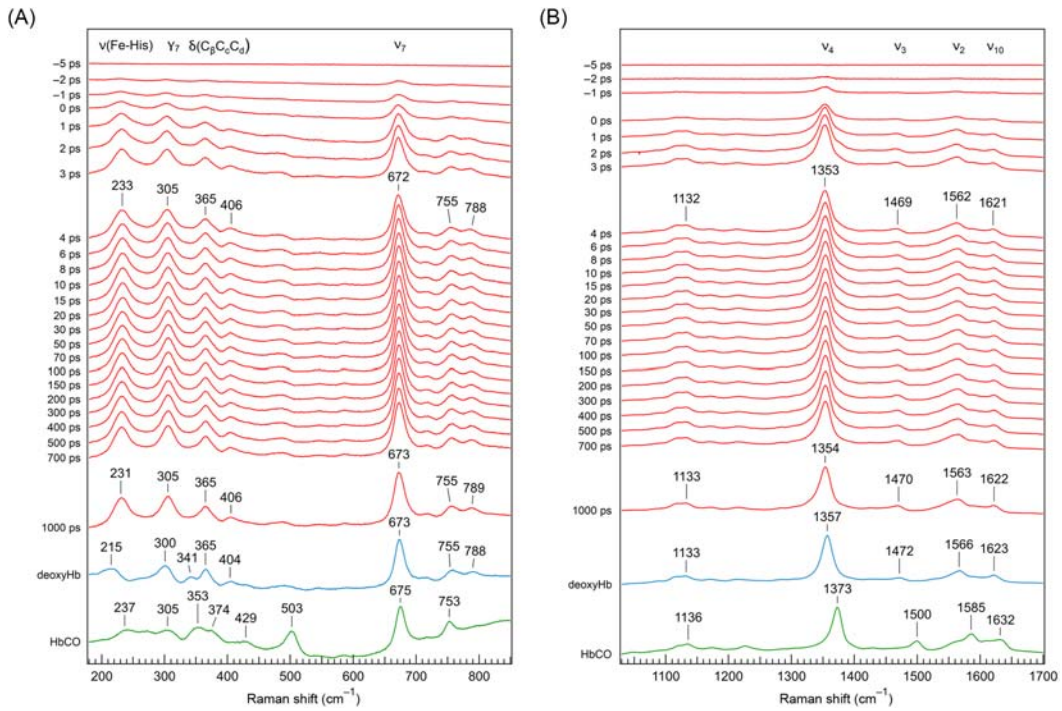


Figure 16.2

Time-resolved resonance Raman (RR) spectra of hemoglobin A (HbA) following CO photolysis. Spectra of the equilibrium states of deoxyHb and HbCO are depicted at the bottom for comparison. Probe and pump wavelengths were 442 and 540 nm, respectively. (A) Time-resolved RR spectra in the 150 to 850-cm^{-1} region. (B) Time-resolved RR spectra in the 1100 to 1700-cm^{-1} region. Source: *Reproduced from Y. Mizutani, M. Nagai, Ultrafast protein dynamics of hemoglobin as studied by picosecond time-resolved resonance Raman spectroscopy, Chem. Phys. 396 (2012) 45–52, with permission from Elsevier B.V.*

However, a close comparison of the spectra showed that the ν_2, ν_3, and ν_4 bands of the transient species at 1000 ps were 2–3 cm^{-1} lower in frequency than those of the deoxy form (Fig. 16.2A), which is consistent with the results reported by Friedman et al. [29] and Dasgupta and Spiro [30] using a *nanosecond* pulse. Differences of the transient spectra from the deoxy spectrum also were observed for the ν(Fe–His), γ_7, and ν_8 bands in the time-resolved RR spectra of HbA (Fig. 16.2B). First, the ν_8 band at 341 cm^{-1} in deoxyHb was absent in the photoproduct. Second, the frequencies of the ν(Fe–His) and γ_7 modes at 231 and 305 cm^{-1}, respectively, of the photoproduct at 1000 ps differed from those of the deoxy form. The difference between the spectrum at 1000 ps and that of the deoxy form showed that HbA adopts a metastable structure within the instrument response time and hardly changes in the subnanosecond to nanosecond time region. The observed spectral differences provided insight into the structure of the metastable form in the picosecond time

region. The form has a more disordered orientation of propionates and a less strained environment compared to the deoxy form [6]. The latter discovery was consistent with the experimental observation that the ν(Fe−His) frequency of the metastable form was higher than the deoxy form. The time-resolved RR measurements revealed the ultrafast structural changes and the structure of the metastable state in the heme group.

16.3.3 Comparison of the primary structural dynamics in hemoglobin A to those in relating proteins

The large amplitude motions at the quaternary level, which form a communication link at the subunit interface, are driven by changes of the tertiary structure upon ligation. In this respect, myoglobin (Mb), which is structurally very similar to a subunit of HbA, serves as a model system for the tertiary relaxation processes. Thus we compared the structural dynamics of Mb to those of HbA and found distinct differences in time-resolved RR spectra of the photolyzed species of HbA and Mb. The transient spectra of HbA showed differences in the ν(Fe−His), γ_7, and ν_8 and bands of the deoxy spectra, while those of Mb did not. RR bands of the isolated subunits of HbA exhibited similar temporal behavior as those of HbA, but were completely different from those of Mb [6]. In solution, isolated α and β chains are supposed to form monomers/dimers and tetramers, respectively, and have no cooperativity in the oxygen binding. Thus the difference in the dynamics between HbA and Mb was not due to intersubunit interactions, such as tetramer formation or cooperativity, but to the intrinsic nature of the HbA subunits. The tertiary structural change in the nanosecond region may be important for allostery of HbA to bridge the structural changes on the subnanosecond scale with the quaternary structural changes in the microsecond region.

16.3.4 Structural response of heme pocket in HbA upon the photodissociation

We revealed that the structural change to the domed form takes place within a few picoseconds. The next question is how the surrounding protein moiety responds to this ultrafast structural change in heme. We examined temporal changes of the ν(Fe−His) frequency because it is a good spectral probe to observe the protein response. The proximal Fe−His linkage in heme proteins plays a pivotal role in communicating structural changes of the heme group and the protein moiety, thus affecting its biological properties [31]. The deoxy form within the quaternary T state has been reported to have a band frequency of 216 cm^{-1}, whereas the R state of the modified deoxy NES-des(Arg141α)-Hb exhibited a Raman band at a frequency of 221 cm^{-1} [32]. Fig. 16.3 shows the temporal changes in the ν(Fe−His) frequency of HbA and Mb following the CO dissociation. The ν(Fe−His) mode of HbA exhibited a 2-cm^{-1} downshift with a time constant of about 284 ± 38 ps, suggesting a structural change in the heme pocket following the ligand dissociation. The ~300-ps process for the Fe−His linkage is the fastest structural response reported for the

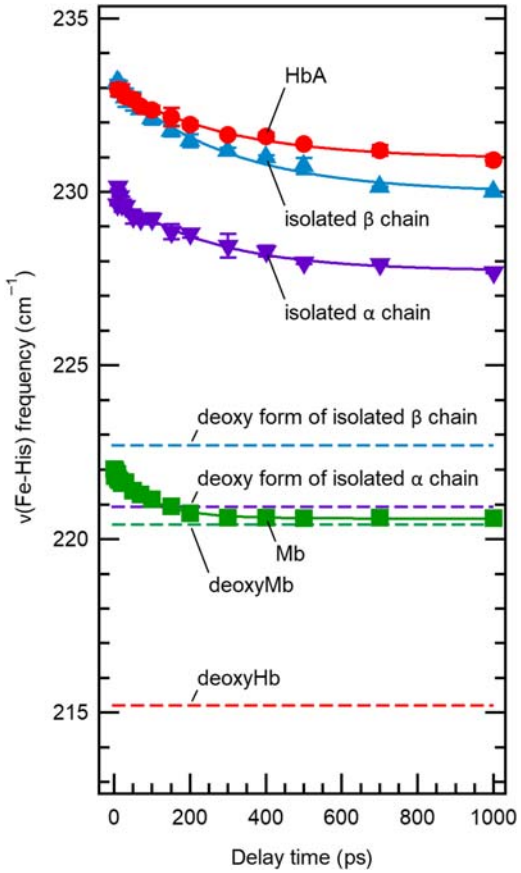


Figure 16.3

Temporal changes in ν(Fe−His) frequencies of hemoglobin A (HbA) (*red circles*) and isolated α (*blue triangles*) and β chains (*purple inverted triangles*) in 20 mM sodium phosphate buffer at pH 7.5. Temporal shift in the ν(Fe−His) band was fitted using a single exponential function, yielding time constants of 284 ± 38, 256 ± 35, and 298 ± 42 ps for HbA, isolated α chains, and isolated β chains, respectively. Data for Mb (*green squares*) are shown for comparison. Broken lines show band positions of the equilibrium deligated form (i.e., deoxy form). Source*: Reproduced from Y. Mizutani, M. Nagai, Ultrafast protein dynamics of hemoglobin as studied by picosecond time-resolved resonance Raman spectroscopy, Chem. Phys. 396 (2012) 45−52, with permission from Elsevier B.V.*

Fe−His linkage of the HbA photoproduct. The red trace in Fig. 16.3 shows temporal changes of the ν(Fe−His) frequency of Mb following the photodissociation. The size and time constant of the frequency change were found to be 1.7 ± 0.1 cm^{-1} and 106 ± 14 ps, respectively. Although the size of the frequency shifts of the ν(Fe−His) band was similar for HbA and Mb, the time constants of the frequency shift of HbA was about threefold slower than that of Mb. This indicates the difference of the primary protein response in HbA from that in Mb.

16.4 Intersubunit communication: direct observation for the Perutz mechanism

16.4.1 Propagation of structural changes between subunits

The changes of the tertiary structure upon the ligation/deligation drive the large amplitude motions at the quaternary level, which form a communication link at the subunit interface. A frequency difference in the ν(Fe−His) mode as large as 16 cm^{-1} observed between the metastable state at 1000 ps and deoxyHb implies a subsequent shift that occurred later (> nanoseconds). It was reported that the subsequent ν(Fe−His) shift occurred stepwise in the nanosecond to microsecond region [8,9,19,33]. These changes were associated with the tertiary and quaternary structure changes induced by the structural changes of heme. Studies using time-resolved ultraviolet RR spectroscopy and time-resolved wide-angle X-ray scattering showed that, upon the CO dissociation, the subunit interface underwent structural changes in microseconds to tens of microseconds [34−36]. However, how the heme of one subunit responds to the ligand dissociation of another heme on a neighboring subunit has not been fully elucidated, despite the importance of understanding this intersubunit communication to clarify the molecular mechanism of HbA cooperativity.

Discriminating between the heme groups in the α and β subunits spectroscopically is difficult because both subunits have identical heme structures. Our strategy to investigate intersubunit communication through the α−β subunit interface involved observation of structural changes in the Ni-heme of the β subunit upon CO the dissociation in the α subunit using carbonmonoxy α_2(Fe)β_2(Ni) [α_2(Fe−CO)β_2(Ni)]. Hybrid Hb α_2(Fe−CO)β_2(Ni) complexes possess two characteristics that make them extremely useful for investigating the properties of individual subunits in time-resolved experiments. First, the resonance enhancement allows selective observation of the Fe-heme and Ni-heme spectra in the hybrid Hb because the peak wavelengths of the Soret band are different for the deoxy Hb and the Ni-substituted Hb. We demonstrated that, because of Soret bands at 430 and 419 nm, RR spectra excited at 436 and 418 nm selectively probe heme and Ni-heme, respectively [8]. Second, no photochemical reaction occurs in the Ni-heme because it does not bind a ligand. Consequently, the RR spectrum of the Ni-heme is a good spectroscopic probe for investigating intersubunit communication. We examined the changes in the RR spectra of the Ni-heme of the β subunits upon the CO dissociation in the α subunits using α_2(Fe−CO)β_2(Ni) [8].

16.4.2 Structural dynamics of subunits with ligated heme in half-ligated hemoglobin

Fig. 16.4A shows time-resolved RR spectra of the α heme in α_2(Fe−CO)β_2(Ni) following CO photolysis, measured with a probe pulse of 436 nm. The spectrum for a 50-ns delay

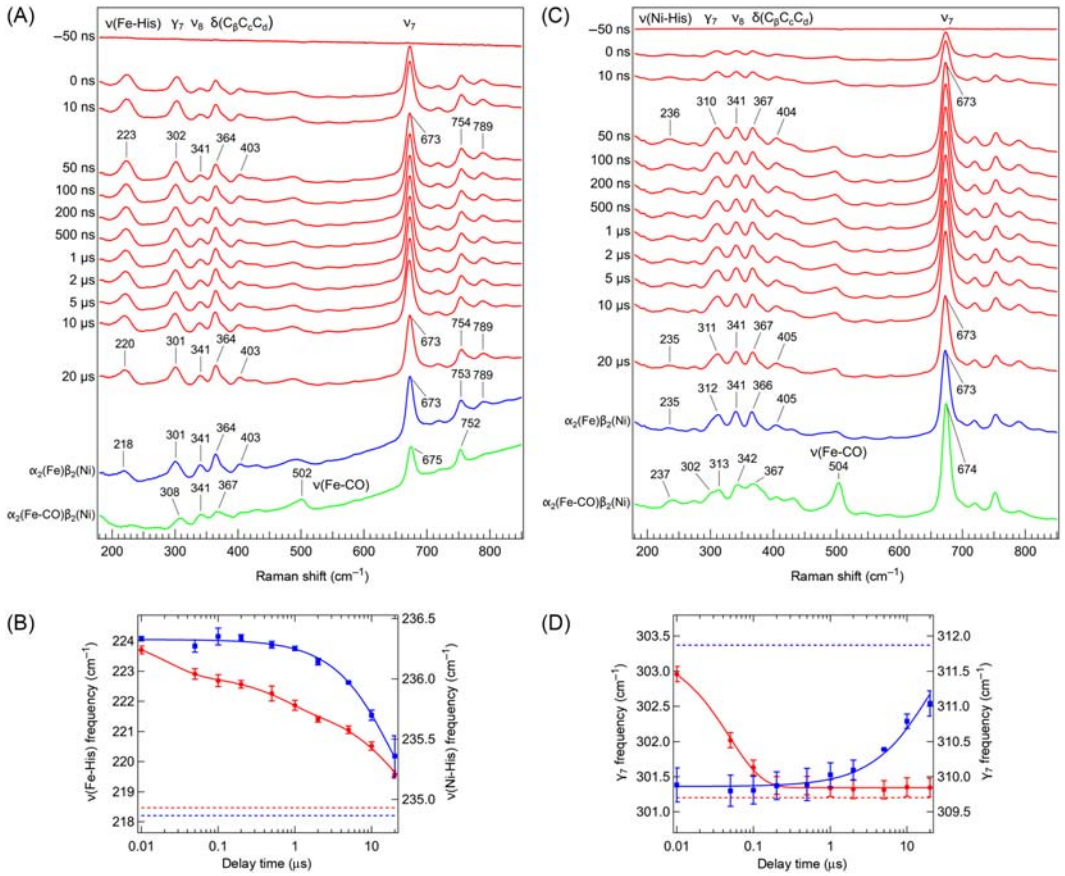


Figure 16.4

Time-resolved resonance Raman (RR) spectra of hybrid Hb α_2(Fe-CO)β_2(Ni) in 50 mM sodium phosphate buffer at pH 8.0 following CO photolysis with 532-nm pulse. Spectra of the equilibrium states of the deoxy and CO-bound forms are shown at the bottom for comparison. Time-resolved difference spectra were generated by subtracting the probe-only spectrum from the pump-probe spectra at each delay time. (A) Time-resolved RR spectra in the 180−850 cm^{-1} region measured with a probe pulse of 436 nm. (B) Time-resolved RR spectra at 180−850 cm^{-1} measured with a probe pulse of 418 nm. (C) Logarithmic time plots of frequencies of ν(Fe−His) and ν(Ni−His) of α heme and β Ni-heme of α_2(Fe)β_2(Ni) following CO dissociation. Solid circles (*red*) and triangles (*blue*) represent frequencies of RR bands of α heme and β Ni-heme, respectively. Solid lines indicate the best fit obtained using an exponential function or a sum of exponential functions. The red and blue broken lines show values of α heme and β Ni-heme, respectively, in deoxy α_2(Fe)β_2(Ni). (D) Logarithmic time plots of frequencies of γ_7 bands of α heme and β Ni-heme of α_2(Fe)β_2(Ni) following CO dissociation. Solid circles (*red*) and triangles (*blue*) represent frequencies of RR bands of α heme and β Ni-heme, respectively. Solid lines indicate the best fit obtained using an exponential function or a sum of exponential functions. The red and blue broken lines show values of α heme and β Ni-heme, respectively, in deoxy α_2(Fe)β_2(Ni). Source*: Reproduced from K. Yamada, H. Ishikawa, M. Mizuno, N. Shibayama, Y. Mizutani, Intersubunit communication via changes in hemoglobin quaternary structures revealed by time-resolved resonance Raman spectroscopy: direct observation of the Perutz mechanism, J. Phys. Chem. B 117 (2013) 12461−12468, with permission from American Chemical Society.*

contained bands arising from vibrations of the heme at 223 [ν(Fe−His)], 302 (γ_7), 341 (ν_8), 364 [$\delta(C_\beta C_c C_d)$], 403 [$\delta(C_\beta C_a C_b)$], and 673 cm^{-1} (ν_7). The time-resolved RR spectra at 50 ns differed from the spectrum of the deoxy form. The frequencies of the ν(Fe−His) and γ_7 bands in the 50-ns time-resolved RR spectra were higher than those in the spectrum of the deoxy form. The spectra exhibited a frequency shift of the ν(Fe−His) and γ_7 bands as the delay time increased. The ν(Fe−His) band was observed at 223 cm^{-1} at 50 ns and downshifted to 220 cm^{-1} at 20 μs. The γ_7 frequency of 302 cm^{-1} at 50 ns shifted to 301 cm^{-1} at 20 μs. The relative intensity of the ν_8 band to the $\delta(C_\beta C_c C_d)$ band increased as the delay time increased. The weaker band intensities observed at 20 μs are due to geminate recombination. Similar temporal changes in the RR spectra were observed for the dissociated form for HbA from blood and recombinant hemoglobin A (rHb) [19].

Fig. 16.4B displays the temporal evolution of the frequencies of the ν(Fe−His) and γ_7 bands of α heme of α_2(Fe)β_2(Ni) following the CO dissociation. The frequency shifts of the ν(Fe−His) and γ_7 bands of α heme in α_2(Fe)β_2(Ni) were fitted using triple-exponential functions and single-exponential functions, respectively. The time constants of the evolutions are summarized in Table 16.1. The temporal evolution of the frequencies of ν(Fe−His) of the α heme in α_2(Fe)β_2(Ni) was similar to that of α and β heme in the rHb [19]. This indicates that structural changes in the α subunit of α_2(Fe)β_2(Ni) are similar to those of the rHb and that replacement of heme with Ni-heme in the β subunit does not influence the allosteric dynamics of the rHb.

The three steps of the ν(Fe−His) shift are consistent with the picture of HbA dynamics obtained by time-resolved absorption [37] and RR [6,34,38] spectroscopic studies, which have established a series of intermediates following the CO dissociation. In the primary intermediate, the heme structure changes to a metastable domed form within 2 ps. The heme pocket rearranges to generate the B intermediate. Further protein relaxation leads to

Table 16.1: Time constants of the temporal evolution of Raman bands of α_2(Fe)β_2(Ni)[a] and recombinant hemoglobin A (rHb).[b]

	Band	τ_1 (ns)	τ_2 (μs)	τ_3 (μs)
α subunit in α_2(Fe)β_2(Ni)	ν(Fe−His)	20.2 ± 10.1 (24 ± 5%)[c]	0.688 ± 0.207 (19 ± 2%)	19.5 ± 1.6 (57 ± 2%)
	γ_7	48.3 ± 1.9		
β subunit in α_2(Fe)β_2(Ni)	ν(Ni−His)			18.0 ± 0.7
	γ_7			18.9 ± 1.5
rHb[b]	ν(Fe−His)	20.0 ± 7.1 (23 ± 4%)	0.631 ± 0.243 (16 ± 2%)	17.3 ± 1.3 (61 ± 2%)
	γ_7	43.4 ± 5.2		

[a]The temporal evolution is fitted to an exponential function or a sum of exponential functions.
[b]From Ref. [19].
[c]Numbers in parentheses represent the relative contribution of the component.

the subsequent intermediate, R_{deoxy}. In the transition from B to R_{deoxy}, the distal E helix displaces toward the heme plane, probably driven by the motion of the proximal F helix in response to Fe−His bond relaxation [34,38,39]. The first component of the ν(Fe−His) shift is due to the step in which the heme structure changes and proximal strain is released via the postulated rotation of the enhancement factor (EF) helical section. The spectral changes of the γ_7 and ν_8 bands in nanoseconds observed in our study are likely associated with EF helical motion. In the step of the second component, interhelical hydrogen bonds are restored following the motion of the A and H helices to generate the R_S intermediate. The third component is assigned to the frequency shift due to the quaternary change to the T structure.

16.4.3 Structural dynamics of subunits with unligated heme in half-ligated hemoglobin

The time-resolved RR spectra of β Ni-heme in α_2(Fe−CO)β_2(Ni) as measured with a probe pulse of 418 nm is shown in Fig. 16.4C. The spectrum at 50-ns delay contains bands arising from vibrations of the heme at 236 [ν(Ni−His)] [40], 310 (γ_7), 341 (ν_8), 367 [$\delta(C_\beta C_c C_d)$], 404 [$\delta(C_\beta C_a C_b)$], and 673 cm^{-1} (ν_7). The ν(Ni−His) band downshifted to 235 cm^{-1} at 20 μs. The two end-point structures showed the ν(Ni−His) band at 237 and 235 cm^{-1} for α_2(Fe−CO)β_2(Ni) and α_2(Fe)β_2(Ni), respectively. Thus the ν(Ni−His) frequency decreased monotonously from that of α_2(Fe−CO)β_2(Ni) to that of α_2(Fe)β_2(Ni). It should be noted that the ν(Ni−His) and γ_7 frequencies are different between wild-type (WT) and βD99N NiHb: the ν(Ni−His) band was observed at 238 cm^{-1} for NiHb(βD99N), while it was observed at 235 cm^{-1} for WT NiHb [8]. The γ_7 band was observed at 310 cm^{-1} for NiHb (βD99N), while it was observed at 312 cm^{-1} for WT NiHb. The deoxy form of Hb Kempsey, which is an abnormal HbA with a βD99N mutation, adopts the R state [41]. Therefore these data showed that the frequencies of the ν(Ni−His) and γ_7 bands were different between the R and T states and that the frequency shifts observed in the time-resolved spectra for α_2(Fe−CO)β_2(Ni) reflect a quaternary structural transition from the R to the T structure.

Fig. 16.4D displays the temporal evolution of the frequencies of the ν(Ni−His) and γ_7 bands of β Ni-heme of α_2(Fe)β_2(Ni) following the CO dissociation. No temporal changes in the ν(Ni−His) and γ_7 bands of β Ni-heme were observed earlier than 1 μs, which is consistent with the lack of structural change of Ni-heme upon excitation by the pump pulse. Later than 1 μs, frequency shifts were observed, indicating structural changes in the β subunit induced by the structural changes in the α subunits. The changes in the ν(Ni−His) and γ_7 frequencies of the β Ni-heme in α_2(Fe)β_2(Ni) were fitted using a single-exponential function. Time constants of the evolutions of the ν(Ni−His) and γ_7 frequencies of β heme were 18.0 and 18.9 μs, respectively, demonstrating that the strain imposed on the Ni−His bond in the β subunit strengthens at $\sim$20 μs changing upon the CO photolysis in the

α subunit. In addition, note that the 20-μs component was commonly observed both in the ν(Fe−His) band of the α subunit and the ν(Ni−His) band of the β subunit, meaning that the bonds between heme and the proximal histidine change concertedly in both subunits upon the R−T quaternary change. Our data not only show that a structural change in the heme of the neighboring subunit occurs in ∼20 μs, but also that the changes include the proximal tension, which is a key factor controlling the oxygen affinity of HbA.

16.4.4 Perutz's strain model

Perutz put the allosteric mechanism of HbA, proposed by Monod, Wyman, and Changeux (MWC) [1], on a stereochemical basis [5,42]. He interpreted the MWC two-state model in terms of an equilibrium between the T and R structures. It was assumed that the T−R equilibrium is governed primarily by the positions of the iron atoms relative to the porphyrin: out-of-plane in the deoxy form and in-plane in the oxy form. In turn, the tension exerted by the intersubunit interactions was transmitted to the Fe−His bond in the T structure and restrained the movement of iron atoms into the porphyrin plane necessary for oxygen binding in the R structure. Thus the heme group and the subunit interface are bidirectionally associated with each other. Matsukawa et al. studied various abnormal Hbs that had a replacement amino acid residue in the subunit interface [43]. They observed a correlation between the ν(Fe−His) frequency and the first Adair constant K_1, which is related to the Gibbs energy of the first oxygenation step. This finding supported the Perutz's strain model from the viewpoint of equilibrium [5,42].

Common time constants of ∼20 μs observed for the ν(Fe−His) (α subunits) and ν(Ni−His) bands (β subunits) in our study support Perutz's strain model from the viewpoint of protein dynamics, that is, the structural change in the subunit interface from the R to T state induces tension in the Fe−His bonds synchronously. Accordingly, the observed synchronous changes of the ν(Fe−His) and ν(Ni−His) bands upon the R−T transition can be regarded as real-time observations of the Perutz mechanism.

16.5 Dynamical coupling of tertiary and quaternary structures in hemoglobin A

16.5.1 Hemoglobin mutants of fixed quaternary structure

In HbA, the quaternary structural changes are induced by the tertiary structural changes and vice versa. Thus the coupling between quaternary structure and tertiary structure is essential in the allostery of HbA. To understand the coupling mechanism, it is intriguing to compare the tertiary structural dynamics of HbA fixed in the R and T states. Accordingly, we examined RR spectra of HbA mutants which are restricted in the T and R quaternary states.

Hb Kansas (β102 Asn→Thr) is a well-studied low O_2 affinity variant [44–46]. Replacement of Asn102 at the $\alpha_1\beta_2$ interface inhibits stabilization of the R structure by the hydrogen bond formation with α94 Asp (Fig. 16.1C). Hb Kansas is an interesting HbA variant whose quaternary structure can be switched in its ligated forms from the R to the T state by inositol hexaphosphate [45,47–50]. In contrast, Hb Kempsey (β99 Asp→Asn) is a high O_2 affinity variant. The amino acid replacement perturbs intersubunit interactions by preventing the formation of a hydrogen bond between β99 Asp and α42 Tyr [51,52], which normally stabilizes the T state, thereby shifting the quaternary equilibrium toward the R state (Fig. 16.1C) [41,53,54]. Accordingly, we observed the structural dynamics of the βN102T and βD99N mutants, which are restricted in the T and R quaternary states, respectively, to examine the tertiary structural changes in fixed quaternary states.

Fig. 16.5A and B show the temporal evolution of the frequency of the ν(Fe–His) band (A) and intensity of the ν_8 band (B) of WT rHb and the βD99N mutant, respectively, following the CO dissociation. The intensity of the ν_8 band is shown as a value relative to the intensity of the ν_7 band. The most striking difference between WT rHb and the βD99N mutant was observed in the frequency and the relative intensity in 10–100 μs. The frequency and the relative intensity of WT rHb did not converge to those of the deoxy form at 100 μs which is the longest delay time in our experiments. In contrast, the frequency and the relative intensity of the βD99N mutant converged to those of the deoxy form at 50 μs, indicating that the structural changes following the CO dissociation are completed in about 50 μs for the βD99N mutant. We fitted the frequency shifts of the ν(Fe–His) bands and the relative intensity changes of the ν_8 bands using a triple-exponential function with common values of the time constants for each sample because the frequency shifts of the ν(Fe–His) bands were reasonably described using three exponential components in the previous studies of WT rHb and HbA from blood [8,19,33].

Fig. 16.5C and D show the temporal evolutions of the frequency of the ν(Fe–His) band (A) and intensity of the ν_8 band (C) of WT rHb and the βN102T mutant, respectively, following the CO dissociation. The evolutions of the ν(Fe–His) and ν_8 bands were completely different between the βN102T mutant and WT rHb. The Raman bands of the βN102T mutant changed much more slowly than those of WT rHb. More than a half fraction of the spectral changes occurred later than 100 μs. This means that the tertiary change of the protein fixed in the T structure is largely decelerated.

16.5.2 Tertiary structural changes dependent on R and T quaternary structures

Spiro and coworkers proposed a model of the R–T structural transition of HbA following the CO dissociation [34]: the photodissociation of HbCO to produce a primary intermediate, B, in which strain is stored in an expansion of the core size of the porphyrin macrocycle and in compression of the Fe–His bond [38,55,56]. This strain is released in the subsequent

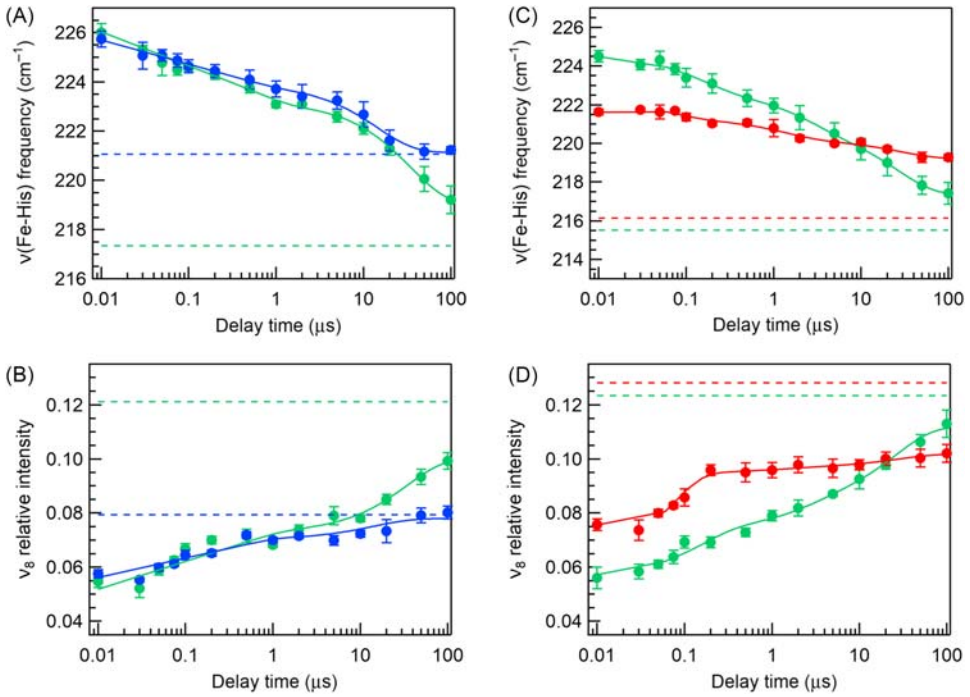


Figure 16.5

Logarithmic time plots of the frequency of ν(Fe—His) band and relative intensity of the ν_8 band of the hemoglobin (Hb) mutant and wild-type recombinant hemoglobin A (WT rHb). (*left*) The frequency of ν(Fe—His) band (A) and relative intensity of ν_8 band (B) of the βD99N mutant and WT rHb in 50 mM Tris—HCl buffer at pH 8.5 containing 100 mM NaCl. The blue and green circles represent the observed values for the βD99N mutant and WT rHb, respectively. The intensity of the ν_8 band was shown in relative value to the intensity of the ν_7 band. The solid lines indicate the best fit obtained using a triple-exponential function with common values of the time constants for each sample. The time constants of the evolutions are summarized in Table 16.1. The blue and green broken lines show the values of the deoxy forms of the βD99N mutant and WT rHb, respectively. (*right*) The frequency of ν(Fe—His) band (A) and relative intensity of ν_8 band (B) of the βN102T mutant and WT rHb in 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid buffer at pH 6.4 containing 1 mM inositol hexaphosphate (IHP). The red and green circles represent the observed values for the βN102T mutant and WT rHb, respectively. The red and green broken lines show the values of the deoxy forms of the βN102T mutant and WT rHb, respectively. Source*: Reproduced from Ref. 9 with permission from the PCCP Owner Societies.*

intermediate, R_{deoxy}, via the postulated clamshell rotation of the E and F helices and breaking of the A···E and F···H interhelical hydrogen bonds. The interhelical hydrogen bonds are then restored by following motions of the A and H helices in hundreds of nanoseconds. These tertiary structural changes lead to the incipient formation of the T quaternary contacts at the critical $\alpha_1\beta_2$ subunit interface in the microsecond regime.

First, the subunits rotate into position to form the "hinge" contacts between the β subunit C helix and the α subunit FG corner in a few microseconds. Later, the "switch" contacts between the α subunit C helix and the β subunit FG corner are formed in tens of microseconds. β99 Asp and β102 Asn are in the β subunit FG corner and participate in the "switch" contacts.

In the βN102T mutant, we found a remarkable difference for the temporal changes of the ν(Fe−His) frequencies and ν_8 intensities from those of WT rHb. This means that the tertiary structural changes in the βN102T mutant were much slower than those in WT rHb and βD99N mutant, demonstrating the difference in the tertiary structural changes between in the R and T quaternary states. It is likely that the cleavage of the β102 Asn···α94 Asp hydrogen bond affects the intrasubunit interactions and decelerates the tertiary structural changes.

The model by Spiro and coworkers suggests the interhelical hydrogen bonds in the EF helices are broken and restored prior to the formation of the T quaternary intersubunit contacts [34]. This suggests that the cleavage of the β102 Asn···α94 Asp intersubunit hydrogen bond decelerates the changes of the interhelical hydrogen bonds in the EF helices. Elimination of the A···E and F···H interhelical hydrogen bonds had no effect on ligand affinity or cooperativity, but accelerated "hinge" contact formation, while decelerating "switch" contact formation [39,57]. These results mean that the allosteric kinetic response between the interhelical hydrogen bonds in the EF helices, which hold the heme group, and the intersubunit hydrogen bonds, which regulate the quaternary state, is bidirectional. Such a bidirectional kinetic response would be essential to the allosteric effect between heme and intersubunit contacts.

16.6 Differences between protein dynamics upon the dissociation of O_2 and CO

Thus far, the time-resolved RR studies of HbA were discussed based upon CO dissociation. Along with the studies of Gibson [58], ligand photolysis techniques have been applied to numerous kinetic studies for heme proteins [22,23,59−61]. However, time-resolved RR studies on HbA were conducted mainly for structural dynamics upon the CO dissociation. Very few studies on structural dynamics of the O_2 dissociation have been reported [62,63], even though the physiological ligand of HbA is O_2. This is due to several reasons. First, by using CO as the ligand rather than O_2, it is possible to avoid the geminate recombination complicating the reaction dynamics in the picosecond regime. Second, O_2 in solution autoxidizes the heme iron. Thus a fraction of the sample becomes inactive for ligand binding during measurement, which makes acquiring data over a long period and getting a good SNR with a limited amount of the protein sample difficult. Third, the quantum yield of O_2 photolysis in heme proteins is low [24]. Forth, CO usually is assumed to be a good

analog for O_2 in kinetic studies of HbA. However, the validity of this assumption has not been examined. Studies of protein dynamics related to the O_2 dissociation are essential for understanding the allostery of HbA. The high stability and sensitivity of our time-resolved RR spectrometers prompted time-resolved measurements of protein dynamics of HbA upon the O_2 dissociation. Fig. 16.6 compares the time-resolved RR spectra and the temporal changes of RR bands upon the CO- and O_2-dissociation. The time-resolved RR spectra of the CO- and O_2-photoproducts at 1-ns delay differed from those of the deoxy form at frequencies of the ν(Fe−His) and γ_7 bands, and at the intensity of the ν_8 band [7]. Spectral changes of the O_2 photoproduct in the submicrosecond region were faster than those of the CO photoproduct, providing evidence that the structural dynamics of the submicrosecond metastable state is ligand-dependent for HbA and that CO is not a perfect analog for O_2 in HbA. In contrast, no ligand dependence of the dynamics was observed for Mb. As discussed in section 16.3, the heme structure in Mb changes to one closely resembling the deoxy form within a few picoseconds [21]. These experimental data indicate that the structural dynamics of the heme occurring later than a nanosecond are specific to HbA, are ligand-dependent, and are associated with allostery of HbA.

The spectral changes of the γ_7 and ν_8 bands in nanoseconds observed in our study are likely associated with EF helical motion in the transition from B to R_{deoxy} and this motion is ligand-dependent in HbA. It is known that the interactions of the heme-bound ligand and the E helix are remarkably different between the O_2 and CO-bound forms of HbA. In the oxy form, the bound O_2 is hydrogen-bonded to the distal histidine in the E helix. In contrast, no such interaction in the CO-bound form exists between the bound CO and the E helix. The difference in the interaction causes ligand-dependence of the EF helical motion.

Both CO and O_2 undergo geminate recombination at room temperature; a fraction of the docked ligand diffuses into the heme pocket and rebinds. The time constants of geminate rebinding for O_2 and CO reported for HbA are 1.5 ns [64] and 70 ns [65], respectively. Interestingly, the rates of the rebinding and those of the $B-R_{deoxy}$ transition are of the same order of magnitude. This similarity can be explained as follows. If intermediates B and R_{deoxy} have slow and fast intrinsic rebinding rates, respectively, the apparent geminate rebinding rate is determined by the rate of the transition from B to R_{deoxy}. This idea is reasonable because the displacement of the E helix toward the heme plane on the $B-R_{deoxy}$ transition reduces the volume accessible for the ligand to diffuse and hence accelerates geminate rebinding in R_{deoxy}. Therefore the structural changes in the $B-R_{deoxy}$ transition are presumably associated with the ligand rebinding process.

Based on these observations, the following hypothesis explains the functional role of the observed structural dynamics. Since the R−T transition occurs in doubly or singly ligated states, HbA needs to release at least two O_2 molecules to change the quaternary structure. In red blood cells, ligands dissociate from HbA through thermal reactions, that is,

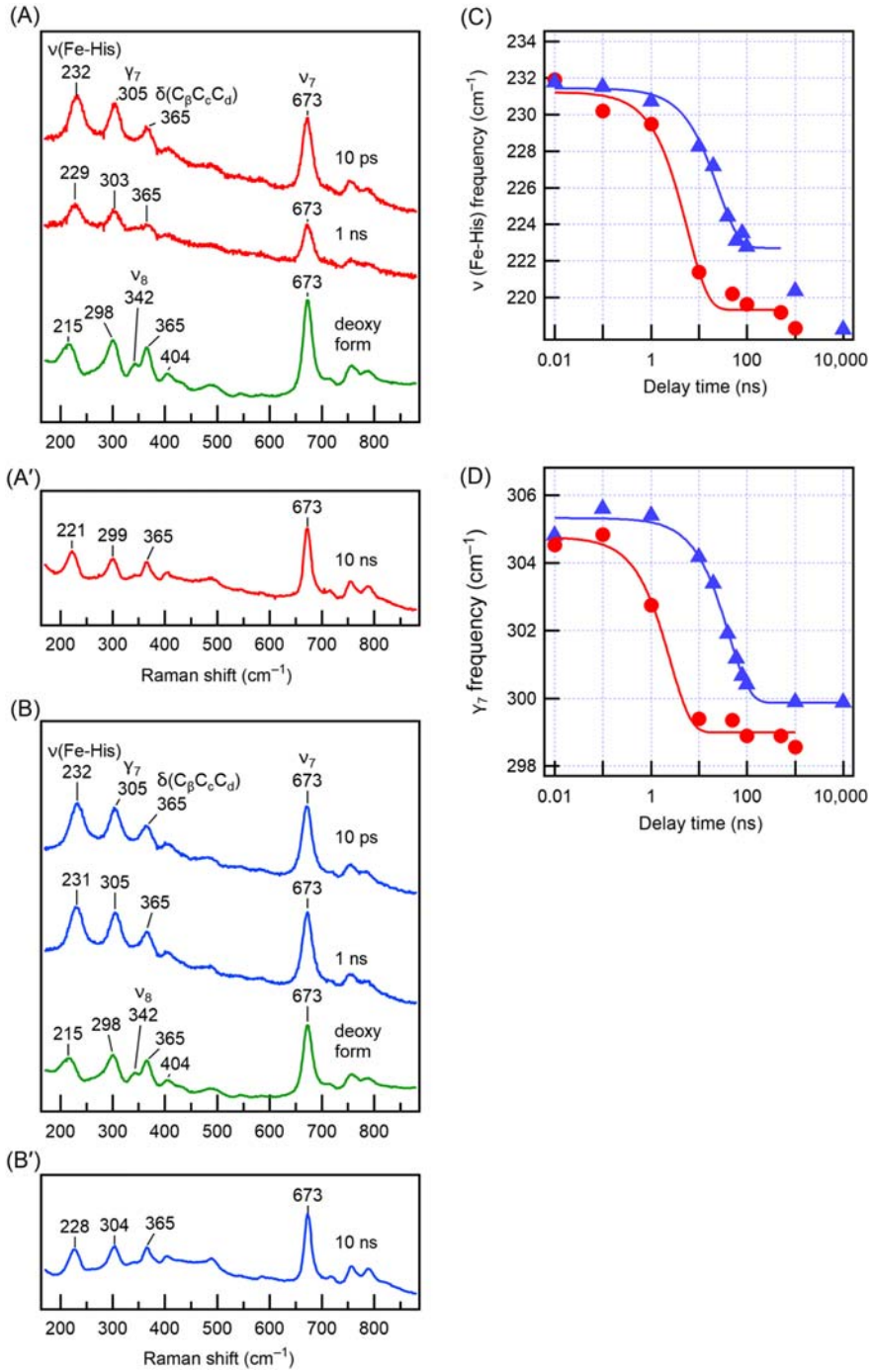


Figure 16.6

Picosecond time-resolved resonance Raman (RR) spectra of hemoglobin A (HbA) upon O_2 dissociation [(A) and (A′)] and CO dissociation [(B) and (B′)]. Temporal changes in (A) ν(Fe−His) and (B) γ_7 frequencies of HbA following ligand photolysis. Circles and triangles

(*Continued*)

stochastically. After the first ligand dissociates, the second ligand needs to dissociate before the dissociated ligand rebinds to undergo the R−T transition. The presence of the slow-binding intermediate allows the second ligand to dissociate before the dissociated ligand rebinds. Therefore the submicrosecond phase specific to HbA may be important for ensuring that HbA has a high R−T transition efficiency.

16.7 Summary and perspectives

For the last decade, we studied the structural evolution of HbA in the time range from picoseconds to submilliseconds by using time-resolved RR spectroscopy. This chapter summarized the insights on heme dynamics characteristic to HbA allostery obtained in the studies.

Distinct differences were observed in the structures of the photodissociated forms between HbA and Mb. Features of the ν(Fe−His), ν_8, and γ_7 bands for the photodissociated form at 1 ns and the deoxy state were substantially different for HbA, while they closely resembled each other for Mb. This indicates the presence of the metastable state in HbA dynamics. Interestingly, the isolated subunits exhibited similar temporal behaviors of their RR bands to those of HbA, but were completely different from those of Mb, meaning that the presence of the metastable state is due to the inherent characters of the HbA subunits.

The rate of structural evolution from the primary metastable state to the subsequent intermediate for HbA was different between the O_2 and CO dissociation, indicating that the structural dynamics of the submicrosecond metastable state specific to HbA is ligand-dependent. The difference in the rate of the structural changes was attributed to the structural difference in the EF helices between the O_2-bound and CO-bound forms.

Our study showed that the structural dynamics of the heme in the submicrosecond region are also dependent on the initial quaternary structure as described in Section 16.5 [9]. If the subunit interactions force the EF helices to shift, the quaternary structure-dependent

◀ represent data for O_2 and CO dissociation, respectively. The temporal shift of the ν(Fe−His) band was fitted using a single exponential function, yielding time constants of 5.7 ± 1.3 and 25.4 ± 3.5 ns, for O_2 and CO dissociation, respectively. The temporal shift of the γ_7 band was fitted using a single exponential function, yielding time constants of 2.5 ± 0.6 and 43.5 ± 3.8 ns, for O_2 and CO dissociation, respectively. For ν(Fe−His) frequency, data up to 100 ns were fitted because a change in a microsecond phase overlapped the region later than 1 μs. Source: *Reproduced from Y. Murakawa, M. Nagai, Y. Mizutani, Differences between protein dynamics of hemoglobin upon dissociation of oxygen and carbon monoxide, J. Am. Chem. Soc. 134 (2012) 1434−1437, with permission from American Chemical Society.*

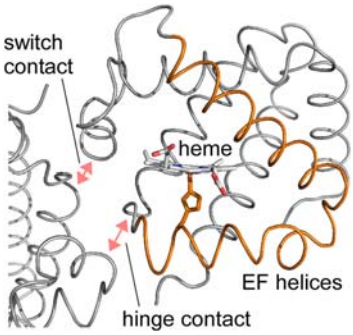


Figure 16.7
A diagram of the hemoglobin A (HbA) subunit (α subunit) and the $\alpha_1\beta_2$ interface. Purple broken lines refer to the A···E and F···H interhelical hydrogen bonds. Pink arrows represent the intersubunit contacts. The heme group and the subunit interface are bidirectionally communicated with each other via the enhancement factor (EF) helices, which is essential to the HbA allostery.

difference in the rates can be explained in a similar manner to the ligand-dependent difference in the rates. Accordingly, the structural changes of the EF helices in the submicrosecond region are inherent to the structure of the HbA subunits and that their rates are affected by the heme-bound ligand and intersubunit interactions. These lead to the picture that the heme group and the subunit interface are bidirectionally communicating with each other via the EF helices, which is essential to the HbA allostery (Fig. 16.7).

HbA allostery is based on its architecture. The propagations of atomic displacements would stop upon encountering internal void spaces in the protein structure. Therefore the absence of internal void spaces is crucial for successive structural changes to propagate to spatially distinct sites, that is, proteins are molecules of functional compactness [10]. Recently, it was proposed that quasi-two-dimensional strained surfaces in regulatory and ligand binding proteins including HbA mediate mechanical couplings between them [66]. HbA is a molecule endowed with functional compactness.

Acknowledgments

We express our gratitude to Prof. Misao Mizuno, Prof. Haruto Ishikawa, Dr. Shanyan Chang, Mr. Kenta Yamada, Ms. Yuka Murakawa, Prof. Masako Nagai, and Prof. Naoya Shibayama for great contributions to the work presented in this chapter. The work was supported by Grants-in-Aid for Scientific Research (Nos. JP15350013, JP17350009, JP19056013, JP20350007, JP23350007, and JP25104006) from the Japan Society for the Promotion of Science and the Ministry of Education, Culture, Sports, Science and Technology.

References

[1] J. Monod, J. Wyman, J.-P. Changeux, On the nature of allosteric transitions: a plausible model, J. Mol. Biol. 12 (1965) 88–118.

[2] G. Fermi, M.F. Perutz, B. Shaanan, R. Fourme, The crystal structure of human deoxyhaemoglobin at 1.74 Å resolution, J. Mol. Biol. 175 (1984) 159–174.

[3] M.F. Perutz, Regulation of oxygen affinity of hemoglobin: influence of structure of the globin on the heme iron, Annu. Rev. Biochem. 48 (1979) 327–386.

[4] M.F. Perutz, G. Fermi, B. Luisi, B. Shaanan, R.C. Liddington, Stereochemistry of cooperative mechanisms in hemoglobin, Acc. Chem. Res. 20 (1987) 309–321.

[5] M.F. Perutz, Stereochemistry of cooperative effects in haemoglobin: haem–haem interaction and the problem of allostery, Nature 228 (1970) 726–734.

[6] Y. Mizutani, M. Nagai, Ultrafast protein dynamics of hemoglobin as studied by picosecond time-resolved resonance Raman spectroscopy, Chem. Phys. 396 (2012) 45–52.

[7] Y. Murakawa, M. Nagai, Y. Mizutani, Differences between protein dynamics of hemoglobin upon dissociation of oxygen and carbon monoxide, J. Am. Chem. Soc. 134 (2012) 1434–1437.

[8] K. Yamada, H. Ishikawa, M. Mizuno, N. Shibayama, Y. Mizutani, Intersubunit communication via changes in hemoglobin quaternary structures revealed by time-resolved resonance Raman spectroscopy: direct observation of the Perutz mechanism, J. Phys. Chem. B 117 (2013) 12461–12468.

[9] S. Chang, M. Mizuno, H. Ishikawa, Y. Mizutani, Tertiary dynamics of human adult hemoglobin fixed in R and T quaternary structures, Phys. Chem. Chem. Phys. 20 (2018) 3363–3372.

[10] Y. Mizutani, Time-resolved resonance Raman spectroscopy and application to studies on ultrafast protein dynamics, Bull. Chem. Soc. Jpn. 90 (2017) 1344–1371.

[11] J. Tang, A.C. Albrecht, Developments in the theories of vibrational Raman intensities, in: H.A. Szymanski (Ed.), Raman Spectroscopy: Theory and Practice, Springer US, Boston, MA, 1970, pp. 33–68.

[12] T. Kitagawa, The heme protein structure and the iron histidine stretching mode, in: T.G. Spiro (Ed.), Biological Application of Raman Spectroscopy, Vol. III, John Wiley and Sons, New York, 1988, pp. 97–131.

[13] D.L. Rousseau, J.M. Friedman, Transient and cryogenic studies of photodissociated hemoglobin and myoglobin, in: T.G. Spiro (Ed.), Biological Application of Raman Spectroscopy, Vol. III, John Wiley and Sons, New York, 1988, pp. 133–216.

[14] T.G. Spiro, X.-Y. Li, Resonance Raman spectroscopy of metalloporphyrins, in: T.G. Spiro (Ed.), Biological Application of Raman Spectroscopy, Vol. III, John Wiley and Sons, New York, 1988, pp. 1–37.

[15] Y. Uesugi, Y. Mizutani, T. Kitagawa, Developments of widely tunable light sources for picosecond time-resolved resonance Raman spectroscopy, Rev. Sci. Instrum. 68 (1997) 4001–4008.

[16] Y. Uesugi, Y. Mizutani, S.G. Kruglik, A.G. Shvedko, V.A. Orlovich, T. Kitagawa, Characterization of stimulated Raman scattering of hydrogen and methane gases as a light source for picosecond time-resolved Raman spectroscopy, J. Raman Spectrosc. 31 (2000) 339–348.

[17] M. Mizuno, Y. Mizutani, Protein response to chromophore isomerization in microbial rhodopsins revealed by picosecond time-resolved ultraviolet resonance Raman spectroscopy: a review, Recent Progress in Colloid and Surface Chemistry with Biological Applications, Vol. 1215, American Chemical Society, 2015, pp. 329–353.

[18] Y. Mizutani, T. Kitagawa, Ultrafast dynamics of myoglobin probed by time-resolved resonance Raman spectroscopy, Chem. Rec. 1 (2001) 258–275.

[19] K. Yamada, H. Ishikawa, Y. Mizutani, Protein dynamics of isolated chains of recombinant human hemoglobin elucidated by time-resolved resonance Raman spectroscopy, J. Phys. Chem. B 116 (2012) 1992–1998.

[20] M. Mizuno, N. Hamada, F. Tokunaga, Y. Mizutani, Picosecond protein response to the chromophore isomerization of photoactive yellow protein: selective observation of tyrosine and tryptophan residues by time-resolved ultraviolet resonance Raman spectroscopy, J. Phys. Chem. B 111 (2007) 6293–6296.

[21] Y. Mizutani, T. Kitagawa, Ultrafast structural relaxation of myoglobin following photodissociation of carbon monoxide probed by time-resolved resonance Raman spectroscopy, J. Phys. Chem. B 105 (2001) 10992–10999.
[22] B.I. Greene, R.M. Hochstrasser, R.B. Weisman, W.A. Eaton, Spectroscopic studies of oxy- and carbonmonoxyhemoglobin after pulsed optical excitation, Proc. Natl Acad. Sci. U.S.A. 75 (1978) 5255–5259.
[23] J.W. Petrich, C. Poyart, J.L. Martin, Photophysics and reactivity of heme proteins: a femtosecond absorption study of hemoglobin, myoglobin, and protoheme, Biochemistry 27 (1988) 4049–4060.
[24] X. Ye, A. Demidov, P.M. Champion, Measurements of the photodissociation quantum yields of Mbno and Mbo2 and the vibrational relaxation of the six-coordinate heme species, J. Am. Chem. Soc. 124 (2002) 5914–5924.
[25] Q.H. Gibson, J.S. Olson, R.E. McKinnie, R.J. Rohlfs, A kinetic description of ligand binding to sperm whale myoglobin, J. Biol. Chem. 261 (1986) 10228–10239.
[26] M. Abe, T. Kitagawa, Y. Kyogoku, Resonance Raman spectra of octaethylporphyrinato-Ni(II) and meso-deuterated and 15n substituted derivatives. II. A normal coordinate analysis, J. Chem. Phys. 69 (1978) 4526–4534.
[27] T. Kitagawa, K. Nagai, M. Tsubaki, Assignment of the Fe—Nϵ (His F8) stretching band in the resonance Raman spectra of deoxy myoglobin, FEBS Lett. 104 (1979) 376–378.
[28] S. Hu, K.M. Smith, T.G. Spiro, Assignment of protoheme resonance Raman spectrum by heme labeling in myoglobin, J. Am. Chem. Soc. 118 (1996) 12638–12646.
[29] J.M. Friedman, R.A. Stepnoski, M. Stavola, M.R. Ondrias, R.L. Cone, Ligation and quaternary structure induced changes in the heme pocket of hemoglobin: a transient resonance Raman study, Biochemistry 21 (1982) 2022–2028.
[30] S. Dasgupta, T.G. Spiro, Resonance Raman characterization of the 7-Ns photoproduct of (carbonmonoxy) hemoglobin: implications for hemoglobin dynamics, Biochemistry 25 (1986) 5941–5948.
[31] B.R. Gelin, A.W.-M. Lee, M. Karplus, Hemoglobin tertiary structural change on ligand binding its role in the cooperative mechanism, J. Mol. Biol. 171 (1983) 489–559.
[32] K. Nagai, T. Kitagawa, Differences in Fe(II)-N epsilon(His-F8) stretching frequencies between deoxyhemoglobins in the two alternative quaternary structures, Proc. Natl Acad. Sci. U.S.A. 77 (1980) 2033–2037.
[33] S. Chang, M. Mizuno, H. Ishikawa, Y. Mizutani, Effect of the N-terminal residues on the quaternary dynamics of human adult hemoglobin, Chem. Phys. 469-470 (2016) 31–37.
[34] G. Balakrishnan, M.A. Case, A. Pevsner, X. Zhao, C. Tengroth, G.L. McLendon, et al., Time-resolved absorption and UV resonance Raman spectra reveal stepwise formation of T quaternary contacts in the allosteric pathway of hemoglobin, J. Mol. Biol. 340 (2004) 843–856.
[35] M. Cammarata, M. Levantino, F. Schotte, P.A. Anfinrud, F. Ewald, J. Choi, et al., Tracking the structural dynamics of proteins in solution using time-resolved wide-angle X-ray scattering, Nat. Methods 5 (2008) 881–886.
[36] M. Cammarata, M. Levantino, M. Wulff, A. Cupane, Unveiling the timescale of the R–T transition in human hemoglobin, J. Mol. Biol. 400 (2010) 951–962.
[37] J. Hofrichter, J.H. Sommer, E.R. Henry, W.A. Eaton, Nanosecond absorption spectroscopy of hemoglobin: elementary processes in kinetic cooperativity, Proc. Natl Acad. Sci. U.S.A. 80 (1983) 2235–2239.
[38] V. Jayaraman, K. Rodgers, I. Mukerji, T. Spiro, Hemoglobin allostery: resonance Raman spectroscopy of kinetic intermediates, Science 269 (1995) 1843–1848.
[39] G. Balakrishnan, C.-H. Tsai, Q. Wu, M.A. Case, A. Pevsner, G.L. McLendon, et al., Hemoglobin site-mutants reveal dynamical role of interhelical H-bonds in the allosteric pathway: time-resolved UV resonance Raman evidence for intra-dimer coupling, J. Mol. Biol. 340 (2004) 857–868.
[40] J.A. Shelnutt, K. Alston, J.Y. Ho, N.T. Yu, T. Yamamoto, J.M. Rifkind, Four- and five-coordinate species in nickel-reconstituted hemoglobin and myoglobin: Raman identification of the nickel-histidine stretching mode, Biochemistry 25 (1986) 620–627.

[41] C.S. Reed, R. Hampson, S. Gordon, R.T. Jones, M.J. Novy, B. Brimhall, et al., Erythrocytosis secondary to increased oxygen affinity of a mutant hemoglobin, hemoglobin Kempsey, Blood 31 (1968) 623–632.
[42] M.F. Perutz, Nature of haem-haem interaction, Nature 237 (1972) 495–499.
[43] S. Matsukawa, K. Mawatari, Y. Yoneyama, T. Kitagawa, Correlation between the iron-histidine stretching frequencies and oxygen affinity of hemoglobins. a continuous strain Model, J. Am. Chem. Soc. 107 (1985) 1108–1113.
[44] J. Bonaventura, A. Riggs, Hemoglobin Kansas, a human hemoglobin with a neutral amino acid substitution and an abnormal oxygen equilibrium, J. Biol. Chem. 243 (1968) 980–991.
[45] Q.H. Gibson, A. Riggs, T. Imamura, Kinetic and equilibrium properties of hemoglobin Kansas, J. Biol. Chem. 248 (1973) 5976–5986.
[46] K.R. Reissmann, W.E. Ruth, T. Nomura, A human hemoglobin with lowered oxygen affinity and impaired heme-heme interactions, J. Clin. Investig. 40 (1961) 1826–1833.
[47] L. Anderson, Structures of deoxy and carbonmonoxy haemoglobin Kansas in the deoxy quaternary conformation, J. Mol. Biol. 94 (1975) 33–49.
[48] J.V. Kilmartin, N.L. Anderson, S. Ogawa, Response of the Bohr group salt bridges to ligation of the T state of haemoglobin Kansas, J. Mol. Biol. 123 (1978) 71–87.
[49] S. Ogawa, A. Mayer, R.G. Shulman, High resolution proton magnetic resonance study of the two quaternary states in fully ligated hemoglobin Kansas, Biochem. Biophys. Res. Commun. 49 (1972) 1485–1491.
[50] J.M. Salhany, S. Ogawa, R.G. Shulman, Correlation between quaternary structure and ligand dissociation kinetics for fully liganded hemoglobin, Biochemistry 14 (1975) 2180–2190.
[51] L.W.M. Fung, C. Ho, Proton nuclear magnetic resonance study of the quaternary structure of human hemoglobins in water, Biochemistry 14 (1975) 2526–2535.
[52] H.W. Kim, T.J. Shen, D.P. Sun, N.T. Ho, M. Madrid, M.F. Tam, et al., Restoring allosterism with compensatory mutations in hemoglobin, Proc. Natl Acad. Sci. U.S.A. 91 (1994) 11547–11551.
[53] H.F. Bunn, R.C. Wohl, T.B. Bradley, M. Cooley, Q.H. Gibson, Functional properties of hemoglobin Kempsey, J. Biol. Chem. 249 (1974) 7402–7409.
[54] T.R. Lindstrom, J.J. Baldassare, H.F. Bunn, C. Ho, Nuclear magnetic resonance and spin-label studies of hemoglobin Kempsey, Biochemistry 12 (1973) 4212–4217.
[55] V. Jayaraman, T.G. Spiro, Structural evolution of the heme group during the allosteric transition in hemoglobin: insights from resonance Raman spectra of isotopically labeled heme, Biospectroscopy 2 (1996) 311–316.
[56] M.C. Schneebeck, L.E. Vigil, J.M. Friedman, M.D. Chavez, M.R. Ondrias, Heme-carbon monoxide religation in photolyzed hemoglobin: a time-resolved Raman study of the iron-carbon monoxide stretching mode, Biochemistry 32 (1993) 1318–1323.
[57] J. Kneipp, G. Balakrishnan, R. Chen, T.-J. Shen, S.C. Sahu, N.T. Ho, et al., Dynamics of allostery in hemoglobin: roles of the penultimate tyrosine H bonds, J. Mol. Biol. 356 (2006) 335–353.
[58] Q.H. Gibson, The photochemical formation of a quickly reacting form of haemoglobin, Biochem. J. 71 (1959) 293–303.
[59] J. Friedman, Structure, dynamics, and reactivity in hemoglobin, Science 228 (1985) 1273–1280.
[60] Y. Mizutani, T. Kitagawa, Direct observation of cooling of heme upon photodissociation of carbonmonoxy myoglobin, Science 278 (1997) 443–446.
[61] C. Shank, E. Ippen, R. Bersohn, Time-resolved spectroscopy of hemoglobin and its complexes with subpicosecond optical pulses, Science 193 (1976) 50–51.
[62] T.W. Scott, J.M. Friedman, Tertiary-structure relaxation in hemoglobin: a transient Raman study, J. Am. Chem. Soc. 106 (1984) 5677–5687.
[63] J. Terner, D.F. Voss, C. Paddock, R.B. Miles, T.G. Spiro, Picosecond resonance Raman spectrum of the oxyhemoglobin photoproduct. evidence for an electronically excited state, J. Phys. Chem. 86 (1982) 859–861.

[64] S.V. Lepeshkevich, J. Karpiuk, I.V. Sazanovich, B.M. Dzhagarov, A kinetic description of dioxygen motion within A- and B-subunits of human hemoglobin in the R-state: geminate and bimolecular stages of the oxygenation reaction, Biochemistry 43 (2004) 1675–1684.

[65] R. Chen, T.G. Spiro, Monitoring the allosteric transition and CO rebinding in hemoglobin with time-resolved FTIR spectroscopy, J. Phys. Chem. A 106 (2002) 3413–3419.

[66] M.R. Mitchell, T. Tlusty, S. Leibler, Strain analysis of protein structures and low dimensionality of mechanical allosteric couplings, Proc. Natl Acad. Sci. U.S.A. 113 (2016) E5847–E5855.

This page intentionally left blank

CHAPTER 17

Immuno-SERS: from nanotag design to assays and microscopy

Yuling Wang[1], Yuying Zhang[2] and Sebastian Schlücker[3]

[1]Department of Molecular Sciences, ARC Centre of Excellence for Nanoscale BioPhotonics, Macquarie University, Sydney, NSW, Australia, [2]Medical School of Nankai University, Tianjin, China, [3]Department of Chemistry and Center for Nanointegration Duisburg-Essen (CENIDE), University of Duisburg-Essen, Essen, Germany

17.1 Introduction

Immuno-SERS (iSERS) is a newly emerging technique that combines target recognition through conventional antigen−antibody interaction and signal output through surface-enhanced Raman scattering (SERS) detection. Analogous to protein detection assays such as enzyme-linked immunosorbent assay (ELISA) or imaging techniques such as immunocytochemistry/immunohistochemistry/immunofluorescence (ICC/IHC/IF), in which antibodies are labeled with either external chromogens or fluorophores, in iSERS, target recognition molecules are labeled with SERS nanotags. A target-specific SERS nanotag is typically composed of a colloidal metal nanoparticle for signal enhancement, Raman reporter molecules adsorbed onto its surface for identification, and biomolecules attached to the SERS nanotag for target-specific recognition. Although there are few review articles on SERS tags for bioanalysis [1−4] and cancer detection [5,6], our focus in this chapter is on the design of SERS nanotags for both iSERS assays and iSERS microscopic imaging.

17.2 Surface-enhanced Raman scattering nanotags

SERS nanotags (Fig. 17.1) contain a metal nanoparticle as the core for signal amplification via localized surface plasmon resonance (LSPR) excitation, Raman reporter molecules adsorbed onto the metal surface for identification, a desired protective layer or shell for particle stabilization, a spacer (linker), and a ligand (antibody, peptide, etc.) for molecular recognition of the target molecule [7−9].

Vibrational Spectroscopy in Protein Research.
DOI: https://doi.org/10.1016/B978-0-12-818610-7.00017-7

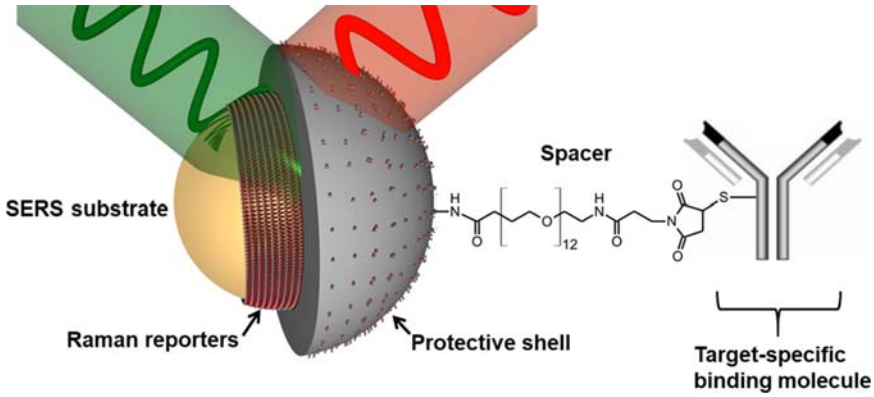


Figure 17.1
Components of surface-enhanced Raman scattering (SERS) nanotags: SERS substrate, Raman reporters, protective shell (optional), spacer (linker), and the target-specific binding molecule. Source: *Cited from Y. Wang, S. Schlücker, Rational design and synthesis of SERS labels, Analyst 138 (8) (2013): 2224–2238..*

17.2.1 Surface-enhanced Raman scattering nanotags: plasmonic nanostructures

The first step in the design of SERS nanotags is the selection of the plasmonic nanostructures with the desired optical properties, in particular the position of the LSPR peak of the nanoparticles. In this section, we will discuss the optical properties of plasmonic nanostructures in terms of SERS activity. Plenty of methods have been reported for the synthesis of plasmonic nanostructures including wet-chemistry (seed-mediated growth), template nanoscale lithography, thin-film, and the template approach, which have been summarized in recent reviews [10–12]. In this section, we will mainly focus on the synthesis and characterization of plasmonic nanostructures that are widely used for iSERS assays. Based on the shape or geometry of the nanostructures, we will give a brief overview on quasispherical metallic nanostructures, anisotropic nanoparticles, as well as plasmonic nanoassemblies, highlighting the plasmonic effect for SERS activity, and the potential to be used as the core of SERS nanotags.

17.2.1.1 Quasispherical metallic nanoparticles

Gold and silver nanoparticles show the most significant SERS effect due to the localized surface plasmon resonance generated on their surfaces, and are most widely used for iSERS assays and imaging. Numerous protocols are available for the synthesis of quasispherical noble metal nanoparticles [13–16]. The simplest and most common approach is still the reduction of metal salts with a variety of reducing and capping agents. To improve the stability and the enhancement of plasmonic nanoparticles, combinations or alloys of metals are one option. For instance, gold/silver nanoshells were designed and synthesized as

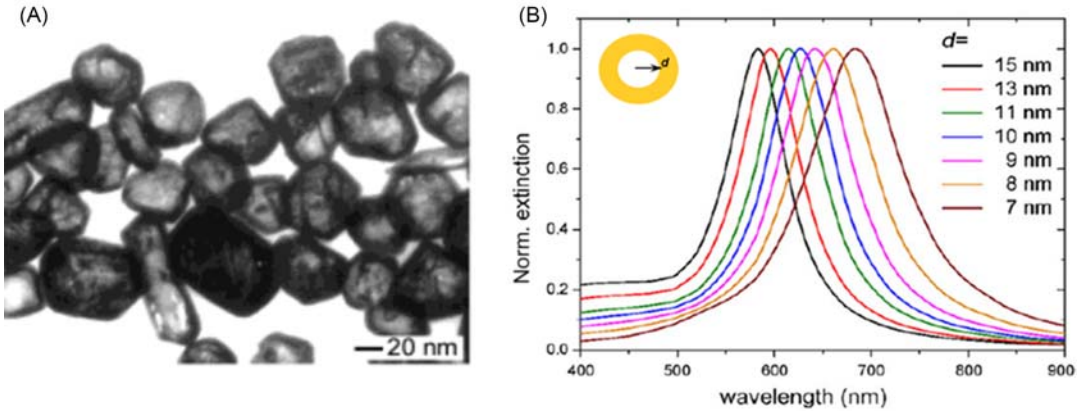


Figure 17.2

Transmission electron microscopy (TEM) image of gold/silver nanoshells (A) and the tunable localized surface plasmon resonance (LSPR) of gold/silver nanoshells with tunable shell thickness (B). Source: *Cited from Y. Sun, B.T. Mayers, Y. Xia, Template-engaged replacement reaction: a one-step approach to the large-scale synthesis of metal nanostructures with hollow interiors, Nano Lett. 2 (5) (2002) 481–485 and M. Gellner, B. Küstner, S. Schlücker, Optical properties and SERS efficiency of tunable gold/silver nanoshells, Vib. Spectrosc. 50 (1) (2009) 43–47.*

composite metallic nanostructures (Fig. 17.2A) that show a higher SERS effect compared to a single metal [17,18]. More importantly, by tuning the shell thickness, the LSPR of gold/silver nanoshells can be tuned (Fig. 17.2B). With regards to the enhancement factor (EF) of quasispherical metallic nanoparticles for SERS, it has been found that individual quasispherical nanoparticles generate very low EF for SERS [8,19]. Single molecule detection [20,21] was reported in case of the strong surface plasmon coupling effect (called "hot spot") between silver nanoparticles (AgNPs). In order to get a high order of EF, well-defined plasmonic nanostructures with one or multiple "hot spots" are required, which will be discussed in the following section.

17.2.1.2 Anisotropic nanoparticles

As it is hard to obtain a detectable SERS signal from individual metallic nanospheres, anisotropic nanoparticles such as nanorods, nanostars, nanocubes, nanoprisms, and nanoplates became very important SERS substrates for iSERS (Fig. 17.3A), as these exhibit significantly higher electromagnetic field enhancements at sharp edges ("lighting rod effect" or "plasmonic antenna effect"), making them more attractive for use as SERS substrates [22–25] for iSERS. For the synthesis of anisotropic nanoparticles, the most widely employed approach is seed-mediated growth, which involves seed formation and growth. A typical example is the synthesis of gold nanorods (Fig. 17.3A), which starts with quasispherical ~4-nm gold seeds and subsequent reduction of more metal salt with a weak reducing agent, such as ascorbic acid, in the presence of structure-directing additives, such

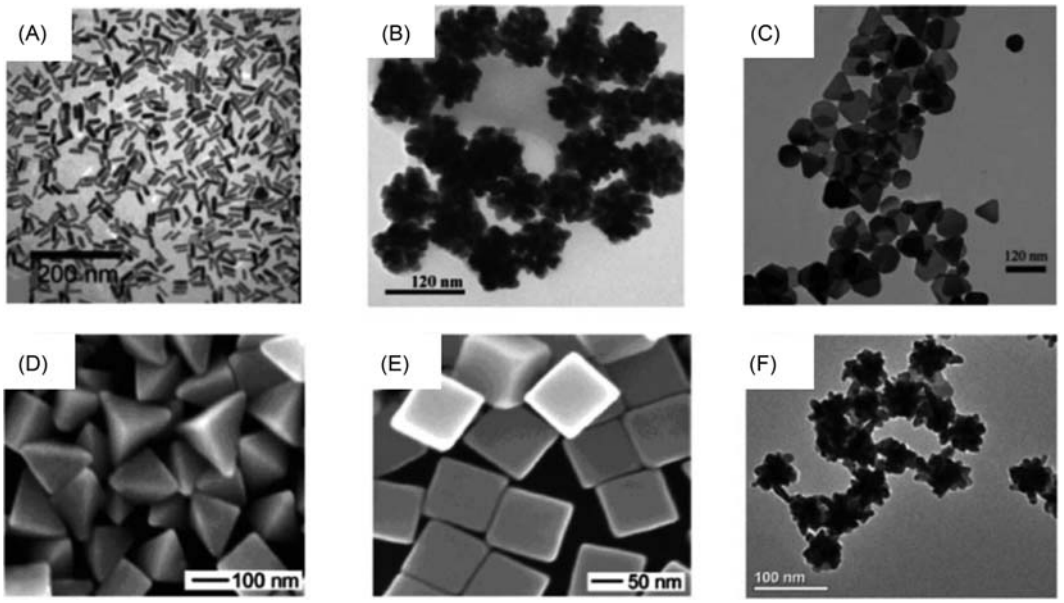


Figure 17.3
Typical electronmicrographs of anisotropic nanostructures for efficient surface-enhanced Raman scattering (SERS). Nanorods (A), nanoflowers (B), nanoplates (C), nanoprisms (D), nanocubes (E), and nanostars (F). Source: *Cited from Y. Wang, S. Schlücker, Rational design and synthesis of SERS labels, Analyst 138 (8) (2013): 2224–2238.*

as cetyltrimethylammonium bromide (CTAB), leading to the formation of gold nanorods [11,22]. The aspect ratio can be controlled by the relative concentrations of the reagents. By tuning the aspect ratio of the gold nanorods, their LSPR can be tuned, which is an important aspect for the application of gold nanorods in SERS as well as in phototherapy.

17.2.1.3 Plasmonic nanoassemblies

To generate the necessary hot spots for high SERS enhancement, plasmonic clusters, including dimers, trimers, and high-order nanoaggregates, have been prepared. As it is very important to understand the relationship between nanoparticle structure and the SERS activity, Van Duyne and coworkers have reported the SERS EF of different nanostructures (mainly dimers and trimers) [26]. As indicated in Fig. 17.4, individual trimers have been encapsulated with a silica shell to avoid interference from the environment as well as to ensure that the contribution is solely from individual nanoparticles. It was demonstrated that the creation of hot spots, where two nanoparticles are in subnanometer proximity or have coalesced to form crevices, is paramount for achieving maximum SERS EF. Specifically, L-shaped trimer nanoantennae comprising three Au cores showed an EF of 10^8–10^{10} (Fig. 17.4A) and the dark-field Rayleigh scattering spectrum of the L-shaped

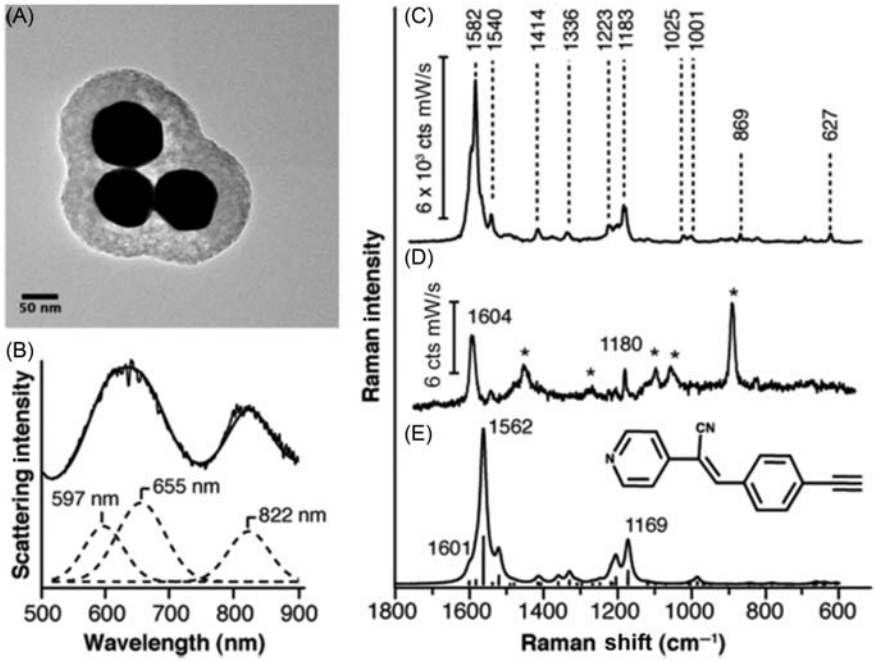


Figure 17.4

L-shaped trimer nanoantenna (A); localized surface plasmon resonance (LSPR) spectrum of the trimer obtained by dark-field Raleigh scattering microscopy (B); SERS spectrum of the Raman molecule (PCEPE) (C); normal Raman spectrum of PCEPE (D); and Raman spectrum of PECEP calculated using density functional theory (DFT) (E). Source: *Cited from K.L. Wustholz, et al., Structure–activity relationships in gold nanoparticle dimers and trimers for surface-enhanced Raman spectroscopy, J. Am. Chem. Soc. 132 (31) (2010) 10903–10910.*

trimer contains three peaks (Fig. 17.4B), corresponding to dipolar and multipolar LSPRs [26]. Single-particle SERS obtained from the trimer nanoantenna showed distinct peaks from the Raman molecule bis(2,3′-bipyridine)-(5-isothiocyanato-phenanthroline) rutheniumbis(hexafluorophosphate) (PCEPE), which correlated with results from density functional theory (DFT) calculations (Fig. 17.4C–E).

Additionally, Yoon et al. [27] reported a method for the preparation of ideal dimers of two highly spherical AuNPs separated by a constant gap for precision plasmonics by a substrate-based assembly strategy with efficient removal of CTAB, as indicated in Fig. 17.5. Both transmission electron microscopy (TEM) image and dark-field scattering spectra show the high structural and plasmonic uniformity of the ideal dimers (Fig. 17.5). Thus it is expected that highly uniform SERS signals from molecules in the hot spots of single ideal dimers will be generated since both gap distance and gap morphology are well defined and highly reproducible across dimers.

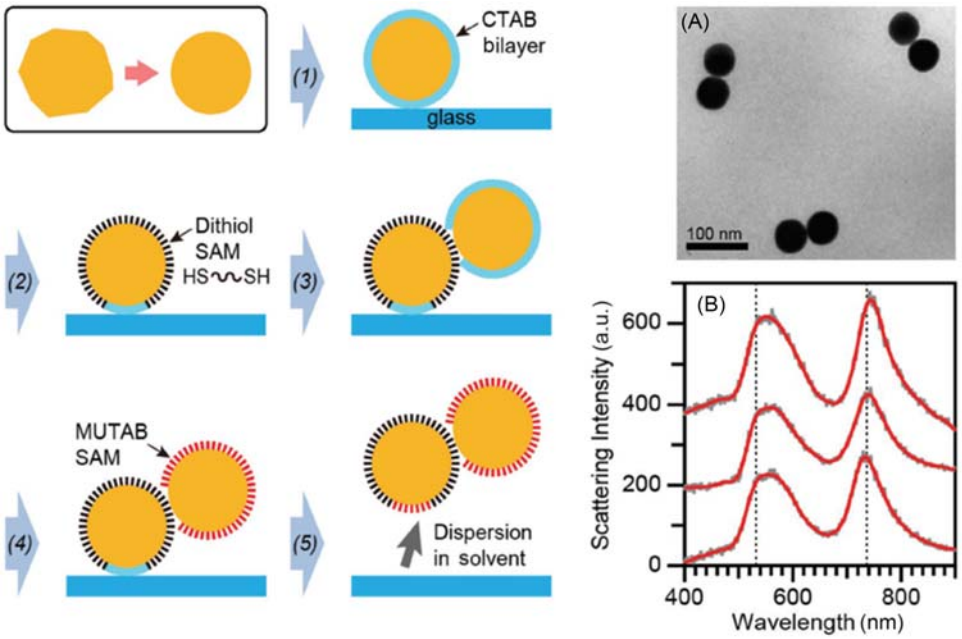


Figure 17.5

(*Left*) Schematic illustration of the substrate-based sequential dimer assembly for the synthesis of ideal dimers. (*Right*) transmission electron microscopy (TEM) image and dark-field scattering spectra of single ideal dimers. Source: *Cite from J.H. Yoon, et al., Ideal dimers of gold nanospheres for precision plasmonics: synthesis and characterization at the single-particle level for identification of higher order modes, Small 14 (4) (2018) 1702754.*

Furthermore, plasmonic nanostructures with multiple hot spots have been reported with satellite nanostructures by using either gold spheres or gold nanorods as the core [28–30]. To obtain the high orders of nanoassemblies, the linker to connect the core and satellite nanoparticles is key. Chemicals with dual/multifunctional groups are often used, such as cystamine and dithiols. As the silica surface is versatile, the functionalization of the silica surface, followed by the assembly of small nanoparticles is a very useful strategy. More importantly, the distance between the satellite nanoparticles and the core nanoparticle can be controlled by the thickness of the silica shell. As indicated in Fig. 17.6, gold nanoparticles (AuNPs) with a diameter of 80 nm were incubated with Raman reporters and encapsulated with a very thin silica shell (Fig. 17.6A). The ultrathin glass shell was then functionalized with a binary mixture of silanes including an aminosilane, which can electrostatically attract the negatively charged gold nanospheres to form the 3D structures (Fig. 17.6B). Correlative high-resolution scanning electron microscopy (HR-SEM)/dark-field/LSPR/SERS experiments on individual 3D SERS-active superstructures together with finite element method (FEM) calculations (Fig. 17.5C) confirmed the plasmonic coupling between the core and the satellite particles, with hot spots occurring between core and satellites as well as between satellites [28].

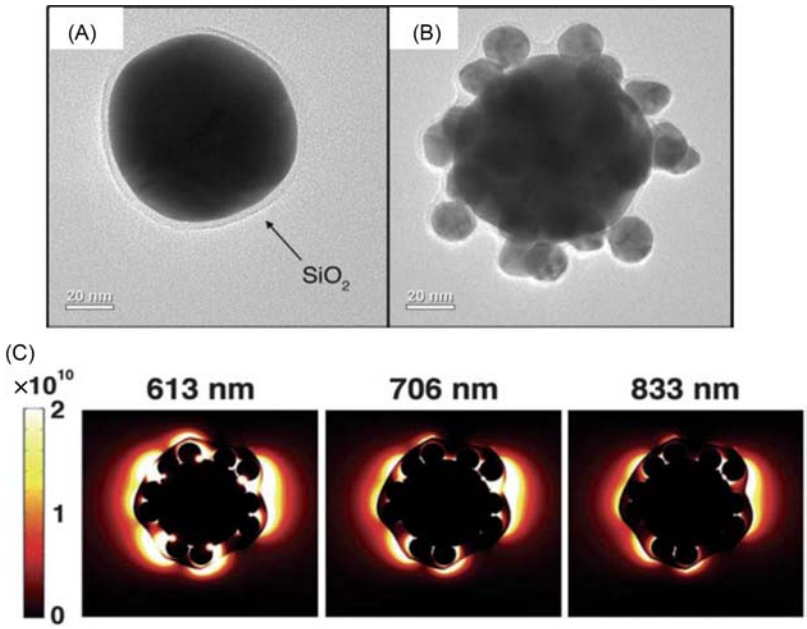


Figure 17.6

Plasmonic nanoassemblies (satellite nanostructures with gold/gold spheres). Transmission electron microscopy (TEM) images of a ca. 80 nm large gold core coated with an ultrathin silica shell (A); 3D plasmonic nanoassemblies with 20 nm gold spheres as the satellites assembled on the silica-coated gold core (B); and results from finite element method (FEM) simulations showing the distribution of the local electric field for three different localized surface plasmon resonances (LSPRs) at 613, 706, and 833 nm (C). Source: *Cited from M. Gellner, et al., 3D Self-assembled plasmonic superstructures of gold nanospheres: synthesis and characterization at the single-particle level, Small 7 (24) (2011) 3445–3451.*

17.2.1.4 Comparison of the surface-enhanced Raman scattering activity on single nanoparticle

As different classes of plasmonic nanoparticles show different SERS activities, recently the single-particle SERS activity of AuNPs, Au nanoshells, Au nanostars, and Au/Au supraparticles has been compared by correlating results from SEM, SERS, and LSPR [31]. Interestingly, it has been found that single gold nanostars and single gold core/gold satellite particles exhibit a detectable SERS signal under the given experimental conditions, while single AuNPs and single hollow gold/silver nanoshells are not detectable (Fig. 17.7). The SERS EF of 10^4 for single nanostars and 10^5 for a single supraparticle were determined under 633 nm laser excitation, thus this may define a further step toward a systematic design of complex NPs with tailored high SERS sensitivities.

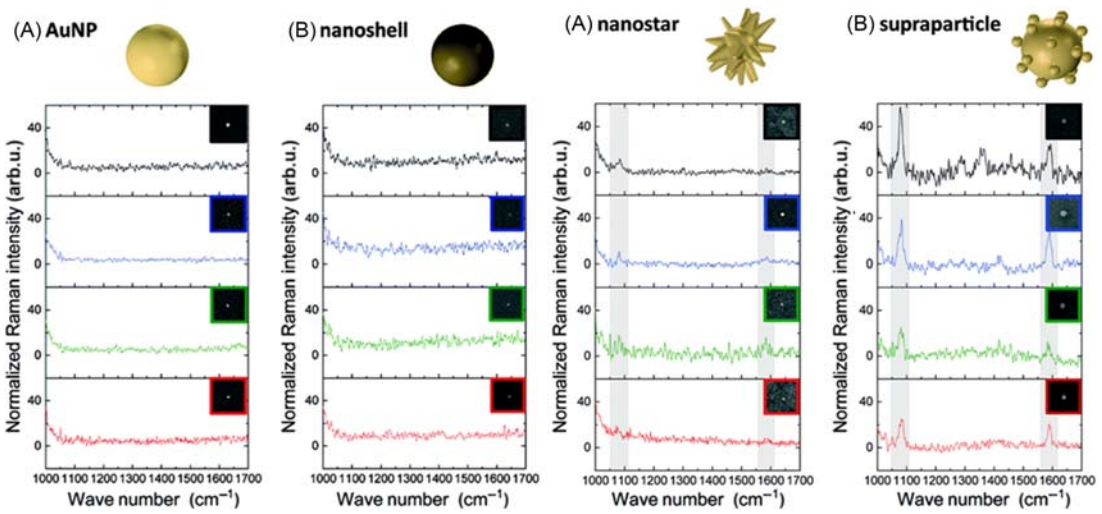


Figure 17.7
Single-particle surface-enhanced Raman scattering (SERS) experiments on Au nanospheres; Au/Ag nanoshells, nanostars, and supraparticles, with the single-particle SERS spectra and corresponding scanning electron microscopy (SEM) images of monomers as insets. Source: *Cited from V. Tran, et al., Probing the SERS brightness of individual Au nanoparticles, hollow Au/Ag nanoshells, Au nanostars and Au core/Au satellite particles: single-particle experiments and computer simulations, Nanoscale 10 (46) (2018) 21721–21731.*

17.2.2 Surface-enhanced Raman scattering nanotags: Raman reporter molecules

The second key step in the design and preparation of SERS nanotags (Fig. 17.1) is the selection of Raman reporter molecules with a characteristic spectral signature for adsorption to plasmonically metal nanoparticles. In principle, a large variety of molecules may be utilized as Raman reporters. Ideally, Raman reporter molecules exhibit (1) high differential Raman scattering cross sections for high signal levels; (2) a small number of atoms and/or high symmetry leading to a minimal number of Raman bands for dense multiplexing; (3) low or no photobleaching for signal stability; and (4) chemisorption to the colloidal metal surface for strong binding.

17.2.2.1 Fluorescent chromophores as Raman reporter molecules for SERRS and multiplex capacity

Fluorescent chromophores possess several aromatic groups, which can generally provide stronger SERS signals than nonfluorescent chromophores because of the additional enhancement via surface-enhanced resonance Raman scattering (SERRS). Generally, the term multiplexing refers to the parallel determination of several parameters within a single experiment. In the context of SERS as a labeling strategy for targeted research,

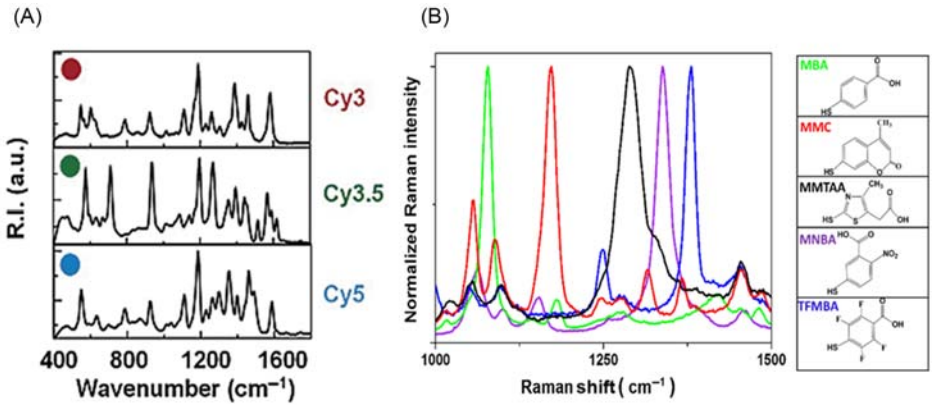


Figure 17.8
Surface-enhanced Raman scattering (SERS) spectra of fluorescent dyes (A) and of small aromatic thiols forming self-assembled monolayers (SAMs) on gold(B). Source: *Cited from Y.C. Cao, et al., Raman dye-labeled nanoparticle probes for proteins, J. Am. Chem. Soc. 125 (48) (2003) 14676–14677 and Y. Lai, et al., Raman-encoded microbeads for spectral multiplexing with SERS detection, RSC Adv. 5 (18) (2015) 13762–13767.*

multiplexing addresses the issue of simultaneously detecting and identifying the spectral fingerprint of several distinct SERS nanotags. Multiplex capacity was first demonstrated for fluorescent chromophores as the labels on nanoparticles for the detection of DNA and proteins in solution as reported by Mirkin and coworkers [32,33]. In their primary work, commercially available dyes (Cy3, Cy3.5, Cy5, rhodamine 6G, tetramethyl rhodamine, and Texas Red) were used as Raman reporter molecules and were performed by a three-component sandwich assay in a microarray format for multiplexed (6-plex) detection of DNA [32] and 3-plex detection of proteins (Fig. 17.8A) [33]. Graham and coworkers have also performed multiplexed experiments for DNA detection, using either custom-made benzotriazoles [35] or further commercially available dyes, including BODIPY, Cy5.5, fluorescein amidites, and carboxy-X-rhodamine [36–38]. Chang's group presented the combinatorial synthesis of a triphenylmethine library of compounds serving as NIR-sensitive Raman reporters for multiplexing [39–41].

17.2.2.2 Self-assembled monolayer of Raman reporters and multiplex capacity

Functional groups with a high affinity towards gold and silver, such as thiols, isothiocyanates, and amines, could adsorb on the metal nanoparticles' surface as the reporter. For example, Porter and coworkers have introduced aryl thiols/disulfides as Raman reporters because they form self-assembled monolayers (SAMs) on gold surfaces via stable Au–S bonds [42–44].

There are several advantages to using a SAM of Raman reporter molecules [9,42–45]. Firstly, a full monolayer ensures maximum SERS sensitivity since it guarantees the maximum surface coverage with Raman reporter molecules in comparison with submonolayer coverage. Secondly, clean and reproducible Raman signatures are obtained due to the uniform orientation of the molecules within the SAM relative to the surface normal of the nanoparticles [9,46,47]. Thirdly, spectral interferences with unwanted contributions, due to the coadsorption of molecules other than Raman reporters on the particle surface, are minimized. In contrast to the submonolayer coverage with Raman reporter molecules, a complete SAM leads to a higher sensitivity. Using hollow gold/silver nanospheres and red laser excitation (632.8 nm), it has been demonstrated experimentally that a complete SAM yields approximately 22 times more intense SERS signals compared with submonolayer coverage [9]. Multiplex detection could also be achieved using SAMs of Raman reporters based on the characteristic Raman peak because of the narrow bandwidth of the Raman signal without spectral overlap and the unique spectra of all Raman reporter molecules as illustrated in Fig. 17.8B. Raman spectra of five different Raman reporter molecules, each of them containing a complete SAM of arylthiols on gold nanospheres, gave rise to the distinct peaks for identification [34].

The choice of the Raman reporter molecules typically depends on the specific application. Fluorescent chromophores generallyimprove the sensitivity due to the resonance effect, however, their packing density on the surface is typically much lower than for SAMs. For multiplexing applications a SAM of Raman reporters is generally considered to be more reliable because of the low possibility of the overlap of Raman peaks, as shown in Fig. 17.8B.

In addition to the multiplexing capacity, quantification is a further very important aspect for the design of SERS nanotags. Various groups have demonstrated that the SERS signal response is proportional to the number of SERS nanotags [48,49]. To accomplish the quantification and avoid the interference of different cross-sections from Raman reporters, isotopic Raman reporter molecules were designed, which show distinct Raman peaks with similar Raman signals, indicating the great potential of isotopic Raman reporter molecules for quantitative detection [50,51].

17.2.3 Surface-enhanced Raman scattering nanotags: protection and stabilization

Although the protection and stabilization of SERS nanoprobes are not strictly necessary, it is highly desired because of the great benefit from the encapsulation. For instance, the resulting chemical and mechanical stability of the colloidal particles allows particle storage and prevents particle aggregation. Furthermore, the desorption of Raman reporter molecules from the metal surface as well as the adsorption of spectrally interfering molecules from the

environment to the surface can be eliminated. More important, the functional chemical groups on the shell surface could provide secure anchoring points for the conjugation of biomolecules. In other words, the protective shell will significantly improve colloidal stability, water solubility, biocompatibility, and provide the functional groups for further bioconjugation. Various encapsulations are available, including hydrophilic SAMs [46,52,53], proteins as biopolymers [54], organic polymers [55–59], and silica [9,47,60–63]. In this section, different encapsulation approaches for SERS nanotags will be briefly introduced.

17.2.3.1 Hydrophilic self-assembled monolayers for stabilizing surface-enhanced Raman scattering nanotags

The first simple and useful stabilization strategy is by employing a hydrophilic unit to form the SAMs on SERS nanotags as depicted in Fig. 17.9A, which is based on terminal hydrophilic spacers attached to the SAM [46,52,53]. This approach combines the stabilization of the SAM by hydrophilic spacers for obtaining water-soluble SERS probes and the option

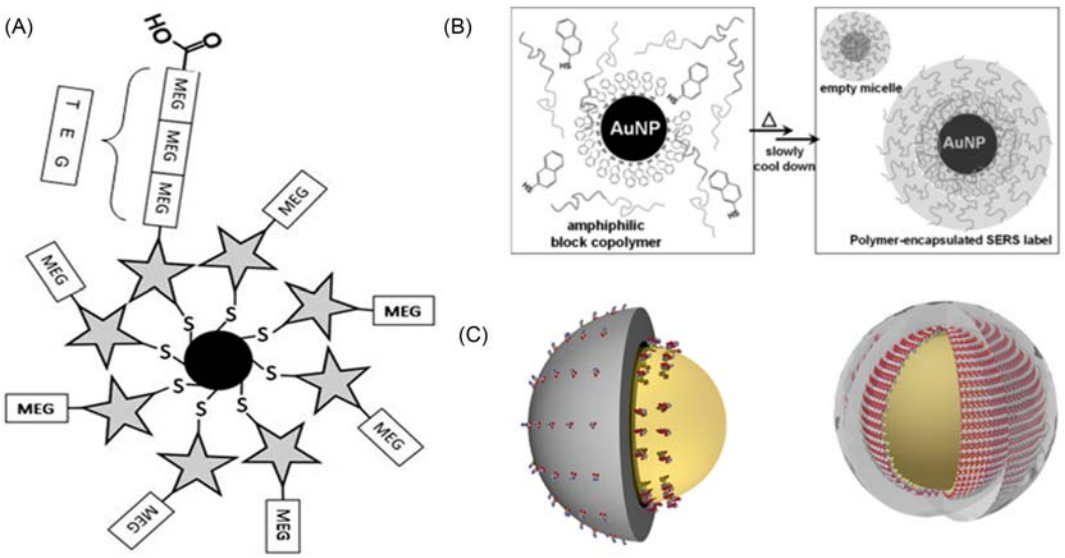


Figure 17.9

Schematic illustration for the hydrophilic self-assembled monolayer (SAM)-protected (A), polymer-encapsulated (B), and silica-encapsulated (C) surface-enhanced Raman scattering (SERS) nanotags. Source*: Cite from (A) C. Jehn, et al., Water soluble SERS labels comprising a SAM with dual spacers for controlled bioconjugation, Phys. Chem. Chem. Phys. 11 (34) (2009) 7499–7504; M. Schütz, Synthesis of glass-coated SERS nanoparticle probes via SAMs with terminal SiO precursors, Small. 6 (6) (2010) 733–737; (B)* M. Yang, Development of polymer-encapsulated metal nanoparticles as surface-enhanced raman scattering probes, Small 5 (2), 2009, 198–202; *and (C) S.P. Mulvaney, Glass-coated, analyte-tagged nanoparticles: a new tagging system based on detection with surface-enhanced Raman scattering, Langmuir 19 (11) (2003) 4784–4790.*

for further controlled bioconjugation. Stabilization of the SAM is achieved by attaching hydrophilic monoethylene glycol (MEG) units with terminal OH groups. A small portion of Raman reporter molecules is conjugated to longer triethylene glycol (TEG) units with terminal COOH moieties for bioconjugation [46]. Four advantages result from this strategy: (1) the colloidal surface is completely covered exclusively by Raman labels for maximum sensitivity; (2) the entire SERS label is hydrophilic and water soluble due to the MEG/TEG units, that is, independent of a particular Raman reporter molecule; (3) increased sterical accessibility of the SAM for bioconjugation via the longer TEG spacers with terminal COOH groups; and (4) the option for controlled bioconjugation by varying the ratio of the two spacer units (MEG-OH: TEG-COOH). A further advantageous property of the hydrophilic EG units is the minimization of nonspecific binding. This is a highly important aspect in many biological and biomedical applications: the binding selectivity is determined by the target-specific binding molecule and should not be diminished by nonspecific binding of the labeling agent, which leads to false-positive results.

17.2.3.2 Polymer and biopolymer-stabilized surface-enhanced Raman scattering nanotags

Stabilization by polymers is a very widely used approach to protect nanoparticles. For instance, Nie and coworkers investigated the protection of SERS nanotags with a layer of thiolated polyethyleneglycol (PEG) [64]. After the binding of the Raman reporters, thiolated PEG was then covalently bound on the SERS particles to avoid aggregation and for further bioconjugation. The Biomedical/Life Sciences Division of the Digital Health Group from Intel Corporation has introduced a concept of Raman reporter molecules−induced controlled aggregation of AgNPs called composite organic−inorganic nanoparticles (COINs). A shell of bovine serum albumin (BSA) protecting COINs was further employed to stabilize the SERS nanotags and for bioconjugation [54,65]. Similar, McCabe et al. developed a SERRS- labeled bead to encapsulate aggregates of SERRS-active nanoparticles into a functional polymer shell for stabilizing the nanoparticles as well as providing the functional group, such as free carboxylates, to conjugate with a wide variety of biomolecules [56]. In contrast, Chen and coworkers [55] developed polymer-encapsulated SERS nanotags by using an amphiphilic diblock copolymer [polystyreneblock-poly (acrylic acid)] to coat the nanoparticles' surface through self-assembly as shown in Fig. 17.9B. The advantages are that polymer shells with uniform thicknesses were easily prepared with a one-pot synthesis. The obtained polymer-encapsulated SERS nanotags provide protection against aggregation even in solutions with high ionic strength and against chemical oxidation as well as good SERS signals.

17.2.3.3 Silica-encapsulation of surface-enhanced Raman scattering nanotags

Among the various encapsulations, silica is attractive because of its high mechanical stability and the option for long-term storage. Mulvaney, Natan, and coworkers have introduced the concept of silica encapsulation for SERS probes through coadsorption of

Raman labels and SiO_2 precursors (typically in a 1:20 stoichiometry) on the surface of metal nanoparticles surface [48], which will leads to a submonolayer coverage of the surface with Raman reporter molecules as shown in Fig. 17.9C (left), followed by silica encapsulation with a modified Stöber method [66]. Soon after, Doering and Nie presented a very similar approach for glass-coated nanoparticles [61]. Generally, a glass shell provides chemical and physical stability to the SERS labels. For example, a 20-nm-thick silica shell leads to a significantly increased lifetime of gold nanospheres in the presence of aqua regia: 3 hours in comparison to 15 seconds for the bare Au colloid [48].

Schlücker and coworkers have introduced the concept of silica-encapsulated SERS nanotags comprising a full SAM (Fig. 17.9C right), thereby combining the advantages of thechemical stability of a glass shell and the maximum and dense surface coverage of Raman reporters for higher sensitivity [9,47].

In principle, the majority of SERS nanotags including nanospheres, anisotropic nanoparticles, and assemblies could be encapsulated by a silica layer by the sol–gel approach. For instance, gold nanorods, gold nanostars, dimers of AuNPs, aggregates, and assemblies [28,67–70] have been demonstrated to be protected by silica encapsulation.

17.2.4 Surface-enhanced Raman scattering nanotags: bioconjugation to surface-enhanced Raman scattering nanotags

Bioconjugation is the last step before the application of SERS nanotags. Specifically, SERS nanotags must be conjugated to target-specific ligands (Fig. 17.1), such as antibodies for antigen recognition.

As discussed, the synthesis of SERS nanotags for bioconjugation starts from the preparation of metal colloids with specific plasmonic properties (Section 17.2.1). The addition of Raman reporter molecules to the nanoparticles leads to the formation of SERS nanotags (Section 17.2.2). Encapsulation of the SERS nanotags, primarily for protection and stabilization, is an optional additional step which is often desirable (Section 17.2.3). Materials such as hydrophilic SAMs, biopolymers, or synthetic polymers and silica can be employed as encapsulants.

Generally, biomolecules (antibody or nucleic acid) can be attached either directly to the Raman labels in the absence of a protective shell (Fig. 17.10A) or to the encapsulants such as a silica shell according to the terminal functional chemical groups (spacer) (Fig. 17.10B). In this section, we will discuss different strategies for the conjugation of biomolecules.

17.2.4.1 Direct conjugation of biomolecules in the absence of a protective shell

Direct conjugation of biomolecules to SERS nanotags could be accomplished either by electrostatic forces or covalent binding. Adsorption of proteins to colloidal gold based on

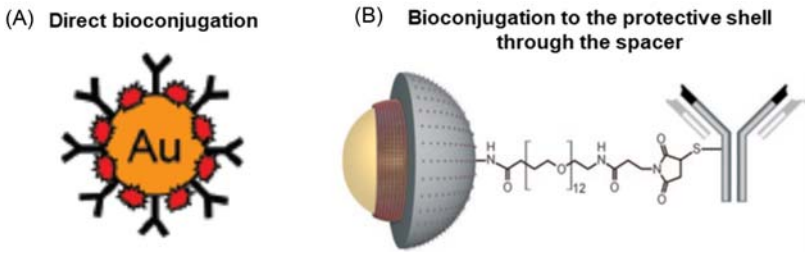


Figure 17.10
Schematic illustration for the bioconjugation of antibodies to surface-enhanced Raman scattering (SERS) nanotags directly to the metal surface (A) or to a protective shell (B). Source: *Cite from B. Küstner, et al., SERS labels for red laser excitation: silica-encapsulated SAMs on tunable gold/silver nanoshells, Angew. Chem. Int. Ed. 48 (11) (2009) 1950–1953 and G.F. Wang, Mixed monolayers on gold nanoparticle labels for multiplexed surface-enhanced Raman scattering based immunoassays, Anal. Chem. 81 (23) (2009) 9643–9650.*

electrostatic forces has been explored for decades to prepare immunogold probes as electron-dense markers [72,73]. Porter and coworkers employed a similar procedure for direct conjugation of antibodies and Raman reporter molecules on SERS nanoparticles via coadsorption [42,71]. However, problems exist when using the coadsorption procedure, including the nonspecific binding and "cross-talk" between different SERS nanotags [44]. Therefore covalent binding is necessary for the conjugation of biomolecules on SERS nanotags to resolve these problems. To achieve the covalent binding between SERS nanoparticles and biomolecules, the surface chemistry of the nanoparticles must be such that the ligand is fixed to nanoparticles and possesses terminal functional groups that are available for the conjugation reaction. Porter and coworkers developed mixed monolayers composed of two different thiolates, one for the Raman signal and the other with the functional terminal group (succinimidyl group) for covalent conjugation with antibodies [43].

17.2.4.2 Conjugation of biomolecules to surface-enhanced Raman scattering nanotags with a protective shell

1. Hydrophilic SAMs: water-soluble SERS nanotags stabilized by hydrophilic SAMs can be conjugated to biomolecules via the terminal carboxyl moieties of the longer TEG spacers [46]. After the activation of the carboxyl group by *N*-hydroxysuccinimide (NHS), the resulting active NHS esters are then conjugated to primary amines such as lysine residues in proteins. Anti-p63 antibodies were successfully attached to SERS nanotags according to this protocol and employed for ex-vivo tissue studies [74,75].
2. Polymer-encapsulated SERS nanotags: polymer-encapsulated SERS nanotags provide an easier way for bioconjugation via terminal functional groups of the polymer. For example, Nie and coworkers conjugated a single-chain variable fragment (scFv) covalently to the carboxyl groups of PEGylated AuNP coated with HS-PEG-COOH.

[64]. Generally, polymer shells on the SERS nanotags should expose a terminal functional group-rich surface to the surrounding in order facilitate the conjugation with target-specific ligands (antibody nucleic acids, etc.) [56].

3. Silica-encapsulated SERS nanotags: an effective bioconjugation approach for silica-encapsulated SERS nanotags is to chemically modify the particle surface with a terminal functional group through a silane coupling agent, to which biomolecules could be covalently linked [76]. Generally, the latter can be attached to SERS nanotags via specific functional moities including amino groups (via succinimide, isothiocyanate, or hydrazine), carboxyl groups (via carbodiimide), thiol groups (via maleimide or acetyl bromide), and azides (via click chemistry) [77]. So far, various ligands such as proteins, antibodies, and oligonucleotides have been successfully conjugated to silica-encapsulated SERS nanotags for targeted research [9,67,69].

Since SERS is really distance-dependent, only the Raman molecular moiety close enough to the nanoparticle surface will experience the drastic electromagnetic near-field enhancement and according to the SERS selection rule [8] the orientation of the molecules relative to the surface normal, z-components of the molecular polarizability tensor experience the largest enhancement (z being the axis parallel to the surface normal). Due to this distance dependence no spectral contributions from the target-specific ligands, for example, the amide I band around 1650 cm^{-1} for α-helices in the case of antibodies as ligands, are observed.

17.3 Immuno-SERS assay

iSERS assay has attracted significant attention for protein detection because of its high sensitivity. The high sensitivity is due to the strong plasmonic effect of the SERS nanotags as discussed in Section 17.2. More importantly, SERS nanotags have demonstrated their unique potential for protein detection [1,3,7,8,44] owing to the advantages of SERS nanotags over fluorescent dyes, including (1) the multiplexing capability for simultaneous target protein detection due to the narrow width of vibrational Raman bands, (2) quantification using the SERS fingerprint of the corresponding reporters, (3) the need for only a single laser excitation wavelength, and (4) high photostability. In this section, iSERS assays for protein biomarker detection will be discussed, ranging from the conventional sandwich immunoassay, dot-blot semisandwich immunoassays, and protein microarrays to microfluidic protein assay.

17.3.1 Sandwich immuno-SERS assay on solid substrates

A SERS sandwich immunoassay is a conventional sensor platform for protein biomarker detection with SERS nanotags. Porter and coworkers were the first to report on a series of works on sandwich iSERS assays for the detection of various protein biomarkers

[42,43,71,78–80]. The detection antibody was directly conjugated to the surface of the AuNPs of the SERS nanotags, also called external reporter labels (ERLs), as discussed in Section 17.2.4 (Fig. 17.10A). The bioconjugation is via either electrostatic forces or covalent binding using dithiobis (succinimidyl propionate) [43] or carboxyl-PEG-SH (poly (ethylene glycol) 2-mercaptoethyl ether acetic acid) [64]. To fabricate the sandwich SERS platform, the substrate could be glass or a gold film (Fig. 17.11A). Additional studies have shown a higher sensitivity with a gold film as the substrate because of the plasmonic coupling effect (gap mode) between the gold film (SPP: surface plasmon polariton) and the SERS labels (localized surface plasmon) [71,80]. Followed by the capture of the target antigen by the captureantibody immobilized on the surface, ERLs can now form the sandwich complex and report on the presence of the target proteinvia the characteristic spectral barcode of the SERS tag.. With a similar platform design, plenty of proteins have been detected including few clinically important biomarkers [44,78–80]. Additionally, the sandwich iSERS assay format has also been performed with magnetic beads conjugated to the capture antibody for capturing and pulling down the antigen. (Fig. 17.11B). With this scheme [34], for example, a cytokine has been sensitively detected using silica-encapsulated SERS nanotags (Fig. 17.11B).

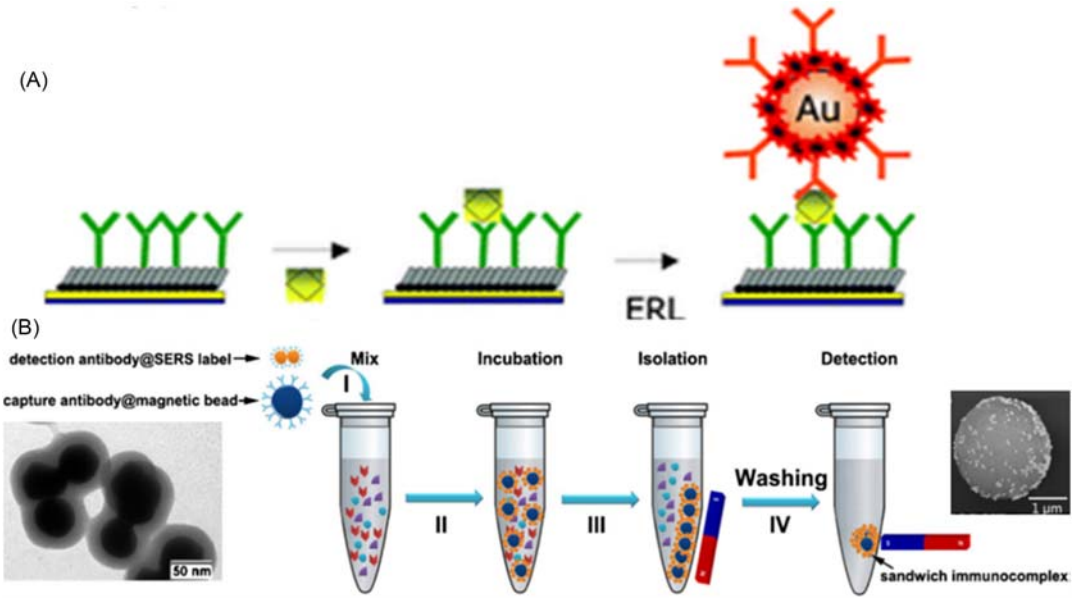


Figure 17.11

Schematic illustration of a typical surface-enhanced Raman scattering (SERS) sandwich immunoassay for protein biomarkers detection on a flat substrate (A) and on a magnetic bead (B). Source*: Cited from Y. Lai, et al., Raman-encoded microbeads for spectral multiplexing with SERS detection, RSC Adv. 5 (18) (2015) 13762–13767 and M.D. Porter, SERS as a bioassay platform: fundamentals, design, and applications, Chem. Soc. Rev. 37 (5) (2008) 1001–1011.*

17.3.2 Protein immuno-SERS microarray

A microarray is a platform that could detect proteins with high throughput using SERS with different Raman reporters. As a typical example indicated in Fig. 17.12, multiple protein targets have been immobilized on the surface, followed by the probe with the Raman reporter-labeled SERS nanoparticles [33]. Upon laser excitation, SERS signals from the reporter indicate the presence of the target proteins. A microfluidic assay (Section 17.3.4) is another platform with high-throughput properties for protein detection, as the design for the channels can separate each of the individual proteins into different channels, as well as enhance the reaction by controlling the flow conditions [81].

17.3.3 Dot-blot semisandwich immuno-SERS assay

Although an iSERS microarray has a high throughput, the requirement for professional training, the long incubation time, as well as the labor-intensive procedures have hindered the application of the platform in a clinical setting. To improve the efficiency of the

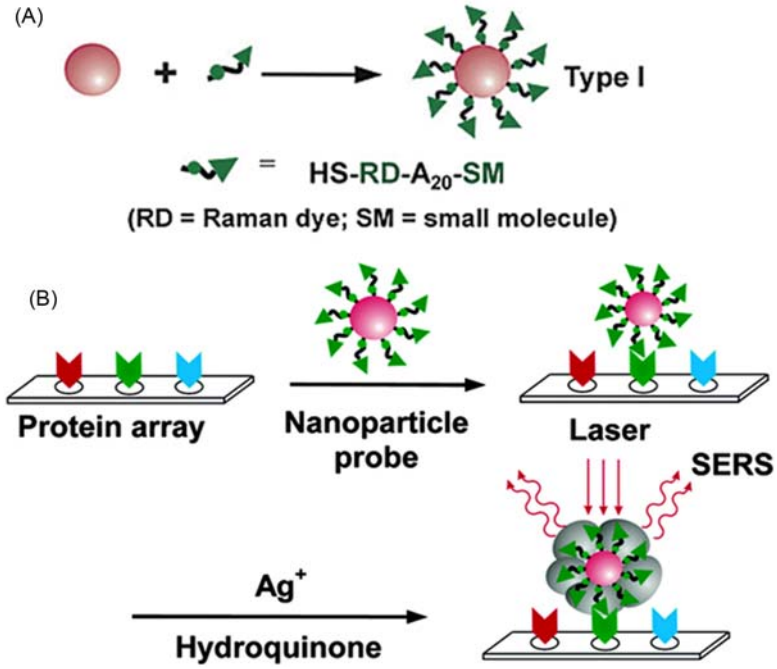


Figure 17.12
Raman reporter-labeled surface-enhanced Raman scattering (SERS) nanoparticles (A) and protein array (B) for high-throughput detection of proteins by SERS. Source: *Cited from Y.C. Cao, et al., Raman dye-labeled nanoparticle probes for proteins, J. Am. Chem. Soc. 125 (48) (2003) 14676–14677.*

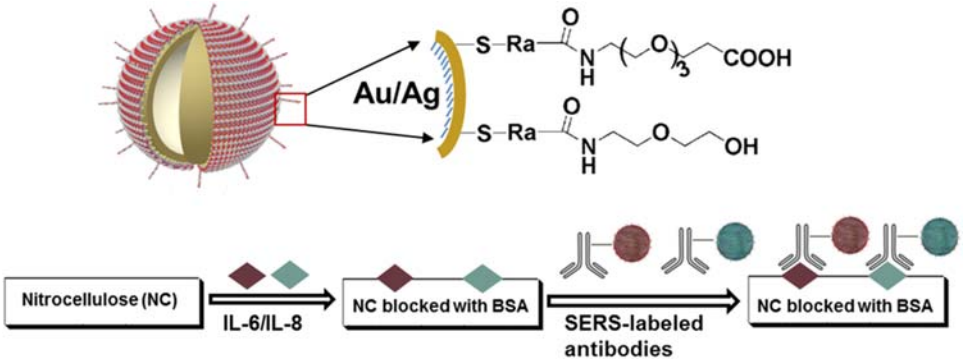


Figure 17.13
Dot-blot assay for direct protein detection by rationally designed surface-enhanced Raman scattering (SERS) nanoparticles. Source: *Cited from Y. Wang, et al., Femtogram detection of cytokines in a direct dot-blot assay using SERS microspectroscopy and hydrophilically stabilized Au—Ag nanoshells, Chem. Commun. 50 (21) (2014) 2711—2714.*

platform, a dot-blot assay for rapid, sensitive, and simultaneous protein detections with a simplified procedure was developed [82]. As shown in Fig. 17.13, two cytokines (interleukin-6 and interleukin-8) were detected simultaneously on the dot-blot assay with femtogram (fg) sensitivity, which was achieved by using Au/Ag nanoshells as the plasmonic substrate.

To minimize the assay time and to enable rapid detection, several strategies have been employed [79,83]. For instance, Driskell, Porter and co-workers developed an approach to increase the flux of antigen and SERS particles to the solid-phase surface by using a rotating capture substrate [79]. As illustrated in Fig. 17.14A, by controlling the rotation frequency for the capture substrate, the reaction kinetics can be improved quickly. Instead of AuNPs, gold nanorods were also used as plasmonic nanostructures for ERLs to improve the sensitivity. The assay time was reduced from 24 hours to 25 minutes, however, there was a 10-fold loss of sensitivity compared to the conventional SERS sandwich immunoassay. To further improve the simplicity of the assay, the syringe pump iSERS assay platform (Fig. 17.14B) was developed to overcome the diffusion-limited binding kinetics that often impedes rapid analysis in conventional SERS immunoassays. The assay time was reduced from 24 hours to 10 minutes with a 10-fold improvement in detection limit [83].

17.3.4 Microfluidic immuno-SERS assay

Despite these attempts being successful in reducing assay times, nonspecific adsorption of nontarget molecules still remains the biggest challenge in immunoassays for protein

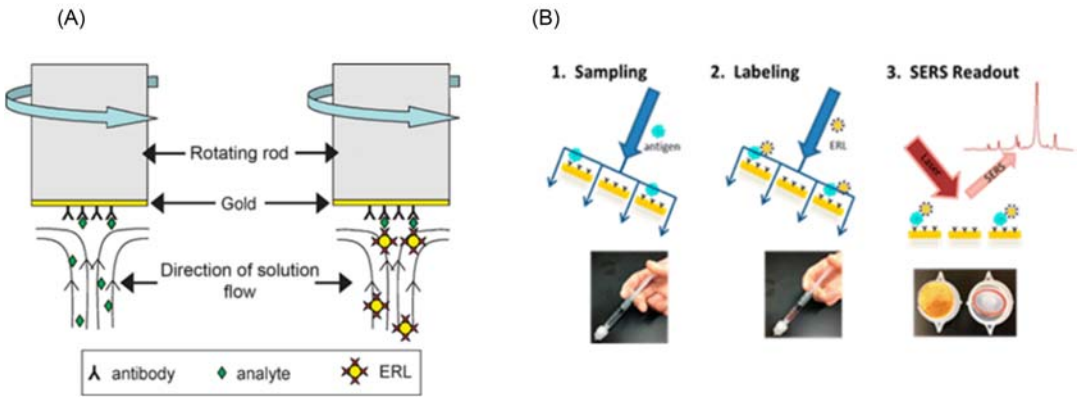


Figure 17.14

Two typical platforms with a rotating capture substrate (A) and a syringe pump (B) to reduce the assay time. Source*: Cited from J.D. Driskell, et al., Surface-enhanced Raman scattering immunoassays using a rotated capture substrate, Anal. Chem. 79 (11) (2007) 4141–4148 and M.A. Penn, D.M. Drake, J.D. Driskell, Accelerated surface-enhanced Raman spectroscopy (SERS)-based immunoassay on a gold-plated membrane, Anal. Chem. 85 (18) (2013) 8609–8617.*

detection. To circumvent this problem, Wang et al. proposed an innovative platform [84] that utilizes nanoscaled alternative current electrohydrodynamic (ac-EHD)-induced surface shear forces to enhance the capture efficiency as well as significantly reduce the nonspecific binding of the molecules on the surface (Fig. 17.15A). Meanwhile, to improve the sensitivity, silica-coated gold/silver nanoshells have been employed as the SERS nanotags. It was found that the detection limit can go down to 1 fg/mL for the detection of HER2. Following the further improvement in the design of the microfluidic channel (Fig. 17.15B), simultaneous detection of four biomarkers was achieved both from serum and patient samples [85].

Due to the great advantages of SERS nanotags in their sensitivity and multiplexed capability with only one single laser excitation and photostability, it is expected that SERS nanotags and iSERS assays will find more applications in point-of-care diagnosis for protein detection by using rationally designed plasmonic nanostructures. SERS nanotags have also been recently used in lateral flow immunoassays which are widely used in point-of-care testing with naked-eye detection since this assay format is known to be robust and easy to use. [86–91]. Thus far, the time-consuming readout using expensive Raman microscopes prevented real-world point-of-care testing (POCT) as users do not want to wait an hour or several hours for a test result. Schlücker and co-workers developed a home-built SERS portable reader [92] for rapid, quantitative, and ultrasensitive point-of-care testing for human chorionic gonadotropin (hCG) hormone. The latter served as a model system in this proof-of-concept study using gold nanoassemblies as SERS nanotags (Fig. 17.16).

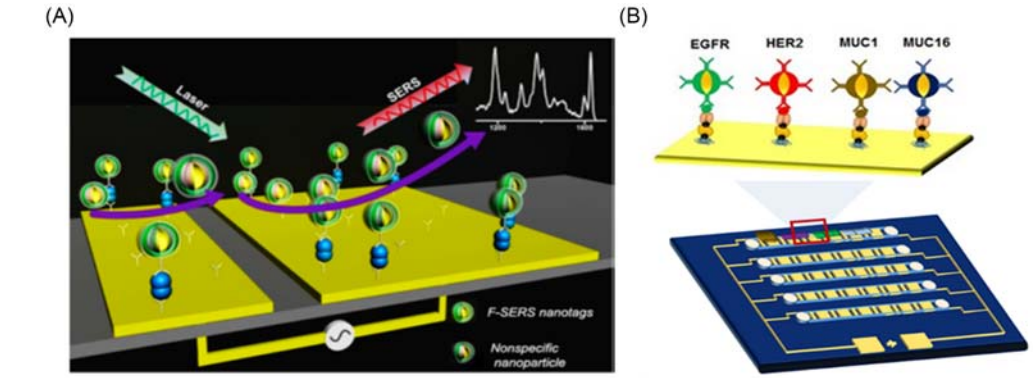


Figure 17.15

Scheme of the microfluidic device platform for protein biomarker sensing by using alternative current electrohydrodynamic (ac-EHD) to reduce the assay time and minimize the nonspecific binding by applying a potential on the unsymmetrical electrode pair. Dual functional Au/Ag nanoshells were used as surface-enhanced Raman scattering (SERS) nanotags for the breast cancer biomarker, human epidermal growth factor receptor 2 (HER2) detection (A), and for simultaneously detecting four biomarkers in a five-channel device (B). Source*: Cited from* Y. Wang, et al., Enabling rapid and specific surface-enhanced Raman scattering immunoassay using nanoscaled surface shear forces, ACS Nano 9 (6), (2015) 6354-6362 *and K. Kamil Reza, et al., Electrohydrodynamic-induced SERS immunoassay for extensive multiplexed biomarker sensing, Small 13 (9) (2017) 1602902.*

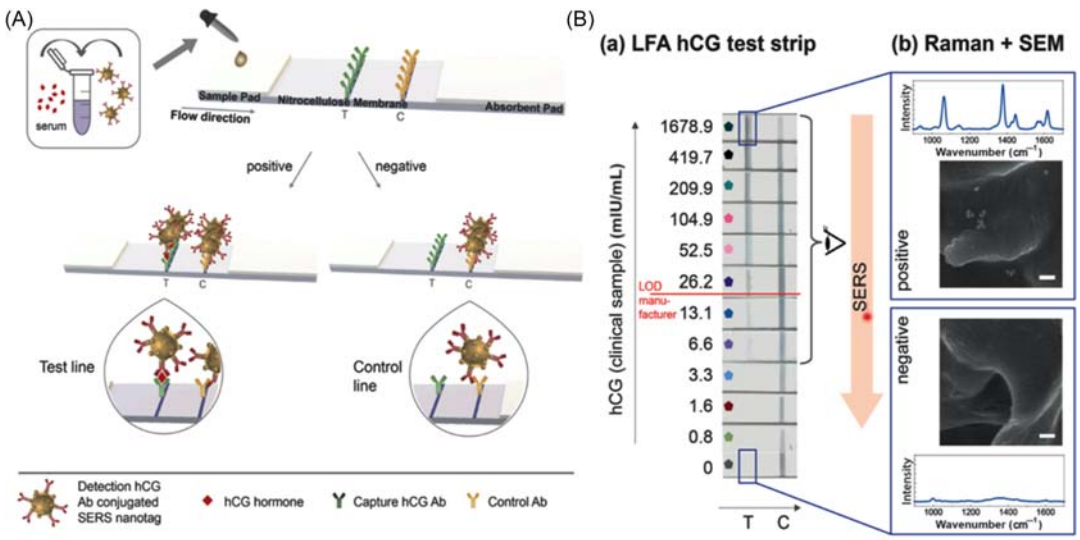


Figure 17.16

Schematic illustration of immuno-SERS (iSERS) assay for human chorionic gonadotropin (hCG) hormone detection on the pad using gold nanoasssemblies as surface-enhanced Raman scattering (SERS) nanotags (A), and concentration-dependent SERS-based lateral flow assay (LFA) of clinical serum samples from pregnant women. (A) Photographs of the test strips; (B) scanning electron microscopy (SEM) images (scale bar: 250 nm) of the test line (TL) from a positive test and a negative control together with the corresponding Raman spectra. Source*: Cited from V. Tran, et al., Rapid, quantitative, ultrasensitive point-of-care testing: a portable SERS reader for lateral flow assays in clinical chemistry, Angew. Chem. Int. Ed. 58 (2) (2019) 442–446.*

The detection is 15-times more sensitive than a commercially available lateral flow assay (LFA), which shows the great potential of SERS nanotags and iSERS assays in real-world applications. Most importantly, the acquisition time for the entire test line using line focus illumination is only 5 sec compared with dozens of minutes to several hours by conventional confocal Raman microscopy, i.e., this approach is two to three order of magnitude faster and therefore very attractive for future POCT where time is a key factor.

17.4 Immuno-SERS microscopy

iSERS microscopy, which employs immunostaining using antibody-SERS tag-conjugates in conjunction with Raman microspectroscopy, is a novel imaging method for protein localization in cells or tissue specimens (Fig. 17.17) [45]. The concept and staining process is analogous to conventional IF or ICC/IHC, that is, the target protein is recognized by a labeled antibody. Cells or tissue specimens, either fresh or frozen or formalin-fixed and paraffin-embedded (FFPE) sections, are usually incubated with blocking agents such as goat serum or BSA solution prior to staining, then incubated with labeled antibodies to form the

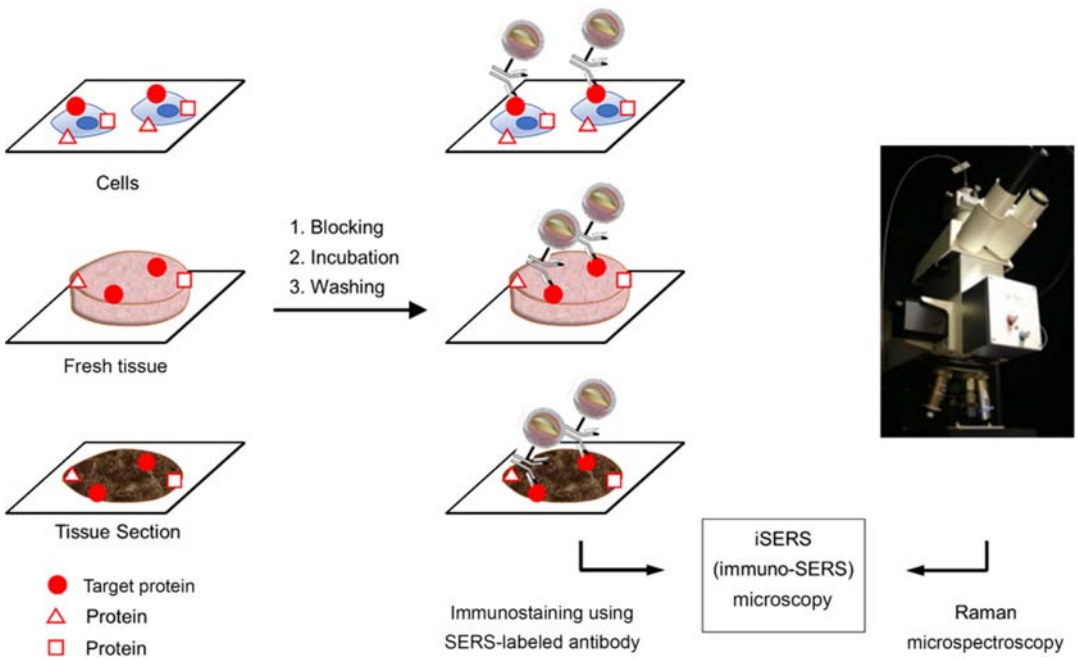


Figure 17.17
Principle of immuno-SERS (iSERS) staining and iSERS microscopy. The surface-enhanced Raman scattering (SERS) tag-labeled antibody recognizes the target protein via its antigen binding sites; After blocking, incubation, and washing, the antibody–antigen complex is detected by the Raman signal of the SERS tags.

corresponding antigen−antibody complexes, and subsequently unbound antibodies are removed by washing. The characteristic signal of the label, i.e., the characteristic Raman spectrum or spectral barcode of the SERS tag in the case of iSERS microscopy, indicates the localization of the target protein because the labels are attached to the antibody, which specifically binds to the target protein.

One key application of iSERS microscopy is profiling of cancer marker proteins on cell membranes or tissue specimens [3]. In the past decades, overexpression of a panel of proteins, such as epithelial cell adhesion molecule (EpCAM), epidermal growth factor receptor (EGFR), human epidermal growth factor receptor 2 (HER2), folic acid receptor, sialic acid (SA) glycans, etc., have been found to be closely associated with malignant and metastatic phenotypes for various types of cancer [93−96]. In particular, EpCAM has been considered as a tumor marker of circulating tumor cells (CTCs) and anti-EpCAM antibody-modified magnetic beads have been employed for CTCs enrichment [97]. Localization of cancer markers on cell membranes and tissue sections using iSERS microscopy has been intensively studied in the past decade for exploring its capability in precise cancer diagnostics. Cultured cells are ideal models for the study of iSERS microscopy, since they are relatively homogenous, can be repeatedly obtained in a large amount, and are accessible to be engineered for various purposes. FFPE tissue sections are the most widely used specimen for histopathological diagnosis in hospitals worldwide since formalin fixation results in cross-linking of macromolecules and keeps the tissue in an excellent morphological condition for histopathological analysis. More recently, several research groups have attempted to stain freshly excised tissue with target-specific SERS tags against cancer markers in order to identify residual tumor at the surgical margins, indicating the great potential of iSERS to be applied as intraoperative guidance [98,99]. In the following sections, we will introduce the studies using iSERS microscopy for cell/tissue-based protein localization. The “state-of-the-art” techniques will be highlighted, and the challenges for their clinical translation will be discussed.

17.4.1 Immuno-SERS microscopy for protein localization on cells

17.4.1.1 Localization of a single target protein on cells

The first study that combined immunostaining of cellular proteins using antibody-SERS tag-conjugates with Raman microspectroscopy was reported by Kim et al. in 2006 [69]. In this study, Raman reporter molecules (4-mercaptotoluene, 2-naphthalenethiol, thiophenol) were adsorbed on Ag NP-embedded silica spheres to fabricate SERS tags and then antibodies against HER2 or CD10 were covalently conjugated on the NP surface for target recognition. The SERS intensity was linearly dependent on the amount of SERS tags, allowing the possibility for quantification of the targets. Immunostaining using the antibody-conjugated SERS tags showed specific localization of HER2 and CD10 on the

membrane of Michigan Cancer Foundation-7 (MCF-7) cells. Soon after, target-specific SERS tags with various shapes and compositions were synthesized and applied for protein localization on different cell membranes. Gold colloidal NPs have advantages such as long-term stability, easily controllable size distribution, and high homogeneity, but their Raman EF is 100−1000 times lower than that of silver NPs. To prepare stable and sensitive SERS tags, Choo et al. fabricated bimetallic Au−Ag NPs by coating a silver layer over Raman reporter-labeled gold NPs (Au/Ag-Rhodamine 6G), followed by antibody conjugation and PEG stabilization. The SERS tag-labeled antibodies were applied for highly sensitive imaging of HEK293 cells expressing phospholipase Cγ1 (PLCγ1), showing the potential application of iSERS microscopy in early cancer diagnosis (Fig. 17.18A) [100]. In the following studies by the same group, crystal violet-labeled hollow Au nanospheres (Fig. 17.18B) [101] and 4-mercaptopyridine (MP)-labeled Au nanorods (Fig. 17.18C) [102] were conjugated with antibodies and used for targeted recognition of HER2 on MCF-7 cells, showing superior staining/imaging quality compared with spherical Au or Ag NPs. To further improve the detection sensitivity of SERS tags, Au@Ag nanorods [103] and core−satellite gold−silver nanocomposites [104] have been synthesized in two recent studies and applied for targeted imaging of cancer markers MHC class I chain-related molecule A (MICA) and EGFR, respectively. Bright SERS signals were individually observed in MICA-overexpressed osteosarcoma cells and EGFR overexpressed-human glioma cell lines, but not in the negative control cells.

Antibodies are typically employed as target recognition molecules in conventional immunostaining techniques. For the purpose of increasing stability, reducing cost, or decreasing size, novel recognition molecules such as affibodies, ScFv antibodies, aptamers, or peptides have also been synthesized for specific binding with the target molecules. Affibody molecules are antibody mimetics which can bind to target proteins with high affinity. Compared to antibodies, the molecular weight of which is ∼150 kDa, affibody molecules are more stable and have a lower molecular weight of ∼7 kDa. Gambhir and coworkers conjugated affibodies molecules onto Au core−silica shell SERS tags for molecular imaging of EGFR. They found that the SERS tags report EGFR-positive A431 tumor cells with a signal nearly 13-fold higher than EGFR-negative MDA-435S tumor cells [98]. Aptamers are single-stranded DNA or RNA with unique tertiary structures that enable them to specifically bind with cognate molecular targets. Their relatively simple chemical structures entail full conformational recovery even after thermal/chemical denaturation and therefore aptamers typically have a long shelf life [105]. AS1411 aptamer is a target moiety against nucleolin, which is overexpressed on the plasma membrane and in the cytoplasm of most of tumor cells, whereas normal cells retain a very low level. Wu et al. fabricated AS1411 aptamer-modified Raman reporter-labeled Au nanoflowers for target recognition of nucleolin. The SERS probes showed high affinity for MCF-7 cells with overexpressed nucleolin and can be used for selective tumor cell screening and tissue imaging [106].

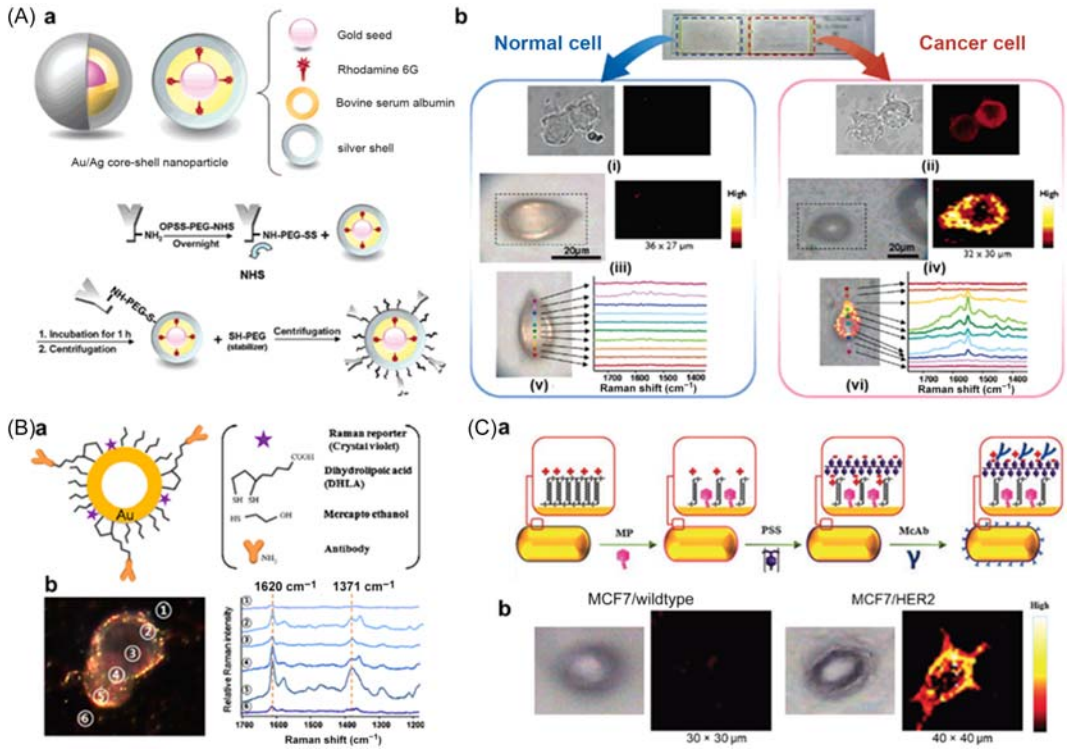


Figure 17.18

(A) (a) Schematic of silver-coated gold nanoprobes and antibody bioconjugation; (b) Fluorescence images (i, ii), surface-enhanced Raman scattering (SERS) images (iii, iv), and SERS spectra (v, vi) of normal HEK293 cells and phospholipase Cγ1 (PLCγ1)-expressing HEK293 cells. (B) (a) Schematic of antibody-conjugated hollow Au nanospheres; (b) Dark-field image of a Michigan Cancer Foundation/human epidermal growth factor receptor 2 (MCF/HER2) cell labeled with antibody-conjugated hollow Au nanospheres and corresponding Raman spectra at six different spots. (C) (a) Schematic illustration of the fabrication process of antibody-modified Au nanorods; (b) Bright-field images and SERS mapping images of a MCF-7/wild-type cell and a MCF-7/HER2(+) cell, respectively. Source*: Adapted with permission from (A) S. Lee, et al., Biological imaging of HEK293 cells expressing PLCgamma1 using surface-enhanced Raman microscopy, Anal. Chem. 79 (3) (2007) 916–922; (B) S. Lee, et al., Surface-enhanced Raman scattering imaging of HER2 cancer markers overexpressed in single MCF7 cells using antibody conjugated hollow gold nanospheres, Biosens. Bioelectron. 24 (7) (2009) 2260–2263; (C) H. Park, et al., SERS imaging of HER2-overexpressed MCF7 cells using antibody-conjugated gold nanorods, Phys. Chem. Chem. Phys. 11 (34) (2009) 7444–7449, copyright 2009 The Royal Society of Chemistry.*

Some small molecules which show specific binding affinity with proteins on cell membranes have also been utilized for target recognition in iSERS microscopy. Folate is a basic component of cell metabolism and DNA synthesis/repair. Rapidly dividing cancer cells require an increased amount of folate to maintain DNA synthesis, and the expression

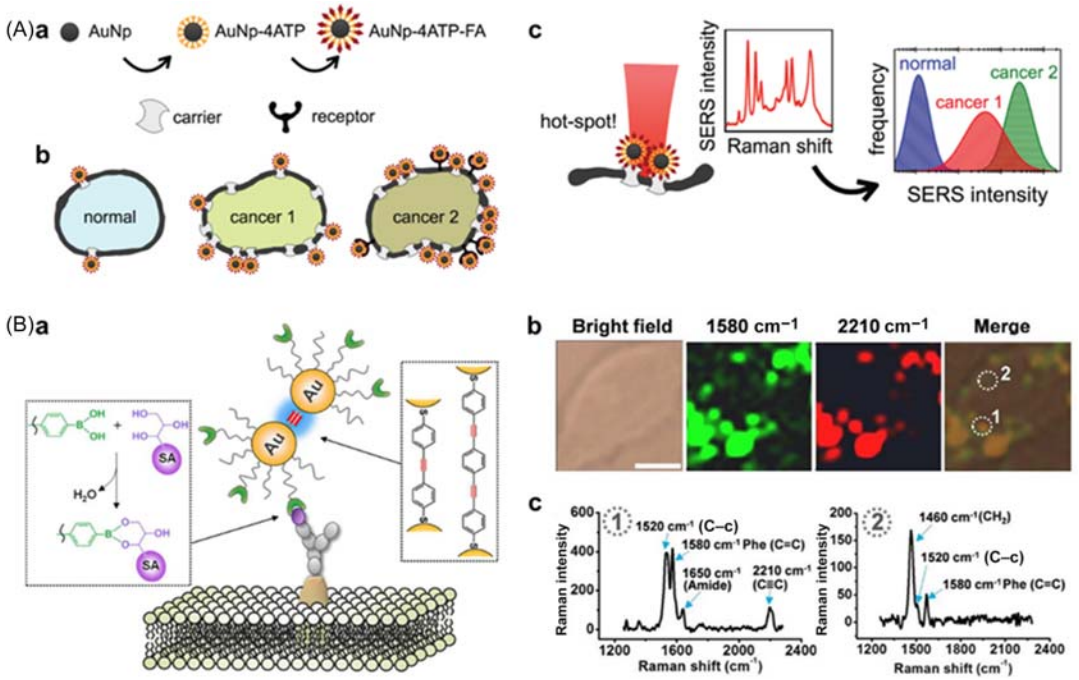


Figure 17.19

(A) Scheme showing modification of gold nanoparticles (AuNPs) with 4-aminothiophenol (4-ATP) and folic acid (a), selective cellular binding of surface-enhanced Raman scattering (SERS) tags on three cellular populations with different levels of folate receptor expressions (HaCaT "normal" cells, PC-3 "cancer 1" cells, HeLa "cancer 2" cells) (b), and SERS measurements on the cells, which allowed the discrimination of the cell populations based on the different density of folate receptors and carriers on the membrane (c). (B) (a) Schematic illustration of SERS tags (alkyne-bridged Au NP dimers) conjugated with phenylboronic acid (PBA) for specific profiling of sialic acids (SAs) on cell membrane; (b) Profiling SA expression on the membrane of a single HeLa cell incubated with SERS tags (the SERS images were analyzed in 1580 and 2210 cm^{-1} channels). Scale bar: 10 μm. (c) Raman spectra of the point 1 and point 2 as indicated in the merged image in (b). Source: Adapted with permission from (A) C. Fasolato, et al., Folate-based single cell screening using surface enhanced Raman microimaging, Nanoscale 8 (39) (2016) 17304-17313, copyright 2016 The Royal Society of Chemistry and (B) H. Di, et al., High precision profiling of sialic acid expression in cancer cells and tissues using background-free surface-enhanced Raman scattering tags, Anal. Chem. 89 (11) (2017) 5874−5881, copyright 2017 American Chemical Society.

levels of folate receptor alpha are usually elevated in specific malignant tumors compared to normal cells [95]. Fasolato et al. coated AuNPs with 4-aminothiophenol (4-ATP) and then conjugated them with folic acid molecules. After treating incubated cells with the folate receptor-targeting SERS tags, they were able to distinguish two cancer cell lines (HeLa and PC-3) from one normal cell line (HaCaT), which have different expression levels of folate-binding proteins (Fig. 17.19A) [108]. Increased expression of SA containing

glycoproteins is symptomatic in cancer progression, and therefore SA may serve as a potential tumor marker [96]. Phenylboronic acid (PBA) induces strong esterification toward SA to produce stable cyclic structures on the cell membrane. In a recent study reported by Liu and coworkers, PBA-modified SERS tags were fabricated and successfully applied for high-precision profiling of SA expression in cancer cells (Fig. 17.19B) as well as differentiation of clinically relevant cancer tissues with various metastasis degrees [107]. Alkyne-bridged plasmonic dimers were synthesized here as background-free SERS tags using alkyne-functionalized Raman reporters terminated with sulfurs on both ends. Alkyne is an exogenous moiety that exhibits a single vibrational peak in the cellular Raman-silent region (between 1800 and 2800 cm^{-1}), in which no Raman signals are generated from endogenous species, thus eliminating background interference and the need for spectral unmixing. However, it should be noted that for SERS nanotags, which are bright enough, there is no need to employ Raman reporters with signatures in the cellular Raman-silent region as the SERS signal of the tag is orders of magnitudes larger than the conventional Raman scattering from cells and tissues. The probe-treated cells were imaged with a confocal Raman microscope and the SERS false-color images were constructed with the 1580 (phenyl C = C bonds, green) and 2210 cm^{-1} (C≡C moieties, red) channels, respectively. As shown in Fig. 17.19B b and c, cellular imaging with a 1580-cm^{-1} channel led to background noise, due to endogenous proteins also containing a large number of phenyl groups with similar intense vibrational peaks, especially when they are in close vicinity to the SERS probes, thus producing spectral overlap. Other endogenous species, such as cytochrome and lipid, might also cause spectral interference. In contrast, C≡C produces a single strong and sharp peak at 2210 cm^{-1}, where the endogenous species as well as contaminates do not generate any signals, thus avoiding background interference. The authors have also found that Prussian blue (PB) possesses a strong and sharp single-band (2156 cm^{-1}) in the cellular Raman-silent region, and simultaneously holds a strong UV−vis absorption band between 500 and 900 nm, which is resonant with the incident laser wavelengths, and therefore can be employed as a highly sensitive and background-free resonant Raman reporter. Through assembly of PB onto plasmonic cores, a new SERRS probe was achieved and applied for high-sensitive immunoassay and cancer cell imaging [109].

17.4.1.2 Localization of multiple target proteins on cells

Colocalization of multiple proteins on a cell membrane is essential for analyzing protein−protein interactions during biological/pathological processes, and also critical for characterizing phenotypes of tumor cells, which is beneficial for precise diagnosis and offers guidance in personalized treatment. One important advantage of SERS tags over other external labels is their tremendous multiplexing capacity, which allows simultaneous detection of a series of target molecules through conjugating with the corresponding antibodies.

In a study performed by Kennedy and coworkers in 2011, Ag NPs were coated with 4-(mercaptomethyl)benzonitrile (MMBN), d7-mercaptomethyl benzene (DMMB), 4-(mercaptomethyl) ethynylbenzene (MMByne), and 4-(mercaptomethyl)nitrobenzene (MMBNO) to fabricate SERS tags with different colors. The 2230 cm^{-1} band of MMBN, the 965 cm^{-1} band of DMMB, the 2020 cm^{-1} band of MMByne, and the 1350 cm^{-1} band of MMBNO are unique and nonoverlapping, thus serving as good candidates for univariate identification of specific reporter molecules. The MMBN- and DMMB-labeled SERS tags were then conjugated to specific antibodies for recognition of two protein receptors, b2-adrenergic receptors and caveolin-3, showing 17% colocalization of the two receptors on the surface of rat cardiomyocytes. This study provides a way to study changes in the degree of receptor colocalization in response to ligand binding and thus adds insights into understanding the biological significance of these proteins in cell signaling [110]. Elevated expression of EGF, ErbB2, and insulin-like growth factor-1 (IGF-1) receptors are usually associated with tumorigenesis and poor prognosis of breast cancer. Lee et al. developed an iSERS-based multiplexing imaging technique to detect these growth factor receptors on breast cancer cell membranes (Fig. 17.20A). In this work, Raman reporter-coded silica-encapsulated hollow gold nanospheres (SEHGNs) were conjugated with specific antibodies against EGFR, ErbB2, and IGF-1R, respectively, and incubated with cultured cells. Expression of EGFR, ErbB2, and IGF-1R in the MDA-MB-468, KPL4, and SK-BR-3 human breast cancer cell lines was assessed by SERS imaging, showing expression patterns coinciding with western blot results, but with a simpler assay, a shorter time, and at the single-cell level [111]. As introduced above, alkyne-labeled SERS tags possess unique Raman emissions in a cellular Raman-silent region and provide background-free signals for cell imaging. Hu and coworkers fabricated alkyne-coded SERS tags and linked them with FA, luteinizing hormone-releasing hormone, or CALNNR8 peptide, respectively. The SERS tags were incubated with HeLa cells for 12 hours before imaging with Raman microspectroscopy. As shown in Fig. 17.20B b, the peaks at 2105, 2158, and 2212 cm^{-1} come from the OPE0, OPE1, and OPE2-coded SERS tags, respectively, and the efficiency of iSERS microscopy for triplex cellular imaging was clearly demonstrated in Fig. 17.20B c [112].

Based on its multiplexed targeting and imaging capability, iSERS microscopy has been increasingly applied for phenotypic analysis of CTCs recently. CTCs are tumor cells that shed from solid tumor sites and enter the circulatory system. Because blood collection is simple and minimally invasive, CTCs could be used as a real-time marker for disease progression and survival [113]. It has been found that the phenotypic information of CTCs represents the molecular characteristics of the corresponding tumor tissue, and the expression of epithelial-to-mesenchymal transition markers is closely associated with an increased capability for CTC migration and invasion [114]. Therefore a phenotypic profiling technique which may detect expression of a panel of protein markers is desired

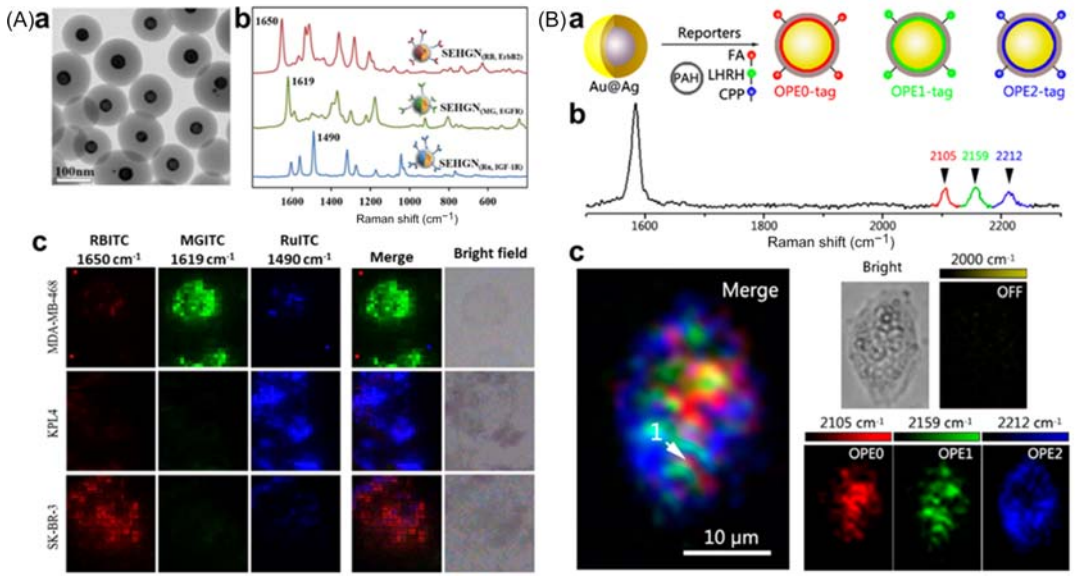


Figure 17.20

(A) (a) A TEM image of silica-encapsulated hollow gold nanospheres (SEHGNs). (b) Surface-enhanced Raman scattering (SERS) spectra of rhodamine B isothiocyanate (RBITC)-, malachite green isothiocyanate (MGITC)-, and bis(2,3′-bipyridine)-(5-isothiocyanato-phenanthroline) rutheniumbis(hexa-fluorophosphate) (RuITC)-coded Raman nanotags. The strongest Raman peaks at 1650 cm^{-1} (RBITC), 1619 cm^{-1} (MGITC), and 1490 cm^{-1} (RuITC) were used for Raman mapping of the respective receptors on breast cancer cells. (c) SERS mapping images of three different types of breast cancer cells. (B) (a) Scheme of SERS tags coded with OPE0 (*red*), OPE1 (*green*), and OPE2 (*blue*) and conjugated with different ligands. (b) Normalized Raman spectra of the three SERS probes. (c) Three-color SERS imaging of a HeLa cell. Source*: Adapted with permission from (A) S. Lee, et al., Rapid and sensitive phenotypic marker detection on breast cancer cells using surface-enhanced Raman scattering (SERS) imaging, Biosens. Bioelectron. 51 (2014) 238–243, copyright 2014 Elsevier; (C) Y. Chen, et al., Alkyne-modulated surface-enhanced Raman scattering-palette for optical interference-free and multiplex cellular imaging, Anal. Chem. 88 (12) (2016) 6115–6119, copyright 2016 American Chemical Society.*

for CTC analysis. Nima et al. fabricated a series of SERS tags by labeling Ag-coated Au nanorods with four different Raman reporter molecules: 4-mercaptobenzoic acid (4-MBA), *p*-aminothiophenol, *p*-nitrothiophenol, and 4-(methylsulfanyl)thiophenol. The SERS tags were conjugated with four kinds of antibodies against different cancer markers; anti-EpCAM, anti-CD44, anti-Keratin 18, and anti-IGF antigen (Fig. 17.21A). The Au/Ag nanorods provide more than two orders of magnitude of SERS signal enhancement compared to conventional Au nanorods. Using this set of highly sensitive target-specific SERS tags, single breast cancer cells could be identified in unprocessed human blood sample [115]. Recent studies have indicated that the molecular phenotype of CTCs evolves

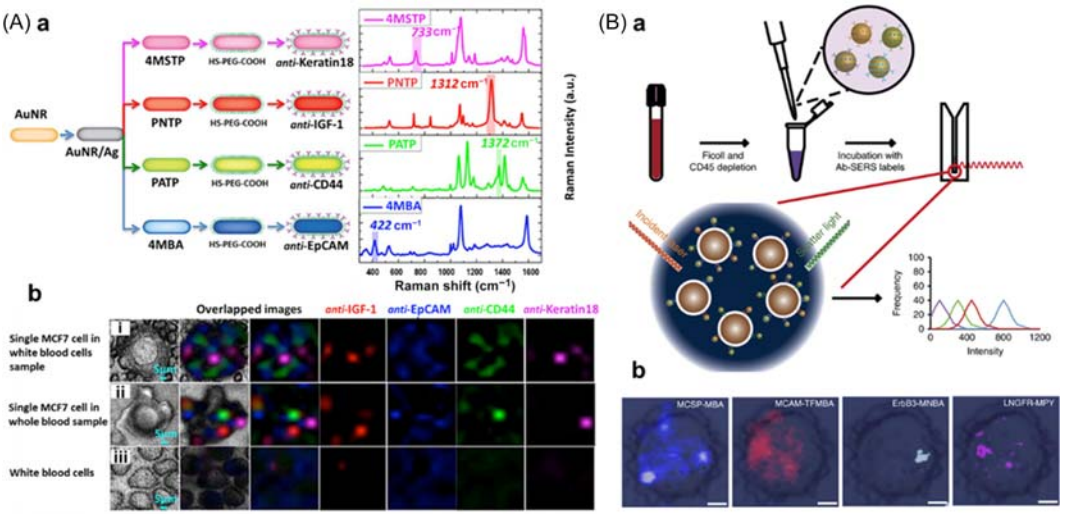


Figure 17.21

(A) (a) Schematic and Raman spectra of four antibody-modified surface-enhanced Raman scattering (SERS) nanotags. (b) Multicolor SERS analysis of a single Michigan Cancer Foundation-7 (MCF-7) cell among white blood cells (WBCs) (i), a single MCF-7 cell in whole blood (ii), and WBCs only (iii). (B) (a) Schematics showing experimental workflow of circulating tumor cell (CTC) detection and characterization with SERS tags. (b) SERS images showing the distribution of four surface markers on a single LM-MEL-64 cell. Source: *Adapted with permission from (A) Z.A. Nima, et al., Circulating tumor cell identification by functionalized silver-gold nanorods with multicolor, super-enhanced SERS and photothermal resonances, Sci. Rep. 4 (2014) 4752, copyright 2014 Springer Nature; (B) S.C. Tsao, et al., Characterising the phenotypic evolution of circulating tumour cells during treatment, Nat. Commun. 9 (1) (2018) 1482, copyright 2018 Springer Nature.*

dynamically with the progression of disease or during medical treatment. For instance, a HER2(+) subpopulation of CTCs was detected in the blood of HER2(-) breast cancer patients after multiple courses of therapy [117]. Therefore real-time monitoring of phenotypic evolution of CTCs is crucial for understanding the progression of cancer and clinical treatment guidance. Tsao and coworkers developed a SERS-based technique to characterize the phenotypic changes of melanoma cell lines as well as CTCs from clinical melanoma patients during molecular targeted therapies [116]. Four Raman reporter molecule−surface marker pairings were decorated on the surface of AuNPs (Fig. 17.21B): 4MBA with melanoma−chondroitin sulfate proteoglycan, 2,3,5,6-tetrafluoro-4-mercaptobenzoic acid with melanoma cell adhesion molecule, 4-mercapto-3-nitrobenzoic acid with erythroblastic leukemia viral oncogene homologue 3, and MP with low-affinity nerve growth factor receptor. The target-specific SERS tags were applied to monitor cellular phenotypic changes of melanoma cell lines harboring BRAF mutations (LM-MEL-33, LM-MEL-64, and LM-MEL-70) in response to PLX4720 (a BRAF inhibitor), showing

that all three tested cell lines formed distinct subpopulations after drug treatment. Then the blood samples from ten stage-IV melanoma patients were serially collected during therapy treatment and the CTC signature changes were monitored, showing that CTC populations shifted after treatment with dabrafenib and trametinib for 40 days and a totally different cluster was formed on day 48 [116]. Compared with current CTC detection technologies such as CellSearch system, the proposed SERS-based technique is extremely sensitive (10 cells in 10 mL of blood), highly multiplexed (simultaneously monitoring expression profiles of several surface proteins), and simple (does not need initial enrichment of CTCs), and therefore holds great potential to be translated into clinical use for disease progress and treatment efficiency monitoring.

17.4.2 Immuno-SERS microscopy for protein localization on tissue specimens

17.4.2.1 Immuno-SERS microscopy for protein localization on frozen or formalin-fixed and paraffin-embedded tissue sections

Detection and localization of tumor markers on tissue specimens is of central importance for cancer diagnostics. Currently, the most widely used method to detect protein expression on tissue specimens is IHC staining, that is, the target proteins are recognized by antigen−antibody interaction and then visualized through enzyme-induced pigment generation. Multiplexed immunostaining technique, in which two or more antigens can be detected simultaneously, is highly desired in tumor diagnosis and classification as it characterizes more precisely the spatial relationship of protein expression in the same tissue specimen, as well as helping to save the tissue samples [118]. However, current readout methods based on dyes or fluorophores are limited because of spectral overlap due to the broad electronic absorption/emission bands. Instead, iSERS microscopy, which combines immunostaining using SERS tag-labeled antibodies with imaging using Raman microspectroscopy, provides the possibility of simultaneous localization of multiple protein markers on a single tissue specimen.

The proof of concept study using iSERS microscopy for protein localization on FFPE tissue sections was performed by Schlücker and coworkers in 2006 [119]. In this work, aromatic thiols were adsorbed on Au nanoshells for the fabrication of SERS tags and the expression of prostate-specific antigen (PSA) in benign prostate tissue was analyzed using SERS tag-labeled anti-PSA antibodies. PSA was chosen here as a target protein because of its high expression levels in prostate tissue and its selective histological abundance in the epithelium of the prostate gland. After incubation of the prostate tissue sections with the SERS-labeled anti-PSA antibodies, SERS signals were detected specifically in the PSA(+) epithelium, but not in the PSA(−) stroma or lumen (Fig. 17.22A). Further applications of iSERS microscopy for localization of different proteins with various types of SERS tags appeared soon after this initial study [9,46,74,120−123]. Among them, Sun and coworkers fabricated

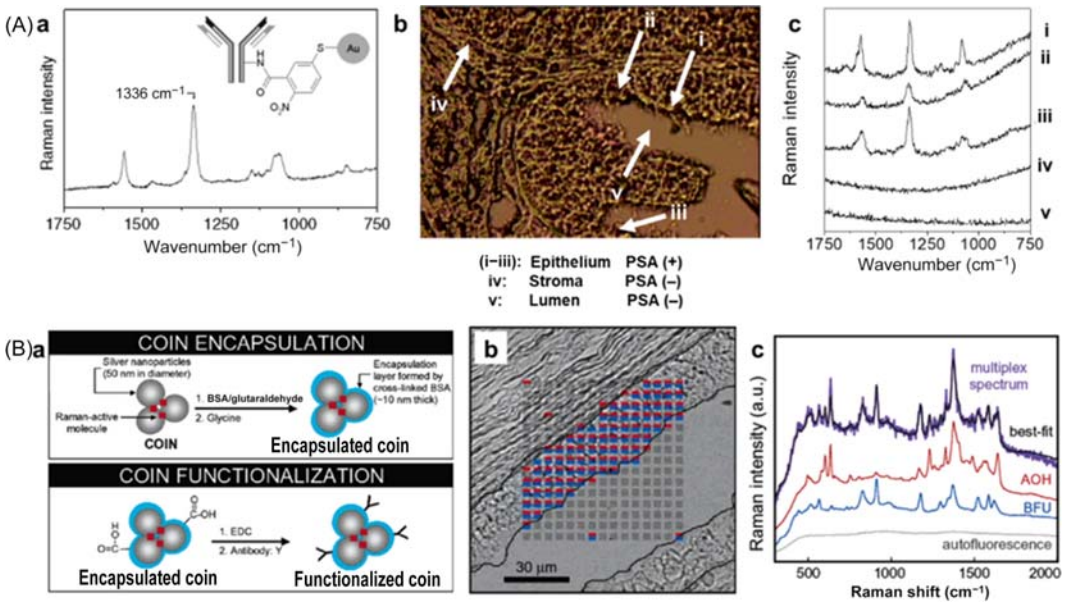


Figure 17.22

(A) Proof of principle study for immuno-SERS (iSERS) microscopy using surface-enhanced Raman scattering (SERS)-labeled primary antibodies for tissue diagnostics. (a) Structural formula of the target-specific SERS tag (*inset*) and its Raman spectrum in solution. (b) White light microscope image of epithelial tissue from the prostate. (c) Raman spectra recorded at the locations indicated by arrows: (i—iii) epithelium, (iv) stroma, and (v) lumen. (B) (a) Schematic illustration of composite organic—inorganic nanoparticle (COIN) encapsulation and functionalization procedure. (b) Prostate-specific antigen (PSA) expression patterns reported by two COINs [basic fuchsin-anti-prostate-specific antigen (BFU-anti-PSA) and acridine orange-anti-prostate-specific antigen (AOH-anti-PSA)] in a simultaneous duplex measurement. Both COINs report PSA-positive expression for nearly every point in the epithelium. Source*: Adapted with permission from (A) S. Schlücker, et al., Immuno-Raman microspectroscopy:In situ detection of antigens in tissue specimens by surface-enhanced Raman scattering, J. Raman Spectrosc. 37 (7) (2006) 719—721, copyright 2006 John Wiley & Sons, Ltd.; (C) L. Sun, K.-B. Sung, C. Dentinger, B. Lutz, L. Nguyen, J. Zhang, H. Qin, et al., Composite organic—inorganic nanoparticles as Raman labels for tissue analysis, Nanoletters 7(2) (2007) 351—356, copyright 2007 American Chemical Society.*

BSA-coated aggregated silver clusters (COINs) as SERS tags for PSA localization in prostate tissue sections [120]. In order to illustrate the multiplexing capabilities of iSERS microscopy in tissue staining, a two color-COIN staining against PSA was performed: SERS tags with two distinct Raman signatures [acridine orange (AOH) and basic fuchsin (BFU)] were conjugated to anti-PSA antibodies, and simultaneously applied to the tissue. The characteristic Raman signatures from both COINs were detected at almost every location in the epithelium, suggesting that steric hindrance from the SERS tags could be neglected (Fig. 17.22B). In a subsequent study performed by the same group, the staining

performances of SERS (COINs)- and fluorophore (Alexa)−PSA antibody conjugates on adjacent tissue sections were compared [124]. Similar staining behavior was observed using these two labels with comparable signal intensities. However, a lower staining accuracy was obtained by COINs, and the authors attributed this phenomenon to an elevated false negative rate of the SERS tags.

This issue was further discussed by Schlücker and co-workers in their study of rapid iSERS microscopy for tissue imaging with single-nanoparticle brightness [125]. They compared SERS signals from a glass-coated monomer, dimer, and trimer of gold NPs by correlative SEM/SERS experiments (Fig. 17.23A). Colloids were deposited on a framed silicon wafer and HR-SEM was employed to identify single NPs/clusters. With 30-ms integration time, the glass-coated trimer (red) and dimer (blue) yielded detectable SERS signals, while the glass-coated monomer (black) did not. When antibodies are labeled with a colloid mixture containing monomers, dimers, and trimers and applied for tissue staining, two situations will be observed: (1) "binding plus detectable SERS signal", where antibodies labeled with a dimer/trimer bind to the antigen, and a Raman signal is detectable within the defined acquisition time of the Raman mapping experiment. Overall, the presence of the antigen is correctly indicated by iSERS microscopy (Fig. 17.23B a); and (2) "binding but no detectable SERS signal", where a monomer is conjugated to antibodies for target recognition. Although the antibody correctly recognizes its antigen, no SERS signal is detectable due to the lack of plasmonic activity of the single AuNP, and a false-negative result is obtained (Fig. 17.23B b). Therefore fabrication of uniform SERS tags with single-particle brightness is essential for obtaining reliable and precise staining results.

IHC is generally performed in a wide field configuration for fast, global analysis of large tissue areas. In contrast, iSERS is typically performed in the Raman mapping mode for a spectrally resolved detection at each pixel. Although acquisition times per pixel have been systematically shortened from 1 second in the initial iSERS experiments on FFPE tissue in 2006 [119] to the millisecond regime in 2013 [125], due to the development of brighter SERS tags, iSERS inherits the disadvantages of normal Raman microscopy in the mapping mode for scanning large areas. One approach to improve the imaging efficiency of iSERS is to combine a fast global analysis by wide field fluorescence with a subsequent local iSERS analysis. To achieve wide field IF-guided local iSERS imaging, Schlücker and coworkers recently fabricated fluorophore/SERS tag-dual-labeled antibodies (Fig. 17.24A) for localization of HER2 on human breast cancer tissue sections (Fig. 17.24B) [126] and localization of PSA on human prostate tissue sections (Fig. 17.24C) [127], respectively. Correlative fluorescence-SERS images were recorded by an in-house integrated confocal Raman-fluorescence microscope: global expression of HER2 or PSA was first imaged using the fluorescence mode, then smaller regions were selected for spectral detection of each pixel using Raman mapping. The latter mode has the potential for a highly multiplexed analysis of cancer biomarkers on the same tissue section.

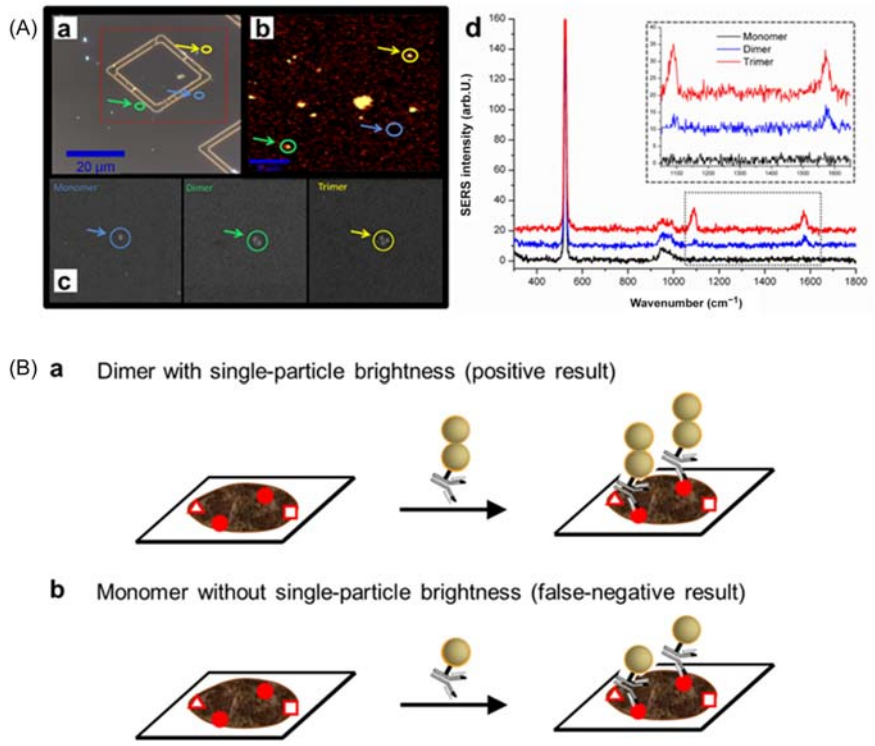


Figure 17.23

(A) Correlative dark-field/surface-enhanced Raman scattering (SERS)/scanning electron microscopy (SEM) single-particle experiments on glass-coated monomers and clusters of AuNPs. (a) Dark-field microscopy, (b) SERS microscopy, and (c) high-resolution scanning electron microscopy (HR-SEM) images. Only the glass-coated gold trimer (*yellow arrow*) and dimer (*green arrow*) exhibit single-particle SERS brightness (30-ms integration time), while the glass-coated monomer (*blue arrow*) does not. (d) Normalized single-particle SERS spectra of the glass-coated monomer, dimer, and trimer from (b). (B) The role of single-particle brightness in iSERS microscopy. (a) Clusters (dimers and trimers) of AuNPs with single-particle brightness yield positive results, while (b) monomers without single-particle brightness give false-negative results. Source*: Adapted with permission from M. Salehi, et al., Rapid immuno-SERS microscopy for tissue imaging with single-nanoparticle sensitivity, J. Biophotonics 6 (10) (2013) 785–792, copyright 2013 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.*

In 2014 Schlücker and co-workers fabricated highly sensitive, silica-encapsulated 4-nitrothiobenzoic acid (4-NTB)- or 4-MBA-encoded AuNP clusters as SERS tags, and performed two-color iSERS microscopy for colocalization of PSA and p63 in nonneoplastic prostate tissue FFPE specimens [128]. P63 is a tumor suppressor protein which is selectively expressed in the nuclei of basal cells of normal prostate glands, but is absent in prostate carcinomas, and has been applied as a biomarker for prostate cancer diagnosis [129]. As shown in Fig. 17.25A a, 4-NTB and 4-MBA are two Raman reporter molecules

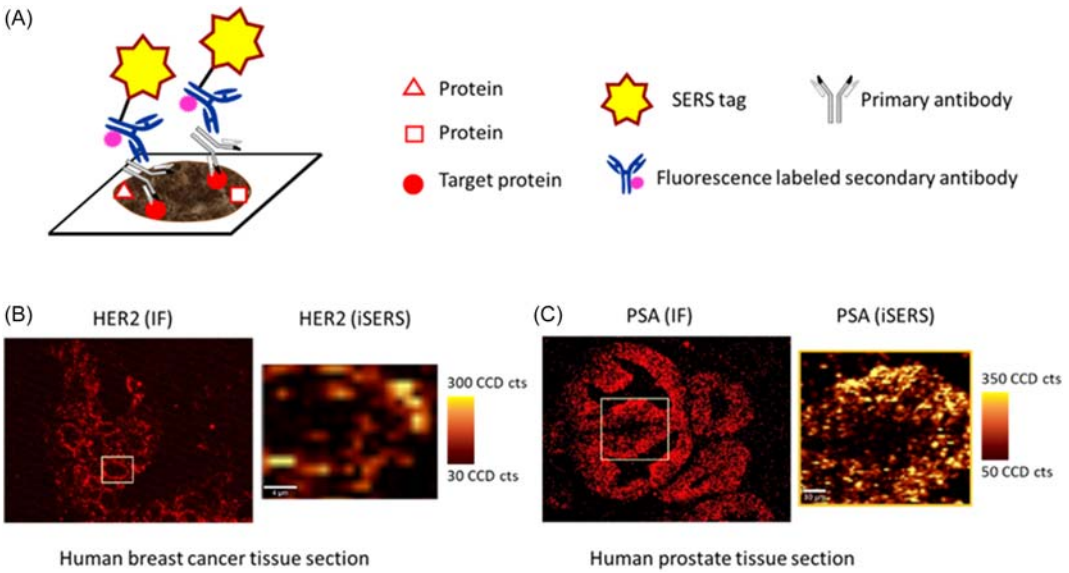


Figure 17.24

(A) Scheme of using fluorophore/surface-enhanced Raman scattering (SERS) tag-dual-labeled antibodies for combined immunofluorescence (IF) and immuno-SERS (iSERS) staining. (B) Correlative IF-iSERS imaging of human epidermal growth factor receptor 2 (HER2) expression on a human breast cancer tissue section. Scale bar: 4 μm. (C) Correlative IF-iSERS imaging of prostate-specific antigen (PSA) expression on a human prostate tissue section. Scale bar: 10 μm. Source: *Adapted with permission from (B) X.P. Wang, et al., iSERS microscopy guided by wide field immunofluorescence: analysis of HER2 expression on normal and breast cancer FFPE tissue sections, Analyst 141 (17) (2016) 5113–5119, copyright 2016 The Royal Society of Chemistry; (C) Y. Zhang, et al., Effect of antigen retrieval methods on nonspecific binding of antibody-metal nanoparticle conjugates on formalin-fixed paraffin-embedded tissue, Anal. Chem. 90 (1) (2018) 760–768, copyright 2018 American Chemical Society.*

with distinct Raman signatures, thus a univariate approach based on the integrated Raman intensities of single peaks (around 1340 cm^{-1} for 4-NTB and 1590 cm^{-1} for 4-MBA) is sufficient for spectral discrimination. The two-color iSERS imaging results were comparable with IHC staining: p63 is selectively observed in the nuclei of the basal cells, while PSA is abundant in the entire epithelium, but not in the stroma (Fig. 17.25A b). In another study performed by Knudsen and coworkers, AOH, BFU, Nile Blue A, and tetramethyl rhodamine isothiocyanate, which have multiple Raman bands, were used for encoding SERS probes. In this case, data analysis based on single Raman peaks is not feasible, instead a spectral fitting method that exploits the entire spectral signature to quantitatively extract individual probe signals from multiplex spectra needs to be developed [123]. Alkynes (C≡C), nitriles (C≡N), azides (N3), and deuterium (C−D) moieties exhibit single narrow bands in the biological Raman-silent spectral region (1800–2800 cm^{-1}). In a recent study performed by Li et al., a universal strategy for the one-pot preparation of

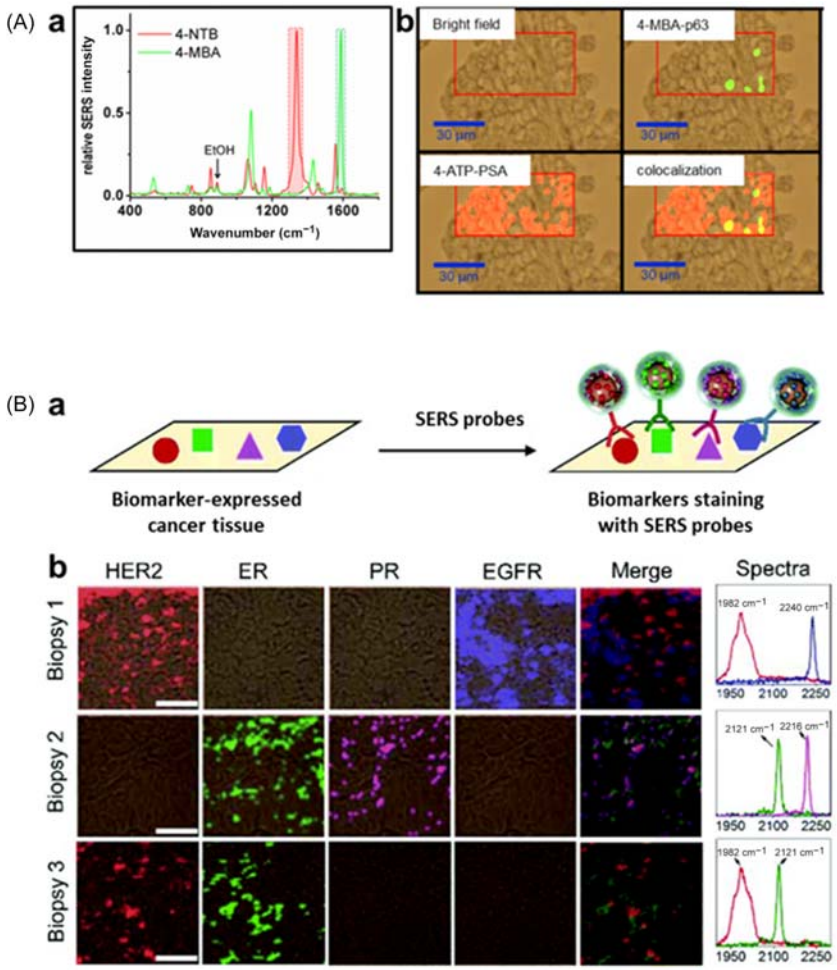


Figure 17.25

(A) (a) Surface-enhanced Raman scattering (SERS) signatures of gold nanoparticles (AuNPs) coded with 4-nitrothiobenzoic acid (4-NTB) and 4-mercaptobenzoic acid (4-MBA), respectively. (b) Two-color immuno-SERS (iSERS) microscopy for the colocalization of p63 (*green*) and PSA (*red*). (B) (a) Schematic of the multiplexed biomarker imaging. (b) Multiplexed iSERS images of four biomarkers on breast cancer biopsies: human epidermal growth factor receptor 2 (HER2) in *red*, oestrogen (ER) in *green*, progesterone (PR) in *magenta*, and epidermal growth factor receptor (EGFR) in *blue*. Scale bar: 50 μm. Source*: Adapted with permission from (A) M. Salehi, et al., Two-color SERS microscopy for protein co-localization in prostate tissue with primary antibody-protein A/G-gold nanocluster conjugates, Nanoscale 6 (4) (2014) 2361–2367, copyright 2014 The Royal Society of Chemistry; J. Li, et al., A universal strategy for one-pot synthesis of SERS tags, Nanoscale 10 (17) (2018) 8292–8297, copyright 2018 The Royal Society of Chemistry.*

SERS tags by incorporating Raman reporters during dopamine polymerization on NP surface was proposed [130]. Four background-free SERS tags were conjugated with corresponding antibodies and applied for multiplexed staining of biomarkers (HER2, ER, PR, and EGFR) in clinical breast cancer biopsies. Based on the easily differentiated single narrow bands in the biological Raman-silent region, Raman signatures from individual SERS probes can be conveniently discriminated (Fig. 17.25B). This method enables easy preparation of a large library of SERS probes for multiplexed protein localization, and promotes the clinical translation of iSERS imaging-based cancer diagnostics.

Finally, it is important to keep in mind that the chosen method of antigen retrieval has an influence on the staining quality in iSERs. Zhang, Schlücker and co-workers, for example, demonstrated that protein-induced eptiope retrieval (PIER) gives much better results compared to the conventionally employed heat-induced epitope retrieval (HIER) [127].

17.4.2.2 Immuno-SERS microscopy for protein localization on fresh tissue specimens

One new application of target-specific SERS tags developed in the last decade is iSERS microscopy-guided tumor resection. The complete removal of tumor tissue is essential in surgical oncology, but it is difficult to achieve when surrounding normal tissues have important functions and need to be maximally preserved. If postoperative pathology reveals that the resection margins are positive for carcinoma, reexcision surgery is typically required, which is not only time-consuming, but also brings additional stress and risk for the patient. Therefore intraoperative methods that can identify residual tumors and guide their complete removal during tumor-resection procedures are required in the clinic.

In 2011 Gambhir and coworkers demonstrated the feasibility of detecting a single biomarker (EGFR) ex vivo on tissues freshly resected from mice using topically applied target-specific SERS tags [98]. SERS signals on EGFR-overexpressing A431 tumors, low EGFR-expressing MDA-435S tumors, and epidermis adjacent to the tumor location were compared, showing a signal nearly 35-fold higher in A431 tumors than in MDA-435S tumors, and 7-fold lower in adjacent healthy tissue than in the positive tumors. However, the authors pointed out that several issues might cause variability of SERS imaging, including particle-to-particle variation in size and intensity, tumor heterogeneity, inconsistencies with the optics, and nonspecific binding arising from both NP and ligand sources. To improve the signal specificity and reliability, Liu and coworkers proposed a ratiometric biomarker detection assay that an equimolar mixture of targeted NPs and one untargeted NP are topically applied to the tissue specimen, and the ratio of signals from targeted NPs to untargeted NPs was used for image construction. This strategy allows an unambiguous assessment of the molecular expression of tumor markers, since the nontargeted SERS tag accurately mimics the nonspecific behavior of the targeted SERS tags in the mixture, and simultaneous scanning eliminates nonspecific binding effects as well as fluctuations of the laser power, charge-coupled device variation, and focal plane

[131]. Using this approach, binding potentials of EGFR in freshly resected healthy and xenograft U251/A431 tumor tissues in mice were estimated from kinetic modeling of targeted and untargeted NP concentrations in response to serial rinsing, showing comparable results with flow cytometry measurements [132]. In the following studies conducted by the same group, the influence of staining conditions such as the sizes of SERS probes and the incubation manners on the staining efficiency were investigated and optimized [133,134].

Due to the fact that disease biomarkers vary greatly between patients, within a tumor over time, and at different locations of a tumor mass, SERS probes that can determine a "quantitative molecular phenotype" of the tissue by simultaneously imaging a diverse panel of disease-related biomarkers are highly desired for improving the diagnostic accuracy [135]. Fig. 17.26 shows the schematic of an iSERS-based intraoperative imaging technique to rapidly identify residual tumors at the margins of freshly resected tissues for the guidance of lumpectomy. In a clinical study recently performed by Liu et al., 57 fresh specimens were incubated with mixed SERS probes for simultaneous quantification of four biomarkers (HER2, ER, EGFR, and CD44). Combined evaluation of these biomarkers enabled highly sensitive (89.3%) and specific (92.1%) detection of breast carcinoma. The excised surgical specimens are stained topically with a mixture of untargeted NPs and targeted NPs against corresponding tumor markers for 5 minutes, followed by a 10 seconds rinse step in PBS and SERS imaging. The entire procedure (including image processing) can be performed within 15 minutes, demonstrating the great potential of iSERS imaging for intraoperative guidance to reduce the rate of reexcision in cancer patients [99].

17.5 Summary and perspective

SERS nanotags as an alternative labeling approach for both immunoassays and ICF/ICH/IF microscopic imaging has experienced tremendous progress in the past few years. This is because of the unique features that SERS offers, including quantification, sensitivity, photostability, and dense multiplexing for simultaneous target detection. Significant advances in the rational design and synthesis of bright metal nanoparticles as the core of SERS nanotags have been accomplished, including the control over size/shape and nanoassemblies with multiple hot spots. Moreover, the diversity of molecules (from dye molecules to small molecules) provides a flexible choice of Raman reporter molecules depending on the applications. Along with the development of protective layers (polymer and silica) as well as spacer/linker molecules, SERS nanotags are expected to find more applications in biomedical and clinical applications.

iSERS assays have been developed from standard sandwich immunoassays, dot-blot semiimmunoassays, microarrays, and microfluidic assays, which provide flexible choices

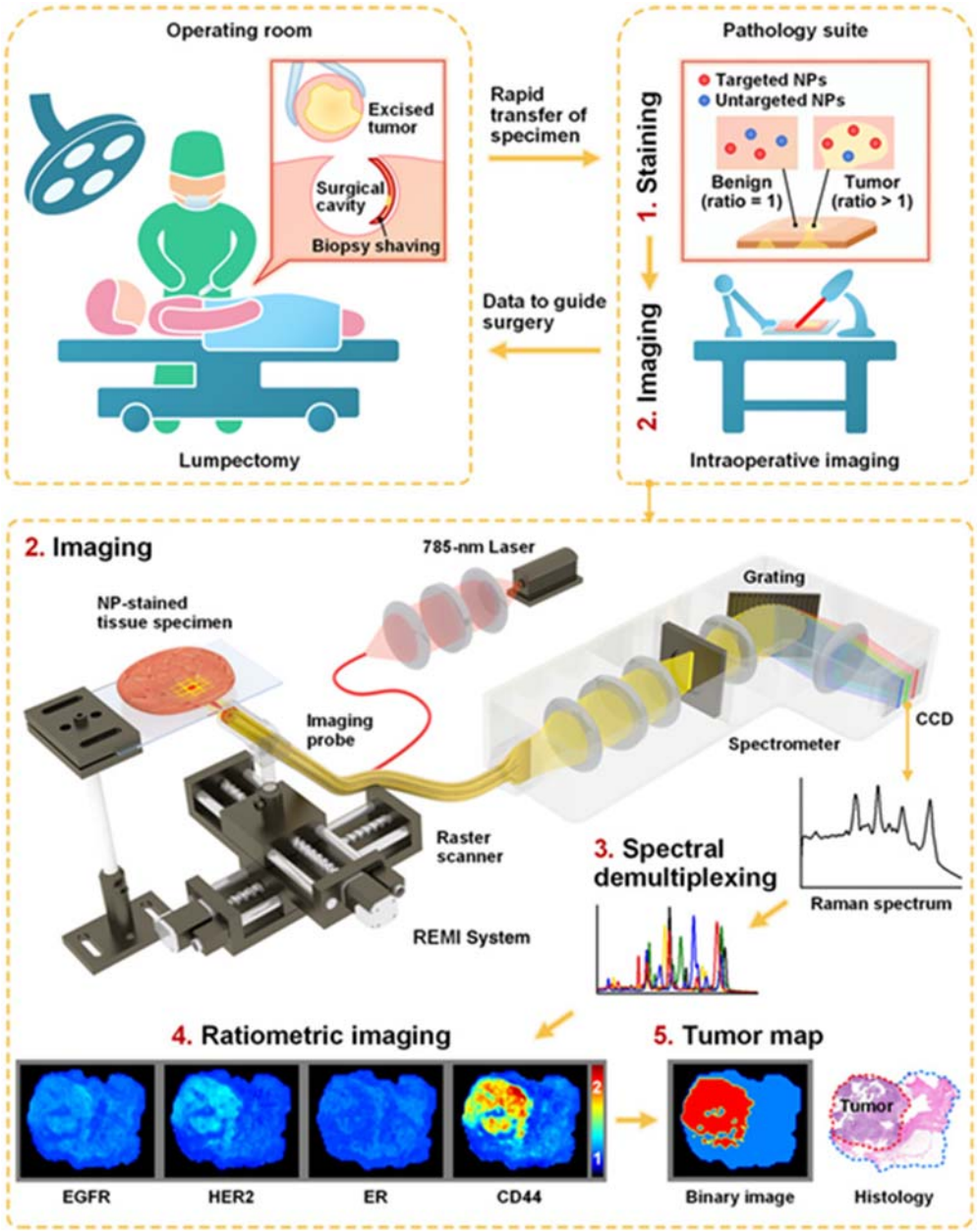


Figure 17.26

(A) Scheme showing iSERS microscopy-based intraoperative guidance of lumpectomy. A ratiometric strategy (right inset) quantifies biomarker expression on freshly resected tissues by comparing the signal from targeted NPs and nontargeted NPs. (B) Raman imaging system. a, Scheme of the spectral-imaging system; b, A photograph of the raster-scanned tissue-imaging device; c, A depiction of the structure of the targeted and nontargeted SERS NPs and d, the Raman spectra of the various SERS NPs. *Adapted with permission from Wang, Y., et al., Quantitative molecular phenotyping with topically applied SERS nanoparticles for intraoperative guidance of breast cancer lumpectomy, Sci Rep 6 (2016) 21242, copyright 2016 Springer Nature.*

for different applications. We thus expect the new direction of iSERS assay to be the miniaturization of the device with high-throughput for the rapid detection of protein biomarkers.

Since the first demonstration of iSERS microscopy, by combining Raman microscopy with SERS nanotags for tissue scanning in 2006 [119], it is currently in a rapid growth phase due to the development of instrument (rapid scanning and superresolution) and imaging methodologies. The challenge here is still the development of rapid confocal Raman microscopes that could be used for large area tissue imaging. One also might expect that in vivo imaging using SERS nanotags will be boosted in the future, toward cancer diagnostics.

As both iSERS assays and iSERS microscopy rely on the quality of SERS nanotags in order to achieve more reliable quantitative and reproducible results, more efforts should be made into the control of the monodispersity and homogenity of SERS nanotags. Furthermore, the stability and the storage of the SERS nanotags are further key issues that need to be considered in the future for their reliable application.

References

[1] Y. Wang, B. Yan, L. Chen, SERS tags: novel optical nanoprobes for bioanalysis, Chem. Rev. 113 (3) (2013) 1391−1428.
[2] Z. Wang, et al., SERS-activated platforms for immunoassay: probes, encoding methods, and applications, Chem. Rev. 117 (12) (2017) 7910−7963.
[3] L.A. Lane, X. Qian, S. Nie, SERS nanoparticles in medicine: from label-free detection to spectroscopic tagging, Chem. Rev. 115 (19) (2015) 10489−10529.
[4] W. Zhang, et al., SERS *nanotags and their applications in biosensing bioimaging*, Journal of Analysis and Testing 2 (1) (2018) 26−44.
[5] L. Fabris, SERS tags: the next promising tool for personalized cancer detection? ChemNanoMat 2 (4) (2016) 249−258.
[6] B. Shan, et al., Novel SERS labels: rational design, functional integration and biomedical applications, Coord. Chem. Rev. 371 (2018) 11−37.
[7] Y. Wang, S. Schlücker, Rational design and synthesis of SERS labels, Analyst 138 (8) (2013) 2224−2238.
[8] S. Schlücker, Surface-enhanced Raman spectroscopy: concepts and chemical applications, Angew. Chem. Int. Ed. 53 (19) (2014) 4756−4795.
[9] B. Küstner, et al., SERS labels for red laser excitation: silica-encapsulated SAMs on tunable gold/silver nanoshells, Angew. Chem. Int. Ed. 48 (11) (2009) 1950−1953.
[10] W. Xie, S. Schlücker, Rationally designed multifunctional plasmonic nanostructures for surface-enhanced Raman spectroscopy: a review, Rep. Prog. Phys. 77 (11) (2014) 116502.
[11] J. Reguera, et al., Anisotropic metal nanoparticles for surface enhanced Raman scattering, Chem. Soc. Rev. 46 (13) (2017) 3866−3885.
[12] H.K. Lee, et al., Designing surface-enhanced Raman scattering (SERS) platforms beyond hotspot engineering: emerging opportunities in analyte manipulations and hybrid materials, Chem. Soc. Rev. 48 (3) (2019) 731−756.

[13] M.C. Daniel, D. Astruc, Gold nanoparticles: assembly, supramolecular chemistry, quantum-size-related properties, and applications toward biology, catalysis, and nanotechnology, Chem. Rev. 104 (1) (2004) 293–346.
[14] D.D. Evanoff, G. Chumanov, Synthesis and optical properties of silver nanoparticles and arrays, ChemPhysChem 6 (7) (2005) 1221–1231.
[15] D. Steinigeweg, S. Schlücker, Monodispersity and size control in the synthesis of 20-100 nm quasi-spherical silver nanoparticles by citrate and ascorbic acid reduction in glycerol-water mixtures, Chem. Commun. 48 (69) (2012) 8682–8684.
[16] Q.B. Zhang, et al., Monodispersity control in the synthesis of monometallic and bimetallic quasi-spherical gold and silver nanoparticles, Nanoscale 2 (10) (2010) 1962–1975.
[17] Y. Sun, B.T. Mayers, Y. Xia, Template-engaged replacement reaction: a one-step approach to the large-scale synthesis of metal nanostructures with hollow interiors, Nano Lett. 2 (5) (2002) 481–485.
[18] M. Gellner, B. Küstner, S. Schlücker, Optical properties and SERS efficiency of tunable gold/silver nanoshells, Vib. Spectrosc. 50 (1) (2009) 43–47.
[19] Y. Zhang, et al., Gold and silver nanoparticle monomers are non-SERS-active: a negative experimental study with silica-encapsulated Raman-reporter-coated metal colloids, Phys. Chem. Chem. Phys. 17 (33) (2015) 21120–21126.
[20] K. Kneipp, et al., Single molecule detection using surface-enhanced Raman scattering (SERS), Phys. Rev. Lett. 78 (9) (1997) 1667–1670.
[21] S.M. Nie, S.R. Emery, Probing single molecules and single nanoparticles by surface-enhanced Raman scattering, Science 275 (5303) (1997) 1102–1106.
[22] C.J. Murphy, et al., Anisotropic metal nanoparticles: synthesis, assembly, and optical applications, J. Phys. Chem. B 109 (29) (2005) 13857–13870.
[23] Y.N. Xia, et al., Shape-controlled synthesis of metal nanocrystals: simple chemistry meets complex physics? Angew. Chem. Int. Ed. 48 (1) (2009) 60–103.
[24] B.J. Wiley, et al., Maneuvering the surface plasmon resonance of silver nanostructures through shape-controlled synthesis, J. Phys. Chem. B 110 (32) (2006) 15666–15675.
[25] X.Q. Zou, E.B. Ying, S.J. Dong, Seed-mediated synthesis of branched gold nanoparticles with the assistance of citrate and their surface-enhanced Raman scattering properties, Nanotechnology 17 (18) (2006) 4758–4764.
[26] K.L. Wustholz, et al., Structure–activity relationships in gold nanoparticle dimers and trimers for surface-enhanced Raman spectroscopy, J. Am. Chem. Soc. 132 (31) (2010) 10903–10910.
[27] J.H. Yoon, et al., Ideal dimers of gold nanospheres for precision plasmonics: *synthesis* and *characterization* at the single-particle level for Identification of higher order modes, Small 14 (4) (2018) 1702754.
[28] M. Gellner, et al., 3D Self-assembled plasmonic superstructures of gold nanospheres: synthesis and characterization at the single-particle level, Small 7 (24) (2011) 3445–3451.
[29] I. Choi, et al., Core–satellites assembly of silver nanoparticles on a single gold nanoparticle via metal ion-mediated complex, J. Am. Chem. Soc. 134 (29) (2012) 12083–12090.
[30] Z. Fan, et al., Assembly gold Nanoparticles on gold nanorods using functionalized poly(N-isopropylacrylamide) as polymeric "glue", Part. Part. Syst. Charact. 33 (9) (2016) 698–702.
[31] V. Tran, et al., Probing the SERS brightness of individual Au nanoparticles, hollow Au/Ag nanoshells, Au nanostars and Au core/Au satellite particles: single-particle experiments and computer simulations, Nanoscale 10 (46) (2018) 21721–21731.
[32] Y.C. Cao, R. Jin, C.A. Mirkin, Nanoparticles with Raman spectroscopic fingerprints for DNA and RNA detection, Science 297 (5586) (2002) 1536–1540.
[33] Y.C. Cao, et al., Raman dye-labeled nanoparticle probes for proteins, J. Am. Chem. Soc. 125 (48) (2003) 14676–14677.
[34] Y. Lai, et al., Raman-encoded microbeads for spectral multiplexing with SERS detection, RSC Adv. 5 (18) (2015) 13762–13767.
[35] D. Graham, K. Faulds, W.E. Smith, Biosensing using silver nanoparticles and surface enhanced resonance Raman scattering, Chem. Commun. 42 (2006) 4363–4371.

[36] D. Graham, Selective detection of deoxyribonucleic acid at ultralow concentrations by SERRS, Anal. Chem. 69 (22) (1997) 4703–4707.
[37] D. Graham, Surface-enhanced resonance Raman scattering as a novel method of DNA discrimination, Angew. Chem. 112 (6) (2000) 1103–1105.
[38] K. Faulds, Quantitative simultaneous multianalyte detection of DNA by dual-wavelength surface-enhanced resonance Raman scattering, Angew. Chem. Int. Ed. 46 (11) (2007) 1829–1831.
[39] S.J. Cho, et al., Combinatorial synthesis of a triphenylmethine library and their application in the development of Surface Enhanced Raman Scattering (SERS) probes, Chem. Commun. 46 (5) (2010) 722–724.
[40] A. Samanta, et al., Ultrasensitive near-infrared Raman reporters for SERS-based in vivo cancer detection, Angew. Chem. Int. Ed. Engl. 50 (27) (2011) 6089–6092.
[41] K.K. Maiti, Multiplex targeted in vivo cancer detection using sensitive near-infrared SERS nanotags, Nano Today 7 (2) (2012) 85–93.
[42] J. Ni, Immunoassay readout method using extrinsic Raman labels adsorbed on immunogold colloids, Anal. Chem. 71 (21) (1999) 4903–4908.
[43] D.S. Grubisha, Femtomolar detection of prostate-specific antigen: an immunoassay based on surface-enhanced Raman scattering and immunogold labels, Anal. Chem. 75 (21) (2003) 5936–5943.
[44] M.D. Porter, SERS as a bioassay platform: fundamentals, design, and applications, Chem. Soc. Rev. 37 (5) (2008) 1001–1011.
[45] S. Schücker, SERS microscopy: nanoparticle probes and biomedical applications, ChemPhysChem 10 (9-10) (2009) 1344–1354.
[46] C. Jehn, et al., Water soluble SERS labels comprising a SAM with dual spacers for controlled bioconjugation, Phys. Chem. Chem. Phys. 11 (34) (2009) 7499–7504.
[47] M. Schütz, Synthesis of glass-coated SERS nanoparticle probes via SAMs with terminal SiO precursors, Small (Weinh. Bergstr., Ger.) 6 (6) (2010) 733–737.
[48] S.P. Mulvaney, Glass-coated, analyte-tagged nanoparticles: a new tagging system based on detection with surface-enhanced Raman scattering, Langmuir. 19 (11) (2003) 4784–4790.
[49] S. Keren, Noninvasive molecular imaging of small living subjects using Raman spectroscopy, Proc. Natl Acad. Sci. (PNAS) 105 (15) (2008) 5844–5849.
[50] D.M. Zhang, Isotope edited internal standard method for quantitative surface-enhanced Raman spectroscopy, Anal. Chem. 77 (11) (2005) 3563–3569.
[51] E. Blackie, et al., Bi-analyte SERS with isotopically edited dyes, Phys. Chem. Chem. Phys. 10 (28) (2008) 4147–4153.
[52] P.A.G. Cormack, et al., Multidentate macromolecules for functionalisation, passivation and labelling of metal nanoparticles, Chem. Commun. 22 (2008) 2517–2519.
[53] F. McKenzie, et al., SERRS coded nanoparticles for biomolecular labelling with wavelength-tunable discrimination, Analyst 134 (3) (2009) 549–556.
[54] X. Su, Composite organic–inorganic nanoparticles (COINs) with chemically encoded optical signatures, Nano Lett. 5 (1) (2005) 49–54.
[55] M. Yang, Development of polymer-encapsulated metal nanoparticles as surface-enhanced raman scattering probes, Small 5 (2) (2009) 198–202.
[56] A.F. McCabe, SERRS labelled beads for multiplex detection, Faraday Discuss. 132 (0) (2006) 303–308.
[57] L.C. Martin, Turning up the lights—fabrication of brighter SERRS nanotags, Chem. Commun. 46 (29) (2010) 5247–5249.
[58] A. McLintock, Controlled side-by-side assembly of gold nanorods and dye molecules into polymer-wrapped SERRS-active clusters, Chem. Commun. 47 (13) (2011) 3757–3759.
[59] G. Chen, Measuring ensemble-averaged surface-enhanced Raman scattering in the hotspots of colloidal nanoparticle dimers and trimers, J. Am. Chem. Soc. 132 (11) (2010) 3644–3645.
[60] W.E.E. Doering, SERS as a foundation for nanoscale, optically detected biological labels, Adv. Mater. 19 (20) (2007) 3100–3108.
[61] W.E. Doering, Spectroscopic tags using dye-embedded nanoparticles and surface-enhanced Raman scattering, Anal. Chem. 75 (22) (2003) 6171–6176.

[62] X. Liu, Synthesis of core−shell surface-enhanced Raman tags for bioimaging, Anal. Chem. 82 (1) (2009) 441−446.
[63] J.R. Anema, et al., Shell-isolated nanoparticle-enhanced Raman spectroscopy: expanding the versatility of surface-enhanced Raman scattering., in: Annu. Rev. Anal. Chem., 4, 2011, pp. 129−150.
[64] X. Qian, In vivo tumor targeting and spectroscopic detection with surface-enhanced Raman nanoparticle tags, Nat. Biotechnol. 26 (1) (2008) 83−90.
[65] C.M. Shachaf, PLoS One 4 (4) (2009) e5206.
[66] W. Stöber, Controlled growth of monodisperse silica spheres in the micron size range, J. Colloid Interface Sci. 26 (1) (1968) 62−69.
[67] C.G. Wang, Monodispersed gold nanorod-embedded silica particles as novel Raman labels for biosensing, Adv. Funct. Mater. 18 (2) (2008) 355−361.
[68] M. Li, Detection of adenosine triphosphate with an aptamer biosensor based on surface-enhanced Raman scattering, Anal. Chem. 84 (6) (2012) 2837−2842.
[69] J.-H. Kim, Nanoparticle probes with surface enhanced Raman spectroscopic tags for cellular cancer targeting, Anal. Chem. 78 (19) (2006) 6967−6973.
[70] D. Steinigeweg, Gold nanoparticles: fast and cost-effective purification of gold nanoparticles in the 20-250 nm size range by continuous density gradient centrifugation, Small 7 (17) (2011) 2406−2448.
[71] G.F. Wang, Mixed monolayers on gold nanoparticle labels for multiplexed surface-enhanced Raman scattering based immunoassays, Anal. Chem. 81 (23) (2009) 9643−9650.
[72] W.P. Page Faulk (Ed.), Communication to the editors, Immunochemistry 8 (11) (1965) 1081−1083.
[73] H. Robenek, Colloidal gold: principles, methods, and applications vols. I and II, (vol. III in preparation) Edited by M. A. Hayat, Academic Press, Inc., New York.
[74] M. Schütz, Hydrophilically stabilized gold nanostars as SERS labels for tissue imaging of the tumor suppressor p63 by immuno-SERS microscopy, Chem. Commun. 47 (14) (2011).
[75] S. Schlücker, Immuno-surface-enhanced coherent anti-stokes Raman scattering microscopy: immunohistochemistry with target-specific metallic nanoprobes and nonlinear Raman microscopy, Anal. Chem. 83 (18) (2011) 7081−77085.
[76] G.T. Hermanson, Chapter 2—Functional targets for bioconjugation, in: G.T. Hermanson (Ed.), Bioconjugate Techniques, *third ed.*, Academic Press, Boston, 2013, pp. 127−228.
[77] G.T. Hermanson, Chapter 3—The Reactions of Bioconjugation, in: G.T. Hermanson (Ed.), Bioconjugate Techniques, third ed., Academic Press, Boston, 2013, pp. 229−258.
[78] J.D. Driskell, et al., Low-level detection of viral pathogens by a surface-enhanced Raman scattering based immunoassay, Anal. Chem. 77 (19) (2005) 6147−6154.
[79] J.D. Driskell, et al., Surface-enhanced Raman scattering immunoassays using a rotated capture substrate, Anal. Chem. 79 (11) (2007) 4141−4148.
[80] G. Wang, et al., Detection of the potential pancreatic cancer marker MUC4 in serum using surface-enhanced Raman scattering, Anal. Chem. 83 (7) (2011) 2554−2561.
[81] Y. Wang, et al., Duplex microfluidic SERS detection of pathogen antigens with nanoyeast single-chain variable fragments, Anal. Chem. 86 (19) (2014) 9930−9938.
[82] Y. Wang, et al., Femtogram detection of cytokines in a direct dot-blot assay using SERS microspectroscopy and hydrophilically stabilized Au−Ag nanoshells, Chem. Commun. 50 (21) (2014) 2711−2714.
[83] M.A. Penn, D.M. Drake, J.D. Driskell, Accelerated surface-enhanced Raman spectroscopy (SERS)-based immunoassay on a gold-plated membrane, Anal. Chem. 85 (18) (2013) 8609−8617.
[84] Y. Wang, et al., Enabling rapid and specific surface-enhanced Raman scattering immunoassay using nanoscaled surface shear forces, ACS Nano 9 (6) (2015) 6354−6362.
[85] K. Kamil Reza, et al., Electrohydrodynamic-induced SERS immunoassay for extensive multiplexed biomarker sensing, Small 13 (9) (2017) 1602902.
[86] M. Sánchez-Purrà, et al., Design of SERS nanotags for multiplexed lateral flow immunoassays, Mol. Syst. Des. Eng. 2 (4) (2017) 401−409.
[87] J. Hwang, S. Lee, J. Choo, Application of a SERS-based lateral flow immunoassay strip for the rapid and sensitive detection of staphylococcal enterotoxin B, Nanoscale 8 (22) (2016) 11418−11425.

[88] X. Fu, et al., A SERS-based lateral flow assay biosensor for highly sensitive detection of HIV-1 DNA, Biosens. Bioelectron. 78 (2016) 530–537.
[89] L. Blanco-Covián, et al., Au@Ag SERRS tags coupled to a lateral flow immunoassay for the sensitive detection of pneumolysin, Nanoscale 9 (5) (2017) 2051–2058.
[90] X. Gao, et al., Paper-based surface-enhanced Raman Scattering lateral flow strip for detection of neuron-specific enolase in blood plasma, Anal. Chem. 89 (18) (2017) 10104–10110.
[91] B.N. Khlebtsov, et al., SERS-based lateral flow. Immunoass. troponin I using. gap-enhanced Raman tags, Nano Res. 12 (2) (2019) 413–420.
[92] V. Tran, et al., Rapid, quantitative, ultrasensitive point-of-care testing: a portable SERS reader for lateral flow assays in clinical chemistry, Angew. Chem. Int. Ed. 58 (2) (2019) 442–446.
[93] O. Stoss, T. Henkel, Biomedical marker molecules for cancer—current status and perspectives, Drug. Discov. Today 3 (6) (2004) 228–237.
[94] Z. Wang, ErbB receptors and cancer, Methods Mol. Biol. 1652 (2017) 3–35.
[95] L.E. Kelemen, The role of folate receptor alpha in cancer development, progression and treatment: cause, consequence or innocent bystander? Int. J. Cancer 119 (2) (2006) 243–250.
[96] M.M. Fuster, J.D. Esko, The sweet and sour of cancer: glycans as novel therapeutic targets, Nat. Rev. Cancer 5 (7) (2005) 526–542.
[97] Z. Shen, A. Wu, X. Chen, Current detection technologies for circulating tumor cells, Chem. Soc. Rev. 46 (8) (2017) 2038–2056.
[98] J.V. Jokerst, et al., Affibody-functionalized gold-silica nanoparticles for Raman molecular imaging of the epidermal growth factor receptor, Small 7 (5) (2011) 625–633.
[99] Y. Wang, et al., Raman-encoded molecular imaging (REMI) with topically applied SERS nanoparticles for intraoperative guidance of lumpectomy, Cancer Res. 77 (16) (2017) 4506–4516.
[100] S. Lee, et al., Biological imaging of HEK293 cells expressing PLCgamma1 using surface-enhanced Raman microscopy, Anal. Chem. 79 (3) (2007) 916–922.
[101] S. Lee, et al., Surface-enhanced Raman scattering imaging of HER2 cancer markers overexpressed in single MCF7 cells using antibody conjugated hollow gold nanospheres, Biosens. Bioelectron. 24 (7) (2009) 2260–2263.
[102] H. Park, et al., SERS imaging of HER2-overexpressed MCF7 cells using antibody-conjugated gold nanorods, Phys. Chem. Chem. Phys 11 (34) (2009) 7444–7449.
[103] J. Yue, et al., Bull serum albumin coated Au@Agnanorods as SERS probes for ultrasensitive osteosarcoma cell detection, Talanta 150 (2016) 503–509.
[104] M. Chen, et al., High-sensitive bioorthogonal SERS tag for live cancer cell imaging by self-assembling core-satellites structure gold-silver nanocomposite, Talanta 172 (2017) 176–181.
[105] G. Zhu, X. Chen, Aptamer-based targeted therapy, Adv. Drug. Deliv. Rev. (2018).
[106] J. Wu, et al., Bioorthogonal SERS nanoprobes for mulitplex spectroscopic detection, tumor cell targeting, and tissue imaging, Chemistry 21 (37) (2015) 12914–12918.
[107] H. Di, et al., High-precision profiling of sialic acid expression in cancer cells and tissues using background-free surface-enhanced Raman scattering tags, Anal. Chem. 89 (11) (2017) 5874–5881.
[108] C. Fasolato, et al., Folate-based single cell screening using surface enhanced Raman microimaging, Nanoscale 8 (39) (2016) 17304–17313.
[109] Y. Yin, et al., Prussian Blue as a highly sensitive and background-free resonant Raman reporter, Anal. Chem. 89 (3) (2017) 1551–1557.
[110] D.C. Kennedy, et al., Development of nanoparticle probes for multiplex SERS imaging of cell surface proteins, Nanoscale 2 (8) (2010) 1413–1416.
[111] S. Lee, et al., Rapid and sensitive phenotypic marker detection on breast cancer cells using surface-enhanced Raman scattering (SERS) imaging, Biosens. Bioelectron. 51 (2014) 238–243.
[112] Y. Chen, et al., Alkyne-modulated surface-enhanced Raman scattering-palette for optical interference-free and multiplex cellular imaging, Anal. Chem. 88 (12) (2016) 6115–6119.
[113] V. Plaks, C.D. Koopman, Z. Werb, Circulating tumor cells, Science 341 (6151) (2013) 1186–1188.

[114] C. Alix-Panabieres, K. Pantel, Challenges in circulating tumour cell research, Nat. Rev. Cancer 14 (9) (2014) 623–631.
[115] Z.A. Nima, et al., Circulating tumor cell identification by functionalized silver-gold nanorods with multicolor, super-enhanced SERS and photothermal resonances, Sci. Rep. 4 (2014) 4752.
[116] S.C. Tsao, et al., Characterising the phenotypic evolution of circulating tumour cells during treatment, Nat. Commun. 9 (1) (2018) 1482.
[117] N.V. Jordan, et al., HER2 expression identifies dynamic functional states within circulating breast cancer cells, Nature 537 (7618) (2016) 102–106.
[118] E.C. Stack, et al., Multiplexed immunohistochemistry, imaging, and quantitation: a review, with an assessment of tyramide signal amplification, multispectral imaging and multiplex analysis, Methods 70 (1) (2014) 46–58.
[119] S. Schlücker, et al., Immuno-Raman microspectroscopy:In situ detection of antigens in tissue specimens by surface-enhanced Raman scattering, J. Raman Spectrosc. 37 (7) (2006) 719–721.
[120] L. Sun, K.-B. Sung, C. Dentinger, B. Lutz, L. Nguyen, J. Zhang, et al., Composite organic–inorganic nanoparticles as Raman labels for tissue analysis, Nano Lett. 7 (2) (2007) 351–356.
[121] Y. Chen, et al., Immunoassay for LMP1 in nasopharyngeal tissue based on surface-enhanced Raman scattering, Int. J. Nanomed. 7 (2012) 73–82.
[122] M. Schütz, et al., Design and synthesis of Raman reporter molecules for tissue imaging by immuno-SERS microscopy, J. Biophotonics 4 (6) (2011) 453–463.
[123] B.R. Lutz, et al., Spectral analysis of multiplex Raman probe signatures, ACS Nano 2 (11) (2008) 2306–2314.
[124] B. Lutz, et al., Raman nanoparticle probes for antibody-based protein detection in tissues, J. Histochem. Cytochem. 56 (4) (2008) 371–379.
[125] M. Salehi, et al., Rapid immuno-SERS microscopy for tissue imaging with single-nanoparticle sensitivity, J. Biophotonics 6 (10) (2013) 785–792.
[126] X.P. Wang, et al., iSERS microscopy guided by wide field immunofluorescence: analysis of HER2 expression on normal and breast cancer FFPE tissue sections, Analyst 141 (17) (2016) 5113–5119.
[127] Y. Zhang, et al., Effect of antigen retrieval methods on nonspecific binding of antibody-metal nanoparticle conjugates on formalin-fixed paraffin-embedded tissue, Anal. Chem. 90 (1) (2018) 760–768.
[128] M. Salehi, et al., Two-color SERS microscopy for protein co-localization in prostate tissue with primary antibody-protein A/G-gold nanocluster conjugates, Nanoscale 6 (4) (2014) 2361–2367.
[129] S. Signoretti, et al., p63 Is a prostate basal cell marker and is required for prostate development, Am. J. Pathol. 157 (6) (2000) 1769–1775.
[130] J. Li, et al., A universal strategy for one-pot synthesis of SERS tags, Nanoscale 10 (17) (2018) 8292–8297.
[131] Y.W. Wang, et al., Rapid ratiometric biomarker detection with topically applied SERS nanoparticles, Technology 2 (2) (2014) 118–132.
[132] L. Sinha, et al., Quantification of the binding potential of cell-surface receptors in fresh excised specimens via dual-probe modeling of SERS nanoparticles, Sci. Rep. 5 (2015) 8582.
[133] Y.W. Wang, et al., Multiplexed Molecular Imaging of Fresh Tissue Surfaces Enabled by Convection-Enhanced Topical Staining with SERS-Coded Nanoparticles, Small 12 (40) (2016) 5612–5621.
[134] S. Kang, et al., Microscopic investigation of topically applied nanoparticles for molecular imaging of fresh tissue surfaces, J. Biophotonics 11 (4) (2018) e201700246.
[135] Y. Wang, et al., Quantitative molecular phenotyping with topically applied SERS nanoparticles for intraoperative guidance of breast cancer lumpectomy, Sci. Rep. 6 (2016) 21242.

CHAPTER 18

Highly localized characterization of protein structure and interaction by surface-enhanced Raman scattering

Janina Kneipp

Department of Chemistry, Humboldt-Universität zu Berlin, Berlin, Germany

18.1 Introduction

Surface-enhanced Raman scattering (SERS) [1–3] is caused by an electromagnetic enhancement of the excitation light and the Raman scattered light in the local optical fields of plasmonic nanostructures [4–6]. The electromagnetic enhancement has the greatest contribution to the total enhancement factor [4–6]. A direct electronic interaction of a molecule with the nanostructure, causing the chemical enhancement, can also contribute to SERS [7,8]. While the latter usually implies a chemical selectivity, and thereby a localization of the probing to those parts of a molecular species that directly adsorb to the nanostructure surface, the strong confinement of the local fields of the nanostructures results in the localization of the SERS signals in a more general sense, in principle independent of the chemical properties of the molecule or of specific functional groups. Provided that they can experience the local optical field, SERS enhancement can be achieved for all parts of a molecule. The intensity of a SERS signal is even more localized and restricted to the immediate environment of a plasmonic nanostructure than the field itself. While the field of a dipole that is induced upon excitation of the localized surface plasmon of a metal nanoparticle decays proportionally to $\sim 1/r^3$, r being the distance between the molecule and the plasmonic nanostructure, the product of the intensities of both the excitation and the Raman field must be considered, each decaying proportionally to $\sim 1/r^6$. Therefore the SERS signal decreases with $\sim 1/r^{12}$. In the probing of large molecules, such as proteins, this high localization of the intense SERS signals to a few nanometers has major implications and provides new possibilities in protein characterization.

In a situation where proteins and SERS substrates can interact freely, the properties of different groups and side chains of a protein molecule, as well as the properties of other

Vibrational Spectroscopy in Protein Research.
DOI: https://doi.org/10.1016/B978-0-12-818610-7.00018-9

molecular species at the surface of the plasmonic nanostructure, will determine the interaction of both, and thereby the spectral information obtained from the protein. This can lead as far as to the protein determining the plasmonic properties of the nanostructures and the distribution and extension of the local field, mainly by the formation of nanostructure aggregates with more [9,10] or less control [11,12]. Therefore, rather than being able to characterize the structure and composition of a molecule as large as a protein in a comprehensive fashion (at least in an individual spectrum, which can only be achieved in some specific configurations), SERS is ideal to study the structure of the protein molecule in its interaction with the nanostructure, which is expected not to be identical to its structure in the absence of the SERS substrate. The idea of single (protein) molecule detection becomes purely technical, considering that on the one hand protein molecules can be detected based on SERS signals of any single side chain, or of a functional or prosthetic group, as was pioneered two decades ago [13], and on the other hand need to have their spectroscopic fingerprint taken in ensemble experiments in order to have their identity and interactions properly determined.

As will be discussed here, SERS being restricted to the immediate proximity of the metal nanostructures, the site of interaction of structurally very similar proteins with gold nanostructures can be characterized in great detail. It will also be shown that the concept of SERS can be extended to two-photon excitation, which leads to complementary vibrational information and additional insight into the interaction of protein side chains with the nanostructures. An example in contrast to this, parts of protein molecules can be decidedly positioned in the hot spots of plasmonic nanoaggregates that are specifically designed to provide a constant and reproducible high enhancement, and thereby specific groups in single protein molecules can be observed. Here, application of so-called nanolenses will be introduced in the context of protein detection, and possibilities and limitations of such an application will be discussed.

18.2 Local probing gives detailed information on the interaction of proteins

18.2.1 The protein corona of plasmonic nanoparticles inside living cells

When a nanoparticle is exposed to a biological environment, biomolecules attach to its surface and form a layer around it, termed the corona. The biomolecular corona changes the original surface properties of the particle, and mainly determines how the nanomaterial interacts with its bioorganic environment. As an example, in animal cells, the properties of the biomolecule corona affect solvation, recognition processes, and nanoparticle stability [14–18]. The formation of a corona around the nanoparticles was originally shown for a number of nanomaterials, including silica, polystyrene, and gold, after their purification and

removal of the pristine nanostructure using mass spectrometry [16,19–22]. It was difficult to obtain data on the nanoparticle surface composition in vivo, until in the past years SERS was shown to be capable of monitoring the biomolecules around gold and silver nanoparticles in cells, and thereby to obtain information on their corona while in the living system [23–25]. In agreement with the mass spectrometry data, the SERS spectra show many contributions from protein molecules [23,26,27]. The interaction of gold and silver nanoparticles with molecules composing the corona has been monitored by SERS in living cells over time, starting from particle incubation in the biological medium followed by endocytotic uptake and endosomal processing [24,28] and compared for different metal nanostructures, as well as for composite materials of plasmonic nanostructures and other components, including silica or iron oxide, which influence the composition and/or structure of the adsorbing molecules [29,30]. The differences of the SERS signatures from gold and silver nanoparticles that were delivered into animal cells by endocytosis from cell culture medium are specifically striking [27].

While the composition of the corona of gold nanoparticles varies a lot over the time course of endosomal maturation, the spectra measured with silver nanoparticles that resided in the same cell line remained constant for many hours. Fig. 18.1 compares the spectra obtained from intracellular particles with those measured when silver nanoparticles were dispersed in Dulbecco's modified eagle medium (DMEM), fetal calf serum (FCS), and L-cysteine solution (Fig. 18.1, top-to-bottom). All of them are components of the cell culture medium to which the silver nanoparticles are added for delivery. The FCS spectrum is similar to the intracellular spectra with intense bands that are attributed to cysteine [31–33]. The carboxylate deformation vibration at 722 cm^{-1} (Fig. 18.1, second spectrum) is very sharp and pronounced. It can be assigned to glutamic or aspartic acid as constituents of the proteins in FCS. The DMEM spectral signature is dominated by Raman signals of amino acids. L-Cysteine, contained in DMEM at a concentration of 48 mg/L, yields a SERS spectrum that has many similarities with the cell spectra (Fig. 18.1, compare first and last spectrum). Apart from occurring as a free amino acid in DMEM, cysteine is also part of proteins contained in FCS. In particular, bovine serum albumin (BSA) as a cysteine-rich protein that accounts for up to 45% of the total protein in FCS could be a major constituent of the corona. Considering the thiophilic character of silver, the strong adsorption of cysteine to the silver nanoparticles is not surprising. The formation of Ag-S–protein bonds can proceed either by reaction with the thiol group or also by cleavage of disulfide bonds of the protein [32]. Meanwhile, several studies that investigate the interaction of silver nanoparticles with BSA [34,35] and human serum albumin (HSA) [36] in artificial solutions have been reported. In our in vivo data here, the spectra from the living cells indicate a persistence of the protein coat of silver nanoparticles inside the cells.

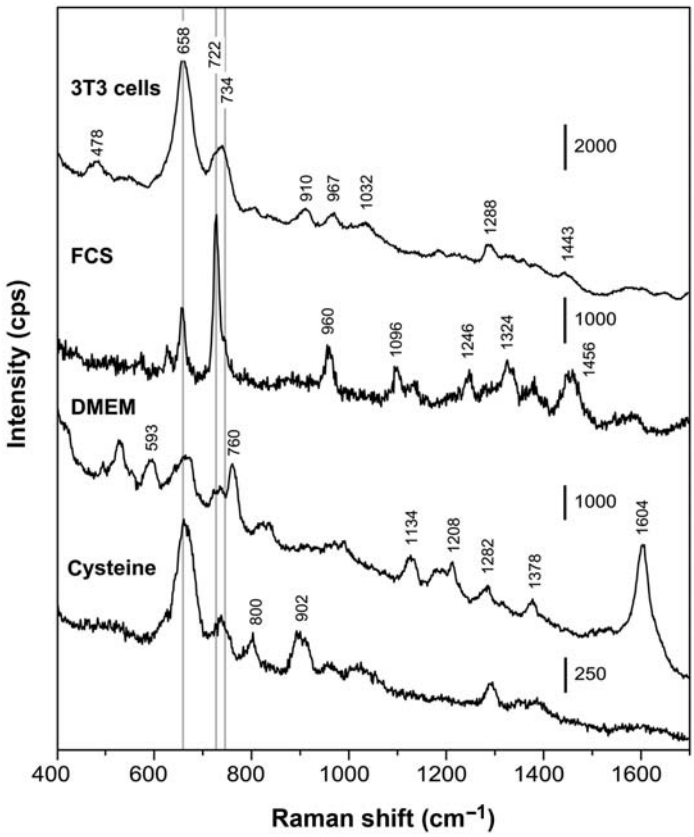


Figure 18.1
SERS spectra of fibroblast cells in comparison with different cell culture relevant media obtained in the presence of silver nanoparticles (excitation wavelength 785 nm; acquisition time 1 s; laser intensity 8×10^4 W/cm). Source: *Reproduced from D. Drescher, P. Guttmann, T. Büchner, S. Werner, G. Laube, A. Hornemann, et al., Specific biomolecule corona is associated with ring-shaped organization of silver nanoparticles in cells, Nanoscale 5 (19) (2013) 9193–9198, with permission from The Royal Society of Chemistry.*

18.2.2 Binding sites of serum albumins in the protein corona of nanoparticles

Serum albumins are not only contained in cell culture media, but are also encountered by nanostructures in other biofluids. A deeper understanding of their adsorption behavior in the presence of nanomaterials is important for diagnostics and drug delivery. Here, we discuss the interaction of BSA and HSA as most abundant serum proteins, with citrate-stabilized gold nanoparticles in the size range of $\sim$30 nm, based on their SERS spectra. The spectral signatures of the citrate ligands of the gold nanoparticles do not show strong interference with the SERS signals from the molecules of interest [37,38]. Even though the high sequence overlap of the two molecules suggests their very similar interaction with the

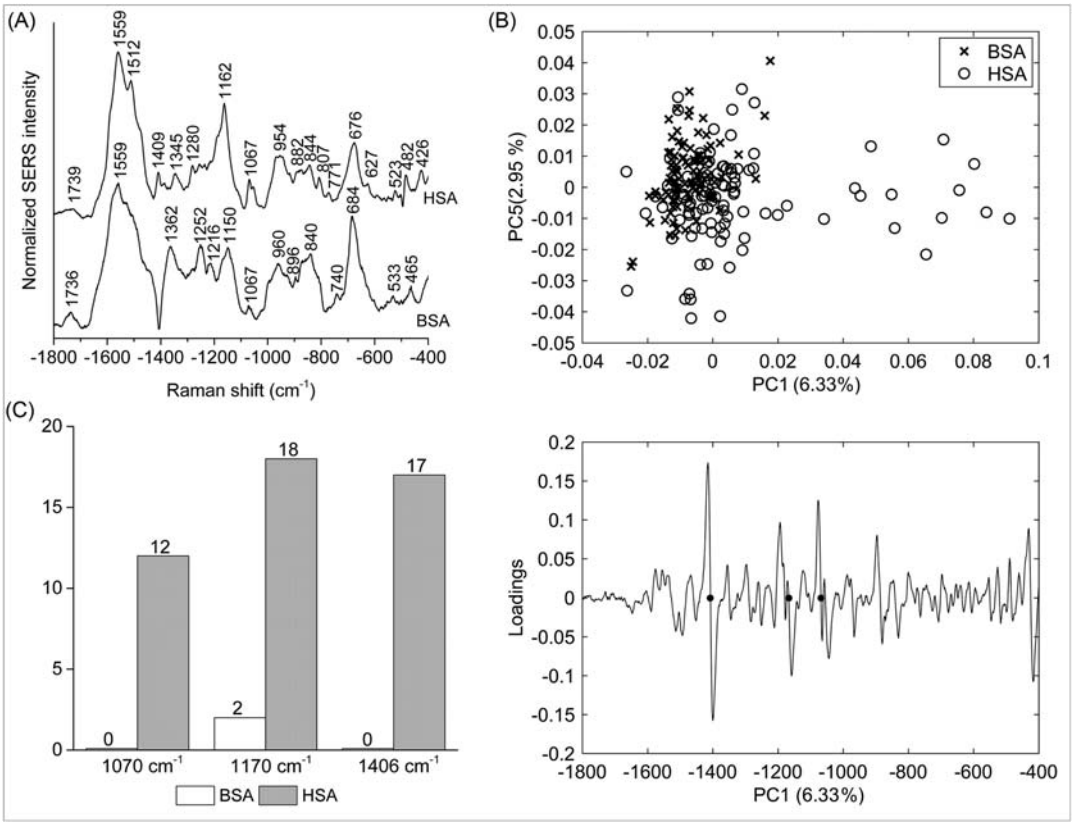


Figure 18.2

(A) Averages of ~100 BSA and ~100 HSA SERS spectra obtained with citrate-stabilized gold nanoparticles and excited at a wavelength of 785 nm and an excitation intensity of 5.7×10^5 W/cm^2. (B) Results of a principal component analysis using a data set of BSA and of HSA, the scores plot shows the first and second principal component obtained in a PCA using the full spectral range of vector-normalized first derivatives the SERS spectra. The loading of the first PC, with dots marking the bands at 1406 and 1170 cm^{-1} of the symmetric COO^- stretching and C—N stretching vibrations [38,39], and at 1070 cm^{-1} of the C-O and/or C-C, NH_2, and C-N stretching vibrations [40,41]. (C) Frequencies of occurrence of the vibrational bands at 1070 cm^{-1} of the C—O and/or C—C, NH_2, and C—N stretching modes [38,40], and at 1170 and 1406 cm^{-1} of the C—N and the symmetric COO^- stretching vibrations [38,39] in the BSA and HSA dataset, respectively. Source: *Reproduced from G.P. Szekeres, J. Kneipp, Different binding sites of serum albumins in the protein corona of gold nanoparticles, Analyst 143 (24) (2018) 6061—6068, with permission from The Royal Society of Chemistry.*

citrate-stabilized gold nanoparticles, very distinct SERS spectra were found (Fig. 18.2A) [42]. Interestingly, the band of lysine (-NH_3^+ deformation vibration at 1523 cm^{-1} [39]) is not pronounced in the normal Raman spectrum of BSA as reported, for example, in Refs. [43,44]. Here, its strong contribution to the SERS spectrum of BSA (Fig. 18.2A) in spite of

its small Raman cross section suggests that the lysine residues must be very close to the gold surface. Such a lysine−citrate interaction has been discussed in previous work using other approaches [21,45−48], and the data in Fig. 18.2A suggest that binding takes place by such an electrostatic interaction [21,45−48]. A strong amide III band at 1253 cm^{-1} and the band at 838 cm^{-1} indicate that the peptide backbone must be very close to the nanoparticle surface as well. Please note that the ring vibration of phenylalanine and tryptophan at 1004 cm^{-1} [43], usually very pronounced, does not appear in the spectrum, excluding the proximity of aromatic side chains to the nanoparticle surface (Fig. 18.2A, top). Since the nonpolar amino acid side chains are located inside the folded structure [49,50], hidden from the hydrophilic environment, the absence of the 1004 cm^{-1} band suggests that the probed BSA molecules preserve their secondary structure in so far as these residues do not become exposed.

Also the SERS spectrum of HSA (Fig. 18.2A, bottom) displays selective enhancement of bands that are not particularly prominent in the normal Raman spectrum of the molecule [51,52], underlining again the selectivity in the SERS experiment that comes from the interaction with the gold nanostructure. The enhancement of the signals from the $-NH_2$ groups (1512 cm^{-1}, 1058 cm^{-1}, 1070 cm^{-1} [38]) and of the protein backbone indicate proximity of both basic amino acid residues and of the peptide backbone to the gold surface. Also here, the 1004 cm^{-1} band of phenylalanine and tryptophan [43] is absent, suggesting that the native structure of the adsorbed HSA molecules is at least partially intact. There are several bands in both the BSA and HSA SERS spectra assigned to aromatic and aliphatic vibrations, indicating their proximity to the nanoparticle surface, which could also point toward a hydrophobic interaction with the nanoparticles.

The SERS spectra were obtained in small volumes of gold nanoparticles and BSA in solution ($\sim$10 fL focal volume created in droplets of 50 μL), and apart from discussing the average spectra displayed in Fig. 18.2A, the variation of the individual spectra in the different data sets must be assessed. Fig. 18.2B contains the results of a principal component analysis (PCA) conducted with a data set of 100 BSA and 100 HSA spectra, utilizing the whole spectral range. The loading of PC1 displayed in Fig. 18.3B suggests that the variance in the band at 1406 cm^{-1}, which can be assigned to the COO^- symmetric stretching [38,40], has the strongest influence on this separation. Furthermore, variances are observed for the band at 1070 cm^{-1}, assigned to stretching vibrations of C-O and/or C-C, NH_2, and C-N groups, and the C-N stretching mode at 1170 cm^{-1} [38,39]. The frequency of occurrence of these bands in the BSA and HSA spectra is presented in Fig. 18.2C. Only two out of the 100 BSA spectra contain the band at 1170 cm^{-1}. In contrast, the 1070, 1170, and 1406 cm^{-1} bands are mostly present in the HSA spectra.

The spectra provide information on the presence of molecular groups in the immediate vicinity of the nanoparticle surface. But the signals appear regardless of the type of

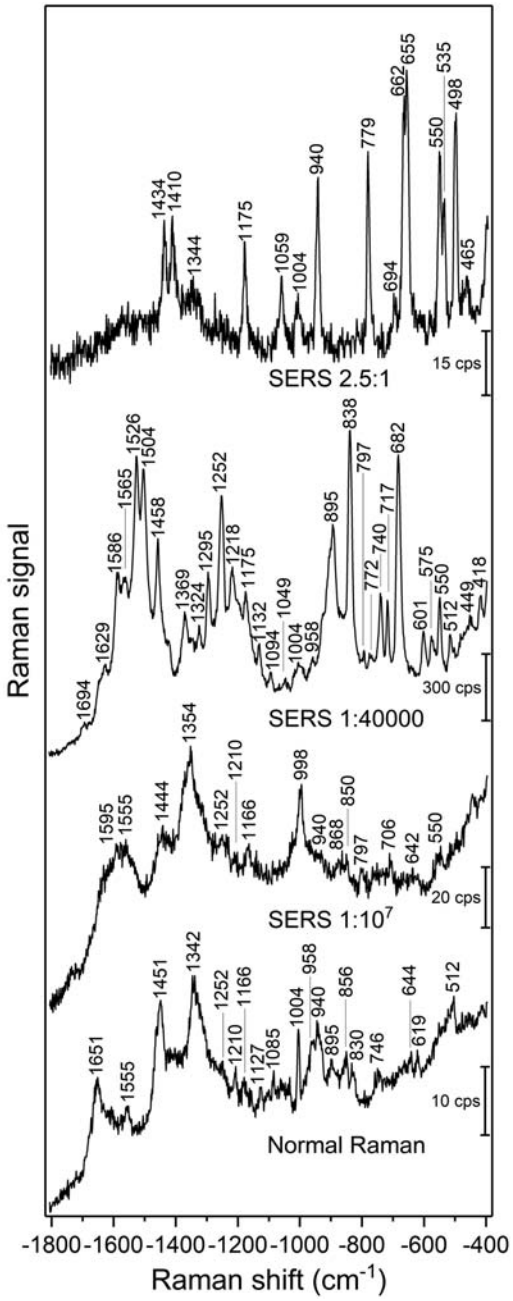


Figure 18.3

SERS spectra of BSA with gold nanoparticles with different gold nanoparticle:BSA molar ratios as indicated in the graph. Excitation wavelength: 785 nm, excitation intensity 5.7×10^5 W/cm^2, acquisition time per spectrum: 5 s for normal Raman, 1 s for SERS spectra. Source*: Reprinted from G.P. Szekeres, J. Kneipp, SERS probing of proteins in gold nanoparticle agglomerates, Front. Chem. 7 (2019) 30. Copyright (2019) Szekeres and Kneipp.*

interaction, and the protein molecules could interact with the citrate ions on the nanoparticles in electrostatic binding [21,47,48], or adsorb directly on the gold surface and replace the citrate ligands [48,53]. Information on the primary protein structure can support the interpretation of the SERS data in this regard. The electrostatic binding of the proteins to the citrate layer without ligand exchange is possible both for BSA and HSA via lysine residues as explained above. In contrast, the direct adsorption of protein molecules on the positively charged surface of gold nanoparticles requires negatively charged residues, namely the deprotonated COO^- groups that are found in the spectra of HSA, as indicated by the band at 1406 cm^{-1}, which can be assigned to the COO^- symmetric stretching [38,40]. The aspartic acid and glutamic acid residues in BSA and HSA (99 and 98 in total in each protein, respectively) are distributed homogeneously along the protein chains. The pH of the local microenvironment can be very different from the global pH, and the protonation state of these amino acid side chains can be altered by their neighboring residues, facilitating deprotonation. In this regard, especially those neighbors that possess basic characteristics can exert an influence. An analysis of all neighboring groups of aspartic acid and glutamic acid residue in both BSA and HSA, and the distribution of the basic residues lysine and arginine, as well as of histidine indicated that BSA contains six more amino acids that could facilitate the deprotonation of their neighboring aspartic acid or glutamic acid than HSA.

However, in the interaction with the nanoparticles, also steric restrictions, implied by the protein secondary structure, play an important role. Specifically, more flexible random coil structures would be energetically favored in the interaction over very much defined α-helices, since flexible segments are distorted more easily, allowing for the most preferred orientation of side chains involved in specific residue–nanoparticle interactions [54]. In BSA, no acidic side chain was found in a completely random-coiled structure, without having neighbors with basic residues that could be repelled by the net positive surface charge of the pristine nanoparticle surface [49,50]. Very different from BSA, a portion of the HSA spectra clearly indicates direct binding of this protein to the nanoparticle surface via the $\cdots$-Val_{293}-Glu_{294}-Asn_{295}-Asp_{296}-Glu_{297}-Met_{298}-$\cdots$ and $\cdots$-Leu_{491}-Glu_{492}-Val_{493}-$\cdots$ segments by the cooccurrence of bands at 1070, 1170, and 1406 cm^{-1} (Fig. 18.2C). These two segments were also identified as flexible segments by modeling of protein structure flexibility, calculating the root-mean-square fluctuations of the residues for the entire HSA chain.

18.2.3 Concentration dependence: implications in live cells

When the BSA concentration is varied in the solutions that contain the gold nanoparticles, a strong variation in enhancement and spectral information is found. Both are the result of a varying interaction of the same molecule when the protein concentration is changed. Fig. 18.3 shows spectra from a data set where BSA concentration was varied by seven orders of magnitude. In spite of the strong influence of the BSA analyte molecule on the

optical properties of the gold nanoparticle aggregates, rendering protein quantification impossible, a comparison of overall SERS signals is interesting regarding the ability to probe protein composition and interaction in cells, where such aggregates are formed from individual gold nanoparticles when they meet with the real biological environment. As indicated by the SERS spectra, the BSA molecules undergo a change in structure, which depends on their concentration. Comparing the normal Raman signals (Fig. 18.3, bottom trace) of a 7.5 mM BSA solution with the SERS spectrum obtained at a BSA concentration of $\sim$150 pM (Fig. 18.1, top trace), we can estimate an enhancement to be on the order of $\sim 10^8$. The top spectrum contains signals from cysteine and methionine side chains that are usually not exposed and indicate unfolding of BSA, while at a higher BSA: nanoparticle ratio amide bands are present, which suggests a less unfolded structure. The highest BSA concentration resulted in the lowest signal intensities (Fig. 18.3, second from below) rather than in a saturated overall intensity as would be expected due to a limited available surface of the SERS substrate [55]. This observation is in agreement with recent discussions [56] that nanoparticle aggregation hardly occurs when the protein concentration is much higher than the concentration needed to form a monolayer on each particle. When protein concentrations in the solutions approximate cellular protein concentrations, the high viscosity and the multilayer protein adsorption on the surface of the nanoparticles hinders the formation of nanoaggregates that can provide high SERS enhancement. Estimating the amount of nanoparticles taken up by a typical cell [57], and the number of typical-size gold nanoaggregates [27] that would fit in a focal volume, the number of particles in the model nanoparticle: BSA ratio of $1{:}10^7$ is very similar. Since the most concentrated sample approximates the cellular protein concentration, the observed high SERS signals from live cells [23,58] must be related to the intracellular inhomogeneity of protein concentrations and to active transport processes, which facilitate the formation and positioning of aggregates in cellular compartments.

18.3 Interaction of protein side chains from surface-enhanced hyper Raman scattering

18.3.1 Surface-enhanced hyper Raman scattering provides complementary information

Two-photon excited, spontaneous Raman scattering that is obtained in the local fields of plasmonic nanostructures, termed surface-enhanced hyper Raman scattering (SEHRS), has been discussed since the early days of SERS. In SEHRS, the scattered photons are shifted relative to the second harmonic of the excitation wavelength, and the excitation with light in the near infrared is combined with the detection of the hyper Raman light in the visible spectral range. Although the cross sections of nonenhanced hyper Raman scattering are

extremely low [59], the hyper Raman signal depends on the square of the incident radiation intensity. As a two-photon process, hyper Raman scattering benefits enormously from the enhancement by the local fields. The total enhancement factor for SEHRS was reported to be on the order of 10^{20} [60,61]. Cross sections that are comparable to good two-photon fluorophores in SEHRS [61] enable two-photon excited vibrational spectra down to the limit of single-molecules [60,62,63] and to collect hyper Raman anti-Stokes spectra [61,64]. Theoretical studies show that the chemical contribution to the overall enhancement in SEHRS can be larger than the corresponding chemical enhancement for SERS [65–68]. The local probing by SEHRS relies even more on very hot hot spots. An extreme localization of optical fields, and aggregate geometry and interparticle distance that have been discussed to exert an important influence on the enhancement in SERS, play an important role in SEHRS as well [69–74].

The hyper Raman process relies on the change in hyperpolarizability of a molecule, and therefore it implies different selection rules, which give complementary spectroscopic information. Depending on the molecular symmetry, we may probe IR active modes or also silent modes, seen neither in Raman nor in IR spectra [75]. In this way, SEHRS can improve structural characterization, in particular of molecule–metal interaction [76], sensing [77–79], and imaging [41,80]. The selection rules for (non-resonantly excited) SEHRS have been discussed by comparing SERS and SEHRS spectra from small organic molecules [81–84], including nucleobases [85,86], amino acids [38], drugs [76], or dyes [87]. In the case of centrosymmetric molecules, the SERS and SEHRS spectra are to great extent complementary [81–83]. When the symmetry is lowered during adsorption to the metal, Raman-active modes become visible in the SEHRS spectrum as well. The high sensitivity of SEHRS with respect to the interaction and orientation of the molecules at the surfaces enables probing of very local surface environment using vibrational information that is inaccessible by SERS [76,83].

In the context of protein characterization, the high sensitivity with respect to changes at the surface of plasmonic nanostructures makes SEHRS ideal to probe the interaction of nanostructures with proteins in biological samples, analogous to those discussed above using one-photon excited SERS. SEHRS spectra obtained from macrophage cells that incorporated gold nanoparticles show, for example, pronounced amide II signals, usually only detected by IR absorption spectroscopy and rare in SERS data [61]. In SEHRS, they can be observed at higher lateral resolution, corresponding to that of an NIR-to-visible wavelength rather than one in the mid-IR. Interestingly, SEHRS observes also spectral bands above 1800 cm^{-1}, which are likely combination modes [61]. In order to better understand the SEHRS spectra from complex samples such as a nanoparticle corona in situ, the nonresonant SEHRS spectra of typical biomolecules, in particular amino acids [38] and nucleic acids [61,85,86], have been discussed in the more recent past.

18.3.2 Toward surface-enhanced Raman scattering spectra of protein side chains

Although SEHRS spectra of proteins apparently contain signals assigned to modes of the protein backbone, such as the amide II band, the interaction of the macromolecules with metal nanostructures occurs to a great extent via the amino acid side chains. Therefore spectra of amino acids, and in particular those of high hyper Raman cross sections, will help to the characterize more complex protein molecules. Fig. 18.4A shows SEHRS spectra of tryptophan, histidine, phenylalanine, and tyrosine, obtained at 1064-nm excitation with a ps pulsed laser in aqueous solution containing silver nanoparticles. Fig. 18.4B shows the corresponding one-photon excited SERS data at 532 nm from the same samples (Fig. 18.4B) for comparison. The spectra were measured under basic pH to promote the interaction with the nanoparticle surface. All four molecules have different SEHRS and SERS spectra (compare each spectrum from Fig. 18.4A with the corresponding spectrum from Fig. 18.4B), as a consequence of the different selection rules that are in place for the one- and two-photon Raman process. The comparison shows that several bands change their relative intensities. For example, while the SEHRS spectrum of tryptophan is dominated by the band at 1552 cm^{-1}, its relative intensity is much smaller in the SERS spectrum (Fig. 18.4, uppermost traces). In contrast, the asymmetric COO^- stretching vibration at 1620 cm^{-1} is relatively weak in the SEHRS spectrum of tryptophan, but very strong in the SERS spectrum (see the shoulder at 1619 cm^{-1}). SEHRS and SERS spectra differ most for histidine (second trace in Fig. 18.4A and B) and phenylalanine (third trace in Fig. 18.4A and B). The SEHRS spectrum of histidine is dominated by three bands, those associated mainly with C = N and C−N stretching modes of the imidazole ring at 1578 and 1047 cm^{-1}. In contrast, the SERS data (second trace in Fig. 18.4B) show other strong contributions in the spectral region between 1100 and 1600 cm^{-1}. Also in the SERS spectrum of phenylalanine (third trace in Fig. 18.4B) very strong signals at 1599 and 1001 cm^{-1}, assigned to the ring stretching and ring breathing modes, respectively, are visible, which are less prominent in the SEHRS spectrum (third trace in Fig. 18.4A). Pronounced contributions from the symmetric COO^- stretching mode are common for the SEHRS spectra of all four molecules: at 1355 cm^{-1} for tryptophan, at 1394 cm^{-1} for histidine, at 1389 cm^{-1} for phenylalanine, and at 1392 cm^{-1} for tyrosine. This may be explained by the fact that at alkaline pH, which was employed in the experiments reported here, the COO^- group is expected to play an important role in the binding of the amino acids to the silver surface.

In general, relatively weak signals of the ring breathing modes of the amino acids are common to all two-photon excited spectra displayed in Fig. 18.4A. In contrast, in SERS, these bands are often used as Raman marker bands because of their high intensity [88–90]. The most prominent example is phenylalanine, where only a very weak band in the SEHRS spectrum at 1002 cm^{-1} is observed (third trace in Fig. 18.4A), while the SERS signal at 1001 cm^{-1} is extremely strong (third trace in Fig. 18.4B). In the case of tyrosine, the typical band around

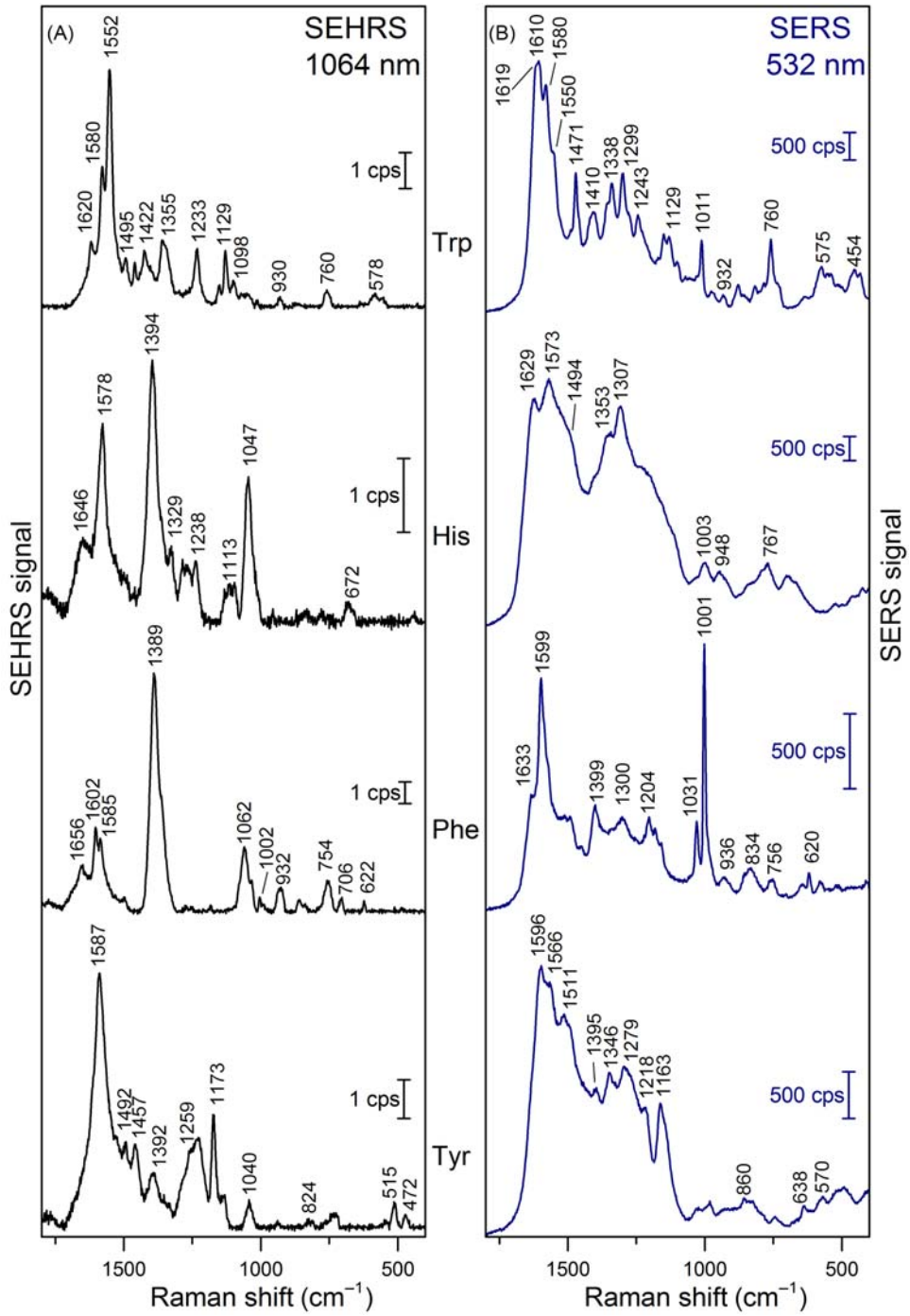


Figure 18.4

SEHRS spectra (A) and SERS spectra (B) of the amino acids tryptophan, histidine, phenylalanine, and tyrosine (from top to bottom) obtained with silver nanoaggregates. Excitation: 1064 nm (A) and 532 nm (B), photon flux density: 3×10^{28} photons/cm^2/s (A) and 6×10^{23} photons/cm/s (B), acquisition time: 60 s (A) and 5 s (B); concentration: 2×10^{-3} M; pH 10. Source*: Reprinted with permission from F. Madzharova, Z. Heiner, J. Kneipp, Surface enhanced hyper-Raman scattering of the amino acids tryptophan, histidine, phenylalanine, and tyrosine, J. Phys. Chem. C 121 (2) (2017) 1235−1242. Copyright (2017) American Chemical Society.*

850 cm^{-1}, containing the ring breathing mode [91], shows very low contributions as a weak broad band around 860 cm^{-1} in the SERS spectrum (bottommost trace in Fig. 18.4B), and is not observed in the SEHRS spectrum (bottommost trace in Fig. 18.4A). The very low intensity in SERS indicates a flat coordination of the species on the nanoaggregates' surface [92–94]. As was found for the ring breathing modes in the nucleic acids [86], also the SEHRS spectra of the aromatic amino acids have high similarity with the corresponding infrared data [95].

The two-photon excited spectra of the amino acids are very sensitive with respect to molecule–nanoparticle interactions. This can pose limitations in protein probing, but could also have enormous advantages. As an example, when SEHRS experiments performed at slightly varied concentration of the amino acids histidine and tyrosine with silver nanoaggregates, the spectra differ a lot, while the SERS spectra of the same samples were almost unchanged [38]. Such changes in the SEHRS spectra indicate that a different interaction with the silver nanostructures takes place when the concentration of the molecules is lowered slightly. It is very likely that a different selectivity leads to probing of different molecular adsorbate species in SEHRS and SERS [78]. It should be noted that, in a nanostructure providing high SEHRS signals, a specific surface chemistry (e.g., surface coverage with magnesium and sulfate ions) can lead to different concentration dependence of the interaction of the molecules than in aggregates that provide high SERS enhancement. Moreover, the positions of highest enhancement, in particular when using nanoparticle aggregates, may vary greatly for SERS and SEHRS experiments, as the enhancement of the excitation field plays a major role in the electromagnetic enhancement in SEHRS. This was demonstrated recently in SEHRS experiments with gold nanoparticles and the corresponding finite-difference time-domain (FDTD) simulations of the distributions of the local fields of gold nanoparticles from nanoparticle/nanoaggregate geometries that are typically used in SERS and SEHRS experiments [96]. The possibility to probe different sites of interaction between the amino acid molecules and parts of the silver nanostructures in combined SEHRS and SERS experiments will help to understand better the formation and composition of biomolecular coronae around nanoparticles in biological environments, which is mainly determined by adsorption of proteins and amino acids [27,97].

18.4 Studies of individual protein molecules in the hot spots of nanolenses

18.4.1 Plasmonic nanolenses for highly localized surface-enhanced Raman scattering probing

When gold and silver particles form nanoaggregates, SERS enhancement can increase dramatically and is very similar for both kinds of aggregates [73,98], in spite of different enhancement found for the different isolated nanoparticles [99,100]. In a solution of

nanoparticles, the addition of analyte can cause aggregation, which leads to an enormous increase of the enhancement, as has been demonstrated extensively both in experiment [98,101,102] and theory [103,104]. Self-similar chains of spherical nanoparticles, termed nanolenses [105] provide very high local fields through a cascaded enhancement mechanism [105–107]. Although the general superiority of the nanolens structures over dimers of two nanoparticles of identical size has been debated [108,109], experiments on nanolenses generated by top-down laser ablation verified the strong SERS enhancement for particular arrangements of gold nanostructures, in agreement with theoretical considerations [110]. Fig. 18.5 shows the distribution of the local field E relative to the excitation field E_0

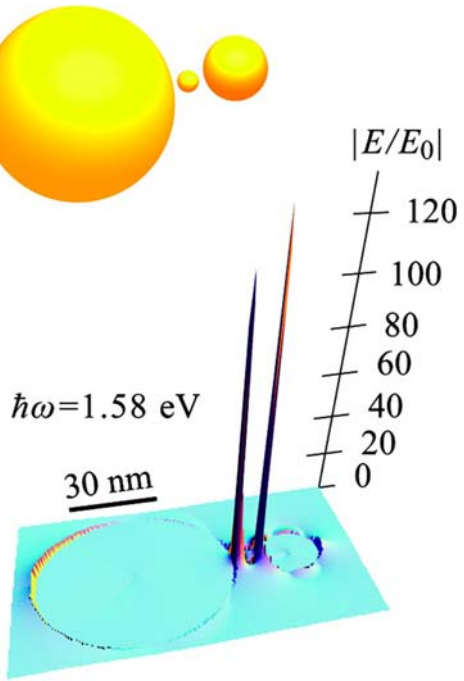


Figure 18.5

Schematic of a nanolens geometry. The radii and small surface-to-surface gaps between the gold nanoparticles are those from typical nanostructures as they can be obtained by laser ablation of a gold foil. The radii of the nanoparticles have been chosen to be 30:3.3:10 nm (from left to right) to better conform with the requirements of the quasistatic approximation. The distribution of the local field E relative to the excitation field E_0 in the symmetry plane of the nanolens is shown at the bottom. In agreement with typical excitation wavelengths applied in experiments with biomolecules in the near infrared, the photon energy of the excitation radiation is chosen here $\hbar\omega = 1.58$ eV, its polarization is along the axis of the nanolens. The local field enhancement $|E/E_0|$ is plotted along the vertical axis. Suggested nanoscale of sizes (30 nm) is indicated at the bottom. Source: *Reprinted with permission from J. Kneipp, X. Li, M. Sherwood, U. Panne, H. Kneipp, M.I. Stockman, et al., Gold nanolenses generated by laser ablation-efficient enhancing structure for surface enhanced Raman scattering analytics and sensing, Anal. Chem. 80 (2008) 4247–4251. Copyright (2008) American Chemical Society.*

in a gold nanolens, for a typical excitation wavelength in the near infrared, as it is often used in nonresonant Raman experiments with biological samples. The maximum local field E is predicted at the axis in the gap between the smallest and the medium nanosphere, where it is enhanced with respect to the excitation field E_0 by a factor $|E/E_0| \sim 130$. As the SERS signal is proportional to $|E/E_0|^4$, this yields an enhancement factor of $\sim 3 \times 10^8$ for this hottest spot.

The defined assembly of groupings of relatively small gold nanoparticles indicated that nucleic acids can serve as molecular templates for nanolens structures [111]. With the advancement of DNA nanotechnology, DNA origami constructs to achieve a defined arrangement of gold [112,113] and recently also silver nanospheres [114] have been used to build nanoparticle dimers and multimers. The assembly of gold nanolenses by using the DNA origami technology enables systematic experiments on the localization of the hottest spots by placing probe molecules in different gaps between the different particles [115]. The simulation results in Fig. 18.6 demonstrate that the electric field enhancement is affected significantly, when gold nanoparticle positions differ in slightly varied nanolens designs. Simulations with gold nanolenses comprised of three nanoparticles of 10, 20, and 60 nm, respectively, provide evidence of the strong confinement of the hottest spots, and that the SERS enhancement strongly depends on how closely individual structures match ideal collinear geometries with smallest gaps (Fig. 18.6A–C). The simulation shows that an additional 5-nm gap between a 10- and a 20-nm gold nanoparticle causes the SERS signal in the design in Fig. 18.6C and F to drop by two orders of magnitude compared to the other arrangements. In an experiment, the actual geometry of an individual nanolens will deviate from the collinear arrangement shown in the simulation though, since in the presence of a surface, the nanolenses lie flat on this surface during probing, the DNA origami scaffold may be planar, and the nanoparticle-anchoring strands in the origami scaffold allow for some flexibility [115].

18.4.2 Detection of single protein molecules in individual gold nanolenses

Recently, a robust coating protocol was established and shown to create stable DNA-coated silver nanoparticles with different coating strand sequences ($(XTT)_3$-T_4-thiol, with X = A, T, G), which enabled the assembly of silver nanoparticles into silver nanolenses and consequently the first SERS study with solid silver nanoparticles on DNA origami [114]. The silver nanolenses were assembled from 10-, 20-, and 60-nm silver nanoparticles, with the 10-nm particle in the center, as illustrated in Fig. 18.7A. The DNA origami platform can control the position and number of analyte protein molecules by incorporating respective binding functionalities. Here, the 5′-end of a single strand in the DNA origami scaffold was modified with a biotin group, located between the attachment points for the 20- and 10-nm particles (Fig. 18.7A). Biotin forms a strong noncovalent bond to the protein streptavidin. By observing a specific sequence in the nanolens assembly process, a single

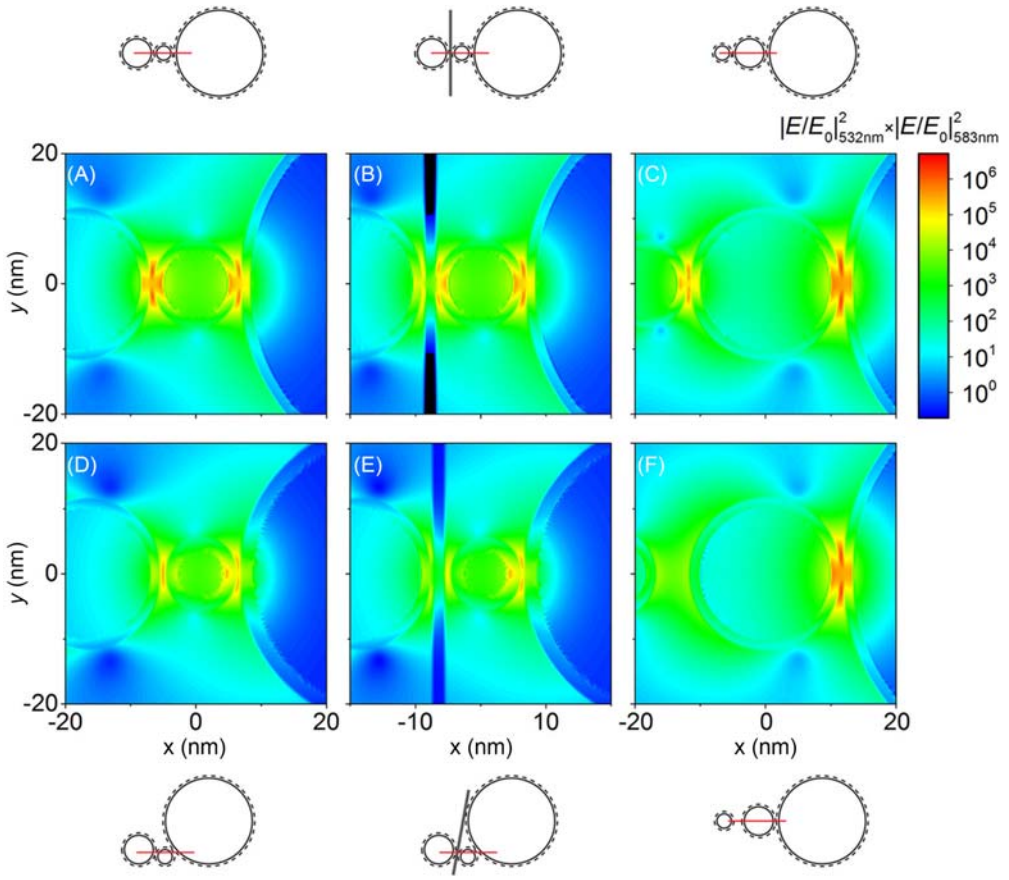


Figure 18.6

Electromagnetic SERS enhancement based on FDTD simulations of the local fields for three designs of gold nanolenses at an excitation wavelength of 532 nm and at the wavelength of a typical Raman shift at 1650 cm^{-1} (583 nm). (A–C) In-line, no gap (D, E) out-of-line, (F) 5-nm-gap arrangement. The polarization direction of the incident light is along the longitudinal axis of the nanolens. The red line in the schematics illustrates the plane of observation; the plane is chosen to intersect the position of greatest field intensity. In two of the arrangements (B, E), the DNA origami scaffold lies between a 20-nm gold nanoparticle and a 10- and 60-nm nanoparticle pair, it is indicated in gray. Source*: Reprinted with permission from C. Heck, J. Prinz, A. Dathe, V. Merk, O. Stranik, W. Fritzsche, et al., Gold nanolenses self-assembled by DNA origami, ACS Photon. 4 (5) (2017) 1123–1130. Copyright (2017) American Chemical Society.*

streptavidin can be placed in a particular gap of a silver nanolens (Fig. 18.7A). In AFM experiments conducted during the stepwise assembly of the nanolenses (Fig. 18.7A), it could be confirmed that the 10-nm silver particles were integrated with yields close to 100% [114]. Due to the pronounced SERS signals originating from the DNA origami scaffold itself, the lysine residues of the streptavidine molecules were modified with alkyne

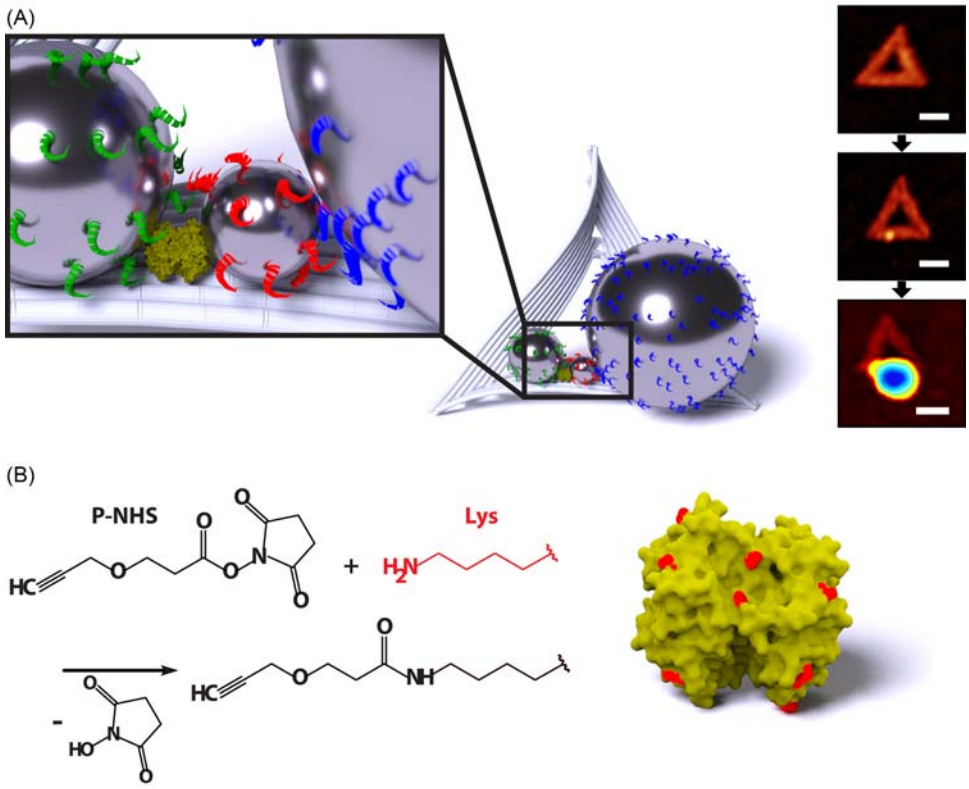


Figure 18.7

(A) A biotin molecule on the DNA origami scaffold immobilizes a single modified streptavidin; after the particles are assembled, it resides in the selected nanolens gap. The AFM images on the right show examples for a bare DNA origami, one with streptavidin and a full silver nanolens. The 60-nm particle is seen as a blue spot, the 20-nm particle as a yellow shoulder. The underlying DNA origami scaffold is dark red. Scale bars: 50 nm. (B) Streptavidin's lysine residues (*red*) are modified with an alkyne group to facilitate an unambiguous Raman detection. Source: *Reprinted with permission from C. Heck, Y. Kanehira, J. Kneipp, I. Bald, Placement of single proteins within the SERS hot spots of self-assembled silver nanolenses, Angew. Chem. Int. Ed. 57 (25) (2018) 7444–7447. Copyright (2018) John Wiley and Sons.*

groups (Fig. 18.7B), which provide a typical signal of the C≡C stretching vibration at 2121 cm^{-1}. Streptavidin was modified with an average of 19 alkyne groups on the full tetrameric protein, as determined by MALDI-TOF MS (Fig. 18.7B). To avoid aggregate formation by several nanolenses in ensemble measurements in solution, SERS spectra were measured from individual nanolenses deposited at high dilution on a silicon wafer, and a statistical analysis of the alkyne signal was carried out. In these analyses, the sample with labeled streptavidin shows a large number of data points for the bin around 2120 cm^{-1}, whereas for the control containing no streptavidin, only very few events were observed.

Enhancement by different individual nanolenses varies considerably, in agreement with the high sensitivity of enhancement toward gap size and precise nanoaggregate geometry [69,116], and indicates that the structural control exerted by the DNA nanostructures is not precise enough yet and the flexibility of the origami scaffold still quite high. Comparison of the experiments with those on self-assembled gold dimers [113], the interparticle gaps are expected to be around 3 nm. The fact that it was possible to detect a signal from the alkyne labels of a single streptavidin indicates that only a few of the side chain residues of the protein, which measures $\sim$5 nm in diameter, protruded into the gap between the particles. Strategies to improve particle alignment and to decrease the coating thickness will be key to develop silver nanolenses that can fulfill the theoretical predictions and provide extreme field enhancements.

Acknowledgments

I am grateful to Daniela Drescher, Christian Heck, Fani Madzharova, and Gergő Péter Szekeres who all have contributed to the research on SERS of proteins in our group over the last decade and to the works discussed here. I would like to thank G.P. Szekeres for assembling Figures 18.2 and 18.3. Funding was provided by DFG GSC 1013 SALSA and by ERC grant 259432.

References

[1] M. Fleischmann, P.J. Hendra, A.J. McQuillan, Raman spectra of pyridine adsorbed at a silver electrode, Chem. Phys. Lett. 26 (2) (1974) 163–166.
[2] M.G. Albrecht, J.A. Creighton, Anomalously intense Raman spectra of pyridine at a silver electrode, J. Am. Chem. Soc. 99 (15) (1977) 5215–5217.
[3] D.L. Jeanmaire, R.P. Van Duyne, Surface Raman spectroelectrochemistry, J. Electroanalyt. Chem. Interf. Electrochem. 84 (1) (1977) 1–20.
[4] M. Moskovits, Surface-enhanced spectroscopy, Rev. Mod. Phys. 57 (3) (1985) 783–826.
[5] A. Otto, I. Mrozek, H. Grabhorn, W. Akemann, Surface-enhanced Raman scattering, J. Phys. Condens. Matter 4 (1992) 1143–1212.
[6] G.C. Schatz, Theoretical studies of surface enhanced Raman scattering, Acc. Chem. Res. 17 (10) (1984) 370–376.
[7] J.R. Lombardi, R.L. Birke, T.H. Lu, J. Xu, Charge-transfer theory of surface enhanced Raman-spectroscopy - Herzberg-Teller contributions, J. Chem. Phys. 84 (8) (1986) 4174–4180.
[8] J.R. Lombardi, R.L. Birke, A unified view of surface-enhanced Raman scattering, Acc. Chem. Res. 42 (6) (2009) 734–742.
[9] I. Pavel, E. McCarney, A. Elkhaled, A. Morrill, K. Plaxco, M. Moskovits, Label-free SERS detection of small proteins modified to act as bifunctional linkers, J. Phys. Chem. C 112 (13) (2008) 4880–4883.
[10] X.X. Han, Y. Kitahama, T. Itoh, C.X. Wang, B. Zhao, Y. Ozaki, Protein-mediated sandwich strategy for surface-enhanced Raman scattering: application to versatile protein detection, Anal. Chem. 81 (9) (2009) 3350–3355.
[11] B. Fazio, C. D'Andrea, A. Foti, E. Messina, A. Irrera, M.G. Donato, et al., SERS detection of biomolecules at physiological pH via aggregation of gold nanorods mediated by optical forces and plasmonic heating, Sci. Rep. 6 (2016) 26952.
[12] G.P. Szekeres, J. Kneipp, SERS probing of proteins in gold nanoparticle agglomerates, Front. Chem. 7 (2019) 30.

[13] H.X. Xu, E.J. Bjerneld, M. Kall, L. Borjesson, Spectroscopy of single hemoglobin molecules by surface enhanced Raman scattering, Phys. Rev. Lett. 83 (21) (1999) 4357–4360.
[14] B.D. Chithrani, W.C. Chan, Elucidating the mechanism of cellular uptake and removal of protein-coated gold nanoparticles of different sizes and shapes, Nano Lett. 7 (6) (2007) 1542–1550.
[15] C.C. Ge, J.F. Du, L.N. Zhao, L.M. Wang, Y. Liu, D.H. Li, et al., Binding of blood proteins to carbon nanotubes reduces cytotoxicity, Proc. Natl Acad. Sci. USA 108 (41) (2011) 16968–16973.
[16] M. Lundqvist, J. Stigler, G. Elia, I. Lynch, T. Cedervall, K.A. Dawson, Nanoparticle size and surface properties determine the protein corona with possible implications for biological impacts, Proc. Natl Acad. Sci. USA 105 (38) (2008) 14265–14270.
[17] A.E. Nel, L. Mädler, D. Velegol, T. Xia, E.M.V. Hoek, P. Somasundaran, et al., Understanding biophysicochemical interactions at the nano-bio interface, Nat. Mater. 8 (7) (2009) 543–557.
[18] P. Nativo, I.A. Prior, M. Brust, Uptake and intracellular fate of surface-modified gold nanoparticles, ACS Nano 2 (8) (2008) 1639–1644.
[19] T. Cedervall, I. Lynch, S. Lindman, T. Berggard, E. Thulin, H. Nilsson, et al., Understanding the nanoparticle-protein corona using methods to quantify exchange rates and affinities of proteins for nanoparticles, Proc. Natl Acad. Sci. USA 104 (7) (2007) 2050–2055.
[20] M.P. Monopoli, D. Walczyk, A. Campbell, G. Elia, I. Lynch, F.B. Bombelli, et al., Physical-chemical aspects of protein corona: relevance to in vitro and in vivo biological impacts of nanoparticles, J. Am. Chem. Soc. 133 (8) (2011) 2525–2534.
[21] S.H. Brewer, W.R. Glomm, M.C. Johnson, M.K. Knag, S. Franzen, Probing BSA binding to citrate-coated gold nanoparticles and surfaces, Langmuir 21 (20) (2005) 9303–9307.
[22] G. Maiorano, S. Sabella, B. Sorce, V. Brunetti, M.A. Malvindi, R. Cingolani, et al., Effects of cell culture media on the dynamic formation of protein-nanoparticle complexes and influence on the cellular response, ACS Nano 4 (12) (2010) 7481–7491.
[23] J. Kneipp, H. Kneipp, M. McLaughlin, D. Brown, K. Kneipp, In vivo molecular probing of cellular compartments with gold nanoparticles and nanoaggregates, Nano Lett. 6 (10) (2006) 2225–2231.
[24] L.A. Austin, B. Kang, M.A. El-Sayed, Probing molecular cell event dynamics at the single-cell level with targeted plasmonic gold nanoparticles: a review, Nano Today 10 (5) (2015) 542–558.
[25] M.A. Ochsenkühn, P.R.T. Jess, H. Stoquert, K. Dholakia, C.J. Campbell, Nanoshells for surface-enhanced Raman spectroscopy in eukaryotic cells: cellular response and sensor development, ACS Nano 3 (11) (2009) 3613–3621.
[26] T. Buchner, D. Drescher, H. Traub, P. Schrade, S. Bachmann, N. Jakubowski, et al., Relating surface-enhanced Raman scattering signals of cells to gold nanoparticle aggregation as determined by la-ICP-MS micromapping, Anal. Bioanal. Chem. 406 (27) (2014) 7003–7014.
[27] D. Drescher, P. Guttmann, T. Buchner, S. Werner, G. Laube, A. Hornemann, et al., Specific biomolecule corona is associated with ring-shaped organization of silver nanoparticles in cells, Nanoscale 5 (19) (2013) 9193–9198.
[28] J. Ando, K. Fujita, N.I. Smith, S. Kawata, Dynamic SERS imaging of cellular transport pathways with endocytosed gold nanoparticles, Nano Lett. 11 (12) (2011) 5344–5348.
[29] T. Buchner, D. Drescher, V. Merk, H. Traub, P. Guttmann, S. Werner, et al., Biomolecular environment, quantification, and intracellular interaction of multifunctional magnetic SERS nanoprobes, Analyst 141 (17) (2016) 5096–5106.
[30] D. Drescher, I. Zeise, H. Traub, P. Guttmann, S. Seifert, T. Buchner, et al., In situ characterization of SiO_2 nanoparticle biointeractions using brightsilica, Adv. Funct. Mater. 24 (24) (2014) 3765–3775.
[31] H. Sugeta, A. Go, S-s and c-s stretching vibrations and molecular conformations of dialkyl disulfides and cystine, Chem. Lett. 1 (1972) 83–86.
[32] T. Watanabe, H. Maeda, Adsorption-controlled redox activity. Surface-enhanced Raman investigation of cystine versus cysteine on silver electrodes, J. Phys. Chem. 93 (1989) 3258–3260.
[33] S. Stewart, P.M. Fredericks, Surface-enhanced Raman spectroscopy of amino acids adsorbed on an electrochemically prepared silver surface, Spectrochim. Acta A 55 (7–8) (1999) 1641–1660.

[34] R.I. MacCuspie, A.J. Allen, V.A. Hackley, Dispersion stabilization of silver nanoparticles in synthetic lung fluid studied under in situ conditions, Nanotoxicology 5 (2) (2011) 140–156.
[35] B. Tang, S. Xu, J. Tao, Y. Wu, W. Xu, Y. Ozaki, Two-dimensional correlation localized surface plasmon resonance spectroscopy for analysis of the interaction between metal nanoparticles and bovine serum albumin, J. Phys. Chem. C 114 (49) (2010) 20990–20996.
[36] S. Laera, G. Ceccone, F. Rossi, D. Gilliland, R. Hussain, G. Siligardi, et al., Measuring protein structure and stability of protein-nanoparticle systems with synchrotron radiation circular dichroism, Nano Lett. 11 (10) (2011) 4480–4484.
[37] A.M. Michaels, M. Nirmal, L.E. Brus, Surface enhanced Raman spectroscopy of individual rhodamine 6g molecules on large Ag nanocrystals, J. Am. Chem. Soc. 121 (43) (1999) 9932–9939.
[38] F. Madzharova, Z. Heiner, J. Kneipp, Surface enhanced hyper-Raman scattering of the amino acids tryptophan, histidine, phenylalanine, and tyrosine, J. Phys. Chem. C 121 (2) (2017) 1235–1242.
[39] A. Hornemann, D. Drescher, S. Flemig, J. Kneipp, Intracellular SERS hybrid probes using bsa-reporter conjugates, Anal. Bioanal. Chem. 405 (19) (2013) 6209–6222.
[40] S.K. Kim, M.S. Kim, S.W. Suh, Surface-enhanced Raman scattering (SERS) of aromatic amino acids and their glycyl dipeptides in silver sol, J. Raman Spectrosc. 18 (3) (1987) 171–175.
[41] Z. Heiner, M. Gühlke, V. Živanović, F. Madzharova, J. Kneipp, Surface-enhanced hyper Raman hyperspectral imaging and probing in animal cells, Nanoscale 9 (23) (2017) 8024–8032.
[42] G.P. Szekeres, J. Kneipp, Different binding sites of serum albumins in the protein corona of gold nanoparticles, Analyst 143 (24) (2018) 6061–6068.
[43] V.J. Lin, J.L. Koenig, Raman studies of bovine serum albumin, Biopolymers 15 (1) (1976) 203–218.
[44] C. Blum, T. Schmid, L. Opilik, S. Weidmann, S.R. Fagerer, R. Zenobi, Understanding tip-enhanced Raman spectra of biological molecules: a combined Raman, SERS and ters study, J. Raman Spectrosc. 43 (12) (2012) 1895–1904.
[45] T. Brulé, A. Bouhelier, A. Dereux, E. Finot, Discrimination between single protein conformations using dynamic SERS, ACS Sens. 1 (6) (2016) 676–680.
[46] J. Shi, Y. Hedberg, M. Lundin, I.O. Wallinder, H.L. Karlsson, L. Moller, Hemolytic properties of synthetic nano- and porous silica particles: the effect of surface properties and the protection by the plasma corona, Acta Biomater. 8 (9) (2012) 3478–3490.
[47] A. Wang, Y.R. Perera, M.B. Davidson, N.C. Fitzkee, Electrostatic interactions and protein competition reveal a dynamic surface in gold nanoparticle-protein adsorption, J. Phys. Chem. C Nanomater. Interf. 120 (42) (2016) 24231–24239.
[48] S. Dominguez-Medina, S. McDonough, P. Swanglap, C.F. Landes, S. Link, In situ measurement of bovine serum albumin interaction with gold nanospheres, Langmuir 28 (24) (2012) 9131–9139.
[49] W. Humphrey, A. Dalke, K. Schulten, Vmd: visual molecular dynamics, J. Mol. Graph. 14 (1) (1996) 33–38. 27-8.
[50] A. Bujacz, K. Zielinski, B. Sekula, Structural studies of bovine, equine, and leporine serum albumin complexes with naproxen, Proteins 82 (9) (2014) 2199–2208.
[51] A.I. Ivanov, R.G. Zhbankov, E.A. Korolenko, E.V. Korolik, L.A. Meleshchenko, M. Marchewka, et al., Infrared and Raman spectroscopic studies of the structure of human serum albumin under various ligand loads, J. Appl. Spectrosc. 60 (1994) 305–309.
[52] N.C. Dingari, G.L. Horowitz, J.W. Kang, R.R. Dasari, I. Barman, Raman spectroscopy provides a powerful diagnostic tool for accurate determination of albumin glycation, PLoS One 7 (2) (2012) e32406.
[53] D.-H. Tsai, F.W. DelRio, A.M. Keene, K.M. Tyner, R.I. MacCuspie, T.J. Cho, et al., Adsorption and conformation of serum albumin protein on gold nanoparticles investigated using dimensional measurements and in situ spectroscopic methods, Langmuir 27 (6) (2011) 2464–2477.
[54] N.C. Fitzkee, G.D. Rose, Reassessing random-coil statistics in unfolded proteins, Proc. Natl Acad. Sci. USA 101 (34) (2004) 12497–12502.
[55] V. Joseph, A. Matschulat, J. Polte, S. Rolf, F. Emmerling, J. Kneipp, SERS enhancement of gold nanospheres of defined size, J. Raman Spectrosc. 42 (9) (2011) 1736–1742.

[56] S.T. Moerz, A. Kraegeloh, M. Chanana, T. Kraus, Formation mechanism for stable hybrid clusters of proteins and nanoparticles, ACS Nano 9 (7) (2015) 6696–6705.
[57] D. Drescher, C. Giesen, H. Traub, U. Panne, J. Kneipp, N. Jakubowski, Quantitative imaging of gold and silver nanoparticles in single eukaryotic cells by laser ablation ICP-MS, Anal. Chem. 84 (22) (2012) 9684–9688.
[58] V. Živanović, G. Semini, M. Laue, D. Drescher, T. Aebischer, J. Kneipp, Chemical mapping of leishmania infection in live cells by SERS microscopy, Anal. Chem. 90 (13) (2018) 8154–8161.
[59] V.N. Denisov, B.N. Mavrin, V.B. Podobedov, Hyper-Raman scattering by vibrational excitations in crystals, glasses and liquids, Phys. Rep. 151 (1) (1987) 1–92.
[60] K. Kneipp, H. Kneipp, I. Itzkan, R.R. Dasari, M.S. Feld, Surface-enhanced non-linear Raman scattering at the single-molecule level, Chem. Phys. 247 (1) (1999) 155–162.
[61] J. Kneipp, H. Kneipp, K. Kneipp, Two-photon vibrational spectroscopy for biosciences based on surface-enhanced hyper-Raman scattering, Proc. Natl Acad. Sci. 103 (46) (2006) 17149–17153.
[62] H. Yoshikawa, T. Adachi, G. Sazaki, T. Matsui, K. Nakajima, H. Masuhara, Surface-enhanced hyper-Raman spectroscopy using optical trapping of silver nanoparticles for molecular detection in solution, J. Opt. A Pure Appl. Opt. 9 (8) (2007) S164.
[63] C.B. Milojevich, B.K. Mandrell, H.K. Turley, V. Iberi, M.D. Best, J.P. Camden, Surface-enhanced hyper-Raman scattering from single molecules, J. Phys. Chem. Lett. 4 (20) (2013) 3420–3423.
[64] S. Nie, L.A. Lipscomb, N.-T. Yu, Surface-enhanced hyper-Raman spectroscopy, Appl. Spectrosc. Rev. 26 (3) (1991) 203–276.
[65] J.T. Golab, J.R. Sprague, K.T. Carron, G.C. Schatz, R.P.V. Duyne, A surface enhanced hyper-Raman scattering study of pyridine adsorbed onto silver: experiment and theory, J. Chem. Phys. 88 (12) (1988) 7942–7951.
[66] W.H. Yang, G.C. Schatz, Ab initio and semiempirical molecular orbital studies of surface enhanced and bulk hyper-Raman scattering from pyridine, J. Chem. Phys. 97 (5) (1992) 3831–3845.
[67] N. Valley, L. Jensen, J. Autschbach, G.C. Schatz, Theoretical studies of surface enhanced hyper-Raman spectroscopy: the chemical enhancement mechanism, J. Chem. Phys. 133 (5) (2010) 054103.
[68] J. Mullin, N. Valley, M.G. Blaber, G.C. Schatz, Combined quantum mechanics (TDDFT) and classical electrodynamics (mie theory) methods for calculating surface enhanced Raman and hyper-Raman spectra, J. Phys. Chem. A 116 (38) (2012) 9574–9581.
[69] P. Nordlander, C. Oubre, E. Prodan, K. Li, M.I. Stockman, Plasmon hybridization in nanoparticle dimers, Nano Lett. 4 (5) (2004) 899–903.
[70] T. Itoh, H. Yoshikawa, K. Yoshida, V. Biju, M. Ishikawa, Evaluation of electromagnetic enhancement of surface enhanced hyper Raman scattering using plasmonic properties of binary active sites in single Ag nanoaggregates, J. Chem. Phys. 130 (21) (2009) 214706.
[71] M.K. Hossain, G.G. Huang, T. Kaneko, Y. Ozaki, Surface-enhanced Raman scattering and plasmon excitations from isolated and elongated gold nanoaggregates, Chem. Phys. Lett. 477 (1) (2009) 130–134.
[72] S.L. Kleinman, B. Sharma, M.G. Blaber, A.-I. Henry, N. Valley, R.G. Freeman, et al., Structure enhancement factor relationships in single gold nanoantennas by surface-enhanced Raman excitation spectroscopy, J. Am. Chem. Soc. 135 (1) (2013) 301–308.
[73] K. Kneipp, H. Kneipp, J. Kneipp, Surface-enhanced Raman scattering in local optical fields of silver and gold nanoaggregatesfrom single-molecule Raman spectroscopy to ultrasensitive probing in live cells, Acc. Chem. Res. 39 (7) (2006) 443–450.
[74] M.I. Stockman, Nanoplasmonics: past, present, and glimpse into future, Opt. Express 19 (22) (2011) 22029–22106.
[75] J.H. Christie, D.J. Lockwood, Selection rules for three- and four-photon Raman interactions, J. Chem. Phys. 54 (3) (1971) 1141–1154.
[76] V. Živanović, F. Madzharova, Z. Heiner, C. Arenz, J. Kneipp, Specific interaction of tricyclic antidepressants with gold and silver nanostructures as revealed by combined one- and two-photon vibrational spectroscopy, J. Phys. Chem. C 121 (41) (2017) 22958–22968.

[77] H.K. Turley, J.P. Camden, A nonlinear approach to surface-enhanced sensing in the short-wave infrared, Chem. Commun. 50 (12) (2014) 1472–1474.
[78] M. Gühlke, Z. Heiner, J. Kneipp, Combined near-infrared excited SEHRS and SERS spectra of pH sensors using silver nanostructures, Phys. Chem. Chem. Phys. 17 (39) (2015) 26093–26100.
[79] M.J. Trujillo, J.P. Camden, Utilizing molecular hyperpolarizability for trace analysis: a surface-enhanced hyper-Raman scattering study of uranyl ion, ACS Omega 3 (6) (2018) 6660–6664.
[80] M. Gühlke, Z. Heiner, J. Kneipp, Surface-enhanced hyper-Raman and Raman hyperspectral mapping, Phys. Chem. Chem. Phys. 18 (21) (2016) 14228–14233.
[81] W.-H. Li, X.-Y. Li, N.-T. Yu, Surface-enhanced hyper-Raman spectroscopy (SEHRS) and surface-enhanced Raman spectroscopy (SERS) studies of pyrazine and pyridine adsorbed on silver electrodes, Chem. Phys. Lett. 305 (3–4) (1999) 303–310.
[82] X.-Y. Li, Q.-J. Huang, V.I. Petrov, Y.-T. Xie, Q. Luo, X. Yu, et al., Surface-enhanced hyper-Raman and surface-enhanced Raman scattering from molecules adsorbed on nanoparticles-on-smooth-electrode (nose) substrate i. Pyridine, pyrazine and benzene, J. Raman Spectrosc. 36 (6-7) (2005) 555–573.
[83] J.C. Hulteen, M.A. Young, R.P. Van Duyne, Surface-enhanced hyper-Raman scattering (SEHRS) on Ag film over nanosphere (fon) electrodes: surface symmetry of centrosymmetric adsorbates, Langmuir 22 (25) (2006) 10354–10364.
[84] J. Kneipp, H. Kneipp, B. Wittig, K. Kneipp, One- and two-photon excited optical pH probing for cells using surface-enhanced Raman and hyper-Raman nanosensors, Nano Lett. 7 (9) (2007) 2819–2823.
[85] H. Kneipp, K. Kneipp, Surface-enhanced hyper Raman scattering in silver colloidal solutions, J. Raman Spectrosc. 36 (6-7) (2005) 551–554.
[86] F. Madzharova, Z. Heiner, M. Gühlke, J. Kneipp, Surface-enhanced hyper-Raman spectra of adenine, guanine, cytosine, thymine, and uracil, J. Phys. Chem. C 120 (28) (2016) 15415–15423.
[87] H.K. Turley, Z. Hu, D.W. Silverstein, D.A. Cooper, L. Jensen, J.P. Camden, Probing two-photon molecular properties with surface-enhanced hyper-Raman scattering: a combined experimental and theoretical study of crystal violet, J. Phys. Chem. C 120 (37) (2016) 20936–20942.
[88] S. Seifert, V. Merk, J. Kneipp, Identification of aqueous pollen extracts using surface enhanced Raman scattering (SERS) and pattern recognition methods, J. Biophoton. 9 (1-2) (2016) 181–189.
[89] H. Schulz, M. Baranska, Identification and quantification of valuable plant substances by IR and Raman spectroscopy, Vibrat. Spectrosc. 43 (1) (2007) 13–25.
[90] M. Aioub, M.A. El-Sayed, A real-time surface enhanced Raman spectroscopy study of plasmonic photothermal cell death using targeted gold nanoparticles, J. Am. Chem. Soc. 138 (4) (2016) 1258–1264.
[91] B. Hernández, Y.-M. Coïc, F. Pflüger, S.G. Kruglik, M. Ghomi, All characteristic Raman markers of tyrosine and tyrosinate originate from phenol ring fundamental vibrations, J. Raman Spectrosc. 47 (2) (2016) 210–220.
[92] M. Moskovits, J.S. Suh, Surface selection rules for surface-enhanced Raman spectroscopy: calculations and application to the surface-enhanced Raman spectrum of phthalazine on silver, J. Phys. Chem. 88 (23) (1984) 5526–5530.
[93] J.A. Creighton, Surface Raman electromagnetic enhancement factors for molecules at the surface of small isolated metal spheres: the determination of adsorbate orientation from SERS relative intensities, Surf. Sci. 124 (1) (1983) 209–219.
[94] T. Deckert-Gaudig, E. Rauls, V. Deckert, Aromatic amino acid monolayers sandwiched between gold and silver: a combined tip-enhanced Raman and theoretical approach, J. Phys. Chem. C 114 (16) (2010) 7412–7420.
[95] B. Hernández, F. Pflüger, A. Adenier, S.G. Kruglik, M. Ghomi, Vibrational analysis of amino acids and short peptides in hydrated media. Viii. Amino acids with aromatic side chains: L-phenylalanine, l-tyrosine, and l-tryptophan, J. Phys. Chem. B 114 (46) (2010) 15319–15330.
[96] F. Madzharova, Z. Heiner, J. Simke, S. Selve, J. Kneipp, Gold nanostructures for plasmonic enhancement of hyper-Raman scattering, J. Phys. Chem. C 122 (5) (2018) 2931–2940.

[97] I. Lynch, T. Cedervall, M. Lundqvist, C. Cabaleiro-Lago, S. Linse, K.A. Dawson, The nanoparticle - protein complex as a biological entity; a complex fluids and surface science challenge for the 21st century, Adv. Colloid Interf. Sci. 134–135 (2007) 167–174.
[98] K. Kneipp, H. Kneipp, R. Manoharan, E.B. Hanlon, I. Itzkan, R.R. Dasari, et al., Extremely large enhancement factors in surface-enhanced Raman scattering for molecules on colloidal gold clusters, Appl. Spectrosc. 52 (12) (1998) 1493–1497.
[99] D.S. Wang, M. Kerker, Enhanced Raman scattering by molecules adsorbed at the surface of colloidal spheroids, Phys. Rev. B 24 (4) (1981) 1777.
[100] Z. Zhu, T. Zhu, Z. Liu, Raman scattering enhancement contributed from individual gold nanoparticles and interparticle coupling, Nanotechnology 15 (3) (2004) 357.
[101] C.G. Blatchford, J.R. Campbell, J.A. Creighton, Plasma resonance – enhanced Raman scattering by absorbates on gold colloids: the effects of aggregation, Surf. Sci. 120 (2) (1982) 435–455.
[102] P. Hildebrandt, S. Keller, A. Hoffmann, F. Vanhecke, B. Schrader, Enhancement factor of surface-enhanced Raman scattering on silver and gold surfaces upon near-infrared excitation. Indication of an unusual strong contribution of the chemical effect, J. Raman Spec. 24 (11) (1993) 791–796.
[103] M. Kerker, O. Siiman, D.S. Wang, Effect of aggregates on extinction and surface-enhanced Raman scattering spectra of colloidal silver, J. Phys. Chem. 88 (15) (1984) 3168–3170.
[104] E. Hao, G.C. Schatz, Electromagnetic fields around silver nanoparticles and dimers, J. Chem. Phys. 120 (1) (2004) 357–366.
[105] K. Li, M.I. Stockman, D.J. Bergman, Self-similar chain of metal nanospheres as an efficient nanolens, Phys. Rev. Lett. 91 (22) (2003) 227402.
[106] M.I. Stockman, V.M. Shalaev, M. Moskovits, R. Botet, T.F. George, Enhanced Raman scattering by fractal clusters: scale-invariant theory, Phys. Rev. B 46 (5) (1992) 2821–2830.
[107] J. Dai, F. Čajko, I. Tsukerman, M.I. Stockman, Electrodynamic effects in plasmonic nanolenses, Phys. Rev. B 77 (11) (2008) 115419.
[108] G. Pellegrini, M. Celebrano, M. Finazzi, P. Biagioni, Local field enhancement: comparing self-similar and dimer nanoantennas, J. Phys. Chem. C 120 (45) (2016) 26021–26024.
[109] J.M. Fitzgerald, V. Giannini, Battling retardation and nonlocality: the hunt for the ultimate plasmonic cascade nanolens, ACS Photon. 5 (6) (2018) 2459–2467.
[110] J. Kneipp, X. Li, M. Sherwood, U. Panne, H. Kneipp, M.I. Stockman, et al., Gold nanolenses generated by laser ablation-efficient enhancing structure for surface enhanced Raman scattering analytics and sensing, Anal. Chem. 80 (2008) 4247–4251.
[111] S. Bidault, F.J. García de Abajo, A. Polman, Plasmon-based nanolenses assembled on a well-defined DNA template, J. Am. Chem. Soc. 130 (9) (2008) 2750–2751.
[112] G.P. Acuna, F.M. Möller, P. Holzmeister, S. Beater, B. Lalkens, P. Tinnefeld, Fluorescence enhancement at docking sites of DNA-directed self-assembled nanoantennas, Science 338 (6106) (2012) 506.
[113] V.V. Thacker, L.O. Herrmann, D.O. Sigle, T. Zhang, T. Liedl, J.J. Baumberg, et al., DNA origami based assembly of gold nanoparticle dimers for surface-enhanced Raman scattering, Nat. Commun. 5 (2014) 3448.
[114] C. Heck, Y. Kanehira, J. Kneipp, I. Bald, Placement of single proteins within the SERS hot spots of self-assembled silver nanolenses, Angew. Chem. Int. Ed. 57 (25) (2018) 7444–7447.
[115] C. Heck, J. Prinz, A. Dathe, V. Merk, O. Stranik, W. Fritzsche, et al., Gold nanolenses self-assembled by DNA origami, ACS Photon. 4 (5) (2017) 1123–1130.
[116] K. Kneipp, H. Kneipp, J. Kneipp, Probing plasmonic nanostructures by photons and electrons, Chem. Sci. 6 (5) (2015) 2721–2726.

This page intentionally left blank

CHAPTER 19

Surface-enhanced Raman scattering chemosensing of proteins

Luca Guerrini[1] **and Ramon A. Alvarez-Puebla**[1,2]

[1]Department of Physical and Inorganic Chemistry and EMaS, Universitat Rovira i Virgili, Tarragona, Spain, [2]ICREA, Barcelona, Spain

19.1 Introduction

Proteins are macromolecules formed by one or more chains of linearly aligned amino acids linked through peptide bonds. These biopolymers are responsible for a large variety of biological functions including structural, immunological, metabolic or catalytic activity, signaling, or transporting. Changes in the composition, spatial organization, and concentration of many proteins are often associated with disease.

The application of nanostructured materials for the development of novel detection techniques with improved sensitivity and/or simplified and faster applicability has rapidly become an appealing alternative to other technologies. Among them, surface-enhanced Raman scattering (SERS) spectroscopy is a powerful analytical technique that has already proven to be particularly effective in chemical biology and medicine [1,2]. Essentially, SERS can be described as the amplified Raman scattering by the presence of a plasmonic structure (most often metallic nanostructures) in the close vicinity of the target analyte. In such a case, the main cause of the excitation of the molecular vibrational levels is the collective oscillation of the conduction electrons in the metal, upon excitation with the appropriate light, which generates an ultrastrong electromagnetic near-field (known as localized surface plasmon resonance, LSPR) in the proximity of the nanostructured surface. Like Raman scattering, SERS provides a complex spectral pattern that contains all the compositional and structural characteristics of the target analyte while affording an extreme experimental flexibility. SERS can be carried out over a wide spectral range, is insensitive to water, and, in many cases, requires no sample preparation. Recent spectacular advances in nanofabrication techniques fueled the development of a large variety of rationally designed SERS substrates with optimized, uniform, and reproducible response [3]. This successfully translated the spectacular analytical potential of SERS to reliable, widely accepted, and commercially viable sensing applications. The dependence of LSPRs with

Vibrational Spectroscopy in Protein Research.
DOI: https://doi.org/10.1016/B978-0-12-818610-7.00019-0

parameters such as size, shape, composition, and surrounding medium provides multiple possibilities for tuning the optical response and, thus, optimizing the SERS performance of the plasmonic nanostructure for a specific application. In conjunction with the control of the signal amplification provided by the optical enhancer, a key step in the practical implementation to sensing applications (including SERS) is the appropriate chemical functionalization of the "bare" metallic (i.e., plasmonic) surface which is required to impart the necessary selectivity and/or sensitivity toward the target analyte, especially in complex media. Two design approaches are commonly used to devise plasmonic nanostructures as SERS sensing platforms: direct and indirect SERS.

Acquisition of the intrinsic SERS spectra of proteins provides a powerful characterization tool for dynamically investigating their structural features under biological conditions. However, when applied for pure sensing purposes, this approach faces major problems as proteins are large molecules with low Raman cross section that typically exist at very low concentrations ($\sim$pg/mL) in complex biological fluids. These biomatrices comprise a multitude of other components that can compete with the target protein for adsorption onto the plasmonic surface. Thus, prepurification steps are required to overcome this issue.

Rapid SERS sensing of proteins in their original media is therefore preferably carried out via indirect approaches, which exploit surface receptors to selectively accumulate the target protein at the metal interface. Antibodies, together with peptide and DNA aptamers, are the preferred surface receptors for indirect SERS determination of proteins thanks to their high binding affinity and selectivity. However, these biomolecules are typically characterized by a very small Raman cross section and thus an extrinsic SERS label has to be incorporated in the sensing platform to provide the distinguishable readout for target quantification.

Sensing with SERS-encoded nanoparticles (SEPs) is by far the most common indirect SERS approach [2,4–6]. SEPs combine at least the three following elements: a plasmonic enhancer, a SERS label bound to the metallic surface yielding an intense and unique SERS fingerprint, and surface molecules affording convenient physicochemical properties (e.g., antibodies). In the SEP-based detection scheme, the identification and quantification of the target protein are correlated to the absolute intensity of the detected SERS signal [7–9].

Alternatively, a rather advantageous sensing design is to integrate within the same surface element both the bioreceptor and the SERS label (Fig. 19.1A). In the resulting "chemoreceptor" molecule, the SERS label moiety firmly binds the metallic surface acting as a linker for the bioreceptor moiety which remains available for the selective interaction with the target protein. In this scenario, the SERS label moiety performs as a SERS transducer converting the receptor/target complex formation into measurable alterations of the label spectral profile (Fig. 19.1B) which can be quantitatively correlated with the number of binding events (Fig. 19.1C). Remarkably, in such sensing scheme, the chemoreceptor acts also as an internal standard, thereby enabling higher reproducibility as

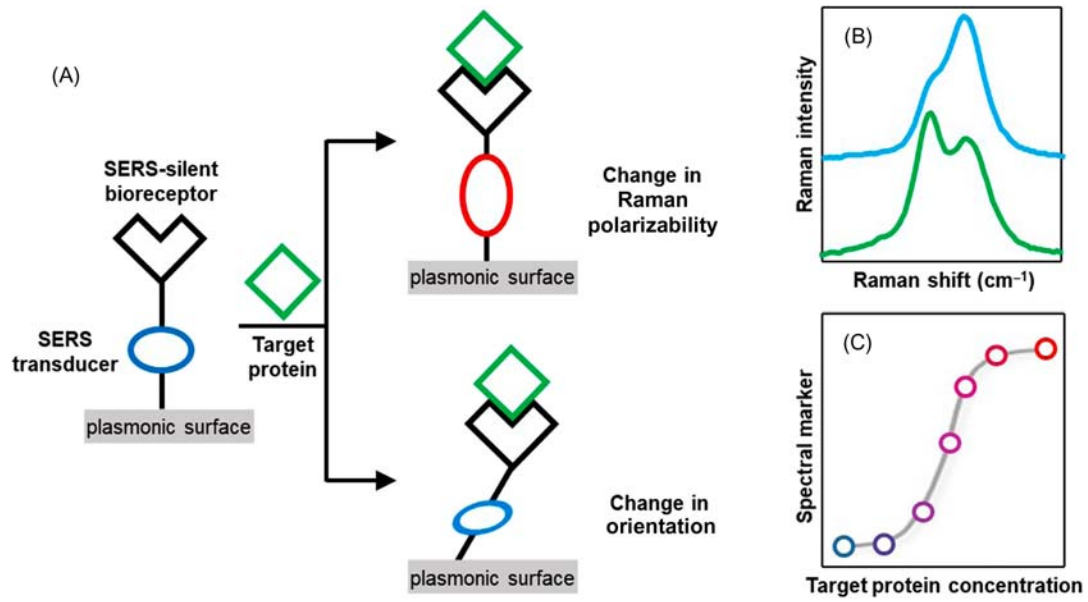


Figure 19.1
Schematic outline of the indirect surface-enhanced Raman scattering (SERS) chemosensing of proteins. (A) Upon binding of the target protein to the SERS-silent bioreceptor, the SERS transducer that connects the ligand moiety to the metallic surface undergoes structural and/or orientation changes that result (B) in spectral alteration of its SERS profile. (C) The extent of these spectral changes can be quantitatively correlated with the number of binding events.

compared to methods that solely rely on the acquisition of absolute intensities for quantification.

These spectral alterations can be mostly classified as the combination of two separate contributions, whose relative extent is case specific and depends on multiple experimental aspects such as the intrinsic physicochemical properties of the chemoreceptor, the surface coverage and surface lateral packing, the measuring conditions, the properties of the enhancing substrate, etc.

In the first case, the formation of the surface complex alters the molecular electronic structure of the SERS transducer resulting in the modification of the Raman polarizability of its vibrational modes. In this scenario, the SERS-active linker can be assimilated, before and after the target binding, as two different molecular entities with their correspondent distinct vibrational fingerprint. On the other hand, the surface complex formation can also simply determine a mere reorientation of the SERS transducer with respect to the metallic surface. In this second case, the observed spectral alterations arise from the differences in the magnitude of the normal (E_n) and tangential (E_t) electric field components experienced

by the SERS transducer at the metallic surface, which can be interpreted according to the "surface selection rules" [10]. In fact, under typical experimental conditions, the normal field component largely exceeds the tangential one; thus those vibrational modes which possess a Raman polarizability with a strong component along the perpendicular axis will be preferably enhanced as compared to other modes that are mostly excited by the tangential component [10]. Ratiometric measurements of the intensity between these two classes of vibrational modes can then inform about the reorientation of the molecule over the plasmonic surface.

Among the different examples reported in the literature, SERS transducers based on mercaptobenzene derivatives are the most widely used [11–20]. The –SH functional group forms strong covalent thiolated bonds with the metallic surface, while the phenyl ring yields strong SERS signals with well-defined spectral features that are particularly sensitive to charge distribution and surface orientation [21]. The presence of additional functional groups (e.g., amine, carboxylic, amide, etc.), typically in *para* position with respect to the thiol group, also provides reactive sites for conjugating the biomolecular receptor.

On the other hand, the plasmonic substrates (Ag, mostly, or Au) commonly employed in these applications can be broadly classified into three categories: nanoparticle films on 2D substrates obtained via *in situ* reduction of silver salts or by deposition of preformed colloidal particles [14,15,18,19,22], nanoparticle films on colloidal microparticles assembled via layer-by-layer protocols [11,20], and nanostructured surfaces prepared via lithographic techniques or glancing angle deposition [12,13,16,17] (Fig. 19.2). Regardless of the class of plasmonic materials, the goal in the fabrication procedures is to maximize the enhancing efficiency by generating a dense collection of electromagnetic hot spots (e.g., closely spaced particle-to-particle junctions) while retaining homogenous SERS responses. A thorough discussion on the concepts, properties, and nanofabrication methods for plasmonic materials for SERS goes beyond the scope of this chapter and thus we refer the interested readers to [3,23–27] and references therein.

19.2 Peptide-based chemosensing

Our group and collaborators [19,20] exploited organic chemoreceptors relying on synthetically designed peptide sequences, with high affinity and selectivity for a target protein, which were orthogonally modified with a 4-mercapto benzoyl unit (MB) at a terminal amino acid residue.

In our first study, a c-Fos peptide sequence (57 aa) was used as the bioreceptor for the oncoprotein c-Jun [19]. c-Jun is a member of the basic zipper family of transcription factors and its overexpression has been linked to multiple carcinogenic processes [28]. c-Jun is known to dimerize via a leucine zipper domain with the native protein partner c-Fos,

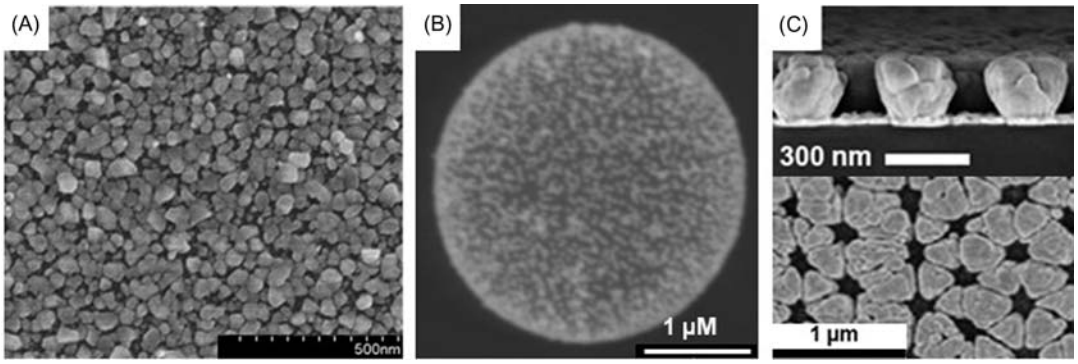


Figure 19.2
Scanning electron microscopy (SEM) images of plasmonic substrates employed in surface-enhanced Raman scattering (SERS) chemosensing of proteins. (A) Ag film obtained by *in situ* reduction of Ag nanoparticles on a glass slide surface modified with 3-mercaptopropyl-triethoxysilane. (B) Polystyrene microbead coated via layer-by-layer by a dense film of gold nanoparticles. (C) Lateral and top views of a nanosculptured thin Ag film on Si substrates prepared by the glancing angle deposition, a technique for fabricating porous thin films with engineered structures technique. Source*: Adapted with permission from (A) B. Tang, et al., Ultrasensitive, multiplex Raman frequency shift immunoassay of liver cancer biomarkers in physiological media, ACS Nano 10 (1) (2016) 871–879. Copyright 2016, American Chemical Society; (B) L. Guerrini, et al., SERS detection of amyloid oligomers on metallorganic-decorated plasmonic beads, ACS Appl. Mater. Interfaces 7 (18) (2015) 9420–9428. Copyright 2015, American Chemical Society; and (C) S.K. Srivastava, et al., SERS biosensor using metallic nano-sculptured thin films for the detection of endocrine disrupting compound biomarker vitellogenin, Small 10 (17) (2014) 3579–3587. Copyright 2014, Wiley-VHC.*

an event that results in the folding of the loosely structured monomeric proteins into a relatively rigid α-helix conformation (Fig. 19.1B). A monolayer of the peptide chemoreceptor (MB-Fos) was deposited on a silver island film prepared by electrostatic deposition of silver colloids on a glass slide, yielding intense SERS features characteristic of the mercaptophenyl group (Fig. 19.3B) [29]. Overall, the SERS spectrum is dominated by in-plane modes such as those centered at 1585 cm^{-1} (C = C stretching), 1075 cm^{-1}, [ring breathing mode coupled with ν(C–S)], 1022 cm^{-1} (CCH deformation), 998 cm^{-1} (ring breathing mode), and 417 cm^{-1} (C–S stretching coupled to in-plane ring bending). Notably, the vibrational modes associated to the 1075 and 1585 cm^{-1} bands are known to display the largest interfacial contribution to the ring polarizability change in the MB–metal surface complex [30]. On the other hand, out-of-plane modes, such as the ring deformation mode at 756 cm^{-1}, are barely detectable or completely absent from the SERS spectra.

Exposure to c-Jun solutions at increasing concentrations generates a broad set of spectral changes. Most prominently, the 1075 cm^{-1} band undergoes a 3 cm^{-1} blue-shift (Fig. 19.3C) while we detect a marked change in relative intensity between the shoulder at

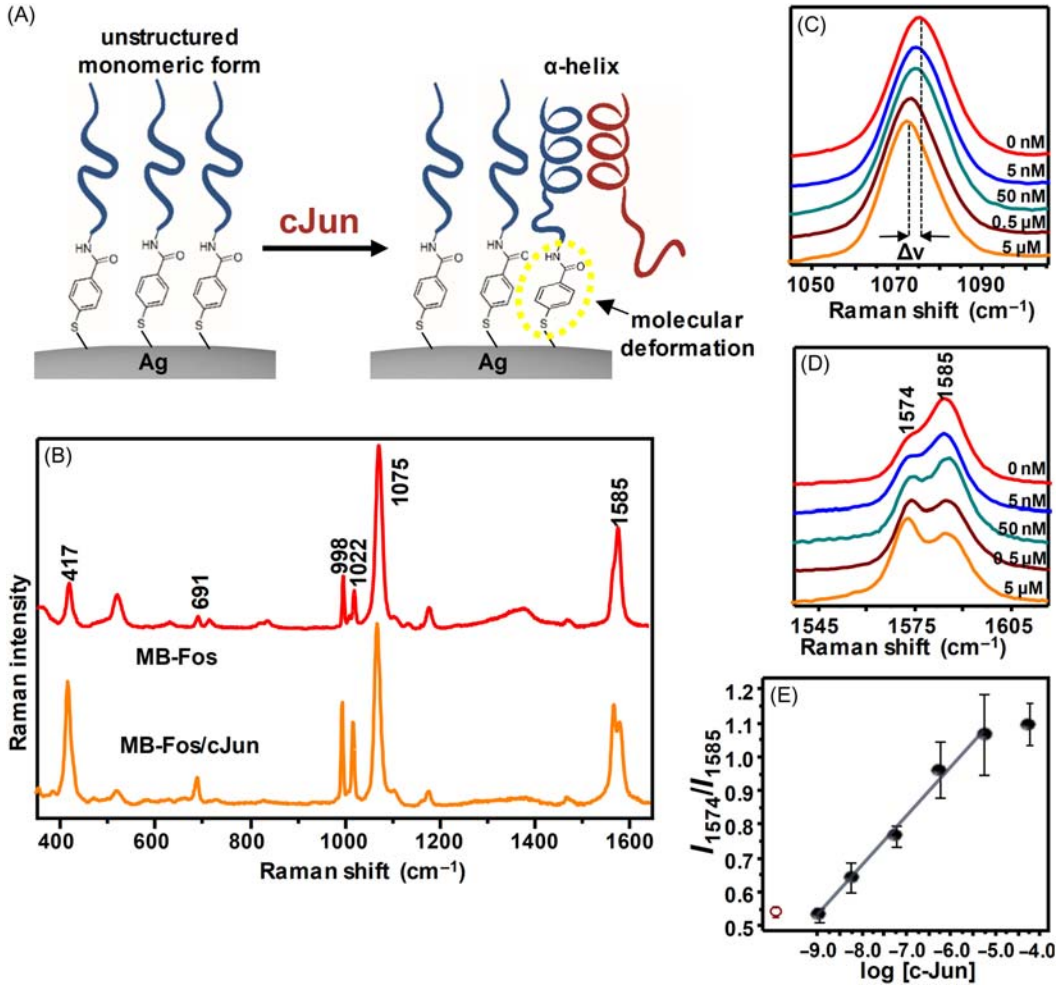


Figure 19.3

(A) Outline of the MB-Fos/c-Jun dimerization on the metal surface and the resulting structural deformation of the SERS transducer. (B) SERS spectra of MB-Fos and MB-Fos/c-Jun onto the silver film. (C, D) Details of the 1000–1100 and 1540–1620 cm^{-1} spectral regions of the SERS spectra of MB-Fos upon exposure to c-Jun solution in 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer at increasing concentrations. (E) Intensity ratio I_{1574}/I_{1585} as a function of c-Jun concentration (logarithmic scale). Source*: Adapted with permission from L. Guerrini, et al., Highly sensitive SERS quantification of the oncogenic protein c-Jun in cellular extracts, J. Am. Chem. Soc. 135 (28) (2013) 10314–10317. Copyright 2012, American Chemical Society.*

approximately 1574 cm^{-1}, assigned to the nontotally symmetric $\nu(C=C)$ vibration, with respect to the totally symmetric $\nu(C=C)$ mode at 1585 cm^{-1} (Fig. 19.3D) [31]. Such redistribution of the scattering cross sections among the C = C stretching vibrations has

been ascribed to the perturbation of the benzene ring symmetry in the surface complex [31]. Here, the SERS transducer appears to act preferably as a molecular spring that responds to the mechanical stress imposed by the coordination of the large oncoprotein. The variation of the corresponding ratiometric peak intensities, I_{1574}/I_{1585}, has been quantitatively correlated with the c-Jun content (Fig. 19.3E), reporting a linear response over a broad range of oncoprotein concentration in cell lysates and a 5 nM detection limit.

Subsequently, we extended this approach for the quantification of the oncoprotein c-MYC in human samples [20]. In this second work, the MB unit has been conjugated to the C-terminus of the peptide sequence H1 that has been reported to display high affinity and selectivity for c-MYC binding. The corresponding MB-H1 chemoreceptor was also attached to the silver film but, in this case, colloidal AgNPs were assembled onto silica beads (8 μm diameter) via layer-by-layer protocol rather than on planar surfaces. These hybrid plasmonic materials combine the advantages of highly colloidally stable particles and plasmonic films, while their micrometric size enables their rapid localization under a conventional microscope. In this manner, the number of microparticle sensors in the sample can be conveniently reduced to very low contents without affecting the SERS performance. With the goal of maximizing the platform sensitivity, MB-H1 surface density was progressively decreased far below the monolayer coverage to reduce the number of binding sites per bead while retaining intense SERS signal from individual microparticles. In this case, the coordination of the large oncoprotein (57 kDa) to the low-molecular-weight chemoreceptor determines major changes in the MB spectral profile which have been ascribed to the striking reorientation of the phenyl ring toward a flatter orientation with respect to the surface (Fig. 19.4). Most notably, this is clearly highlighted by the dramatic rising of the out-of-plane ring deformation at 756 cm^{-1}. Thus the ratiometric peak intensity of the

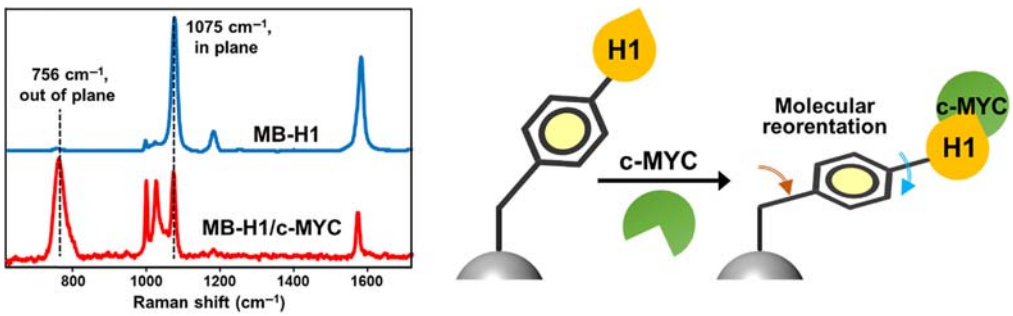


Figure 19.4

SERS spectra of MB-H1 before and after the complexation with c-MYC. A schematic of the MB reorientation upon the biorecognition is also illustrated. Source*: Adapted with permission from E. Pazos, et al., Surface-enhanced Raman scattering surface selection rules for the proteomic liquid biopsy in real samples: efficient detection of the oncoprotein c-MYC, J. Am. Chem. Soc. 138 (43) (2016) 14206−14209. Copyright 2016, American Chemical Society.*

756 cm^{-1} band vs 1075 cm^{-1} (in-plane ring breathing) was selected as a spectral marker and showed to scale linearly with the c-MYC concentration. The efficiency of the sensor was demonstrated in the quantification of c-MYC in cell lysates from both tumor cell lines and human samples. ELISA was performed to validate the results, demonstrating similar detection limits in the 1 pM level but with major advantages in terms of cost and rapidity (SERS results are obtained within minutes compared to the 4−5 hours preparations required for ELISA).

Every chemoreceptor/target protein interaction at the metal surface is defined by a complex set of multiple factors that have to be tackled on a case-by-case basis. However, we can provide a first interpretation of the disparity of the SERS transducer behavior in the two studies based primarily on the different surface coverage. In the first case, a compact monolayer of the chemoreceptor is expected to provide a high packing density which reduces the freedom of movement of the bioreceptor element favoring, in turn, the retention of a radial orientation over the surface upon protein binding. Conversely, in the latter detection scheme, the absence of sterical hindrance for largely spaced chemoreceptors allows for the reorientation of the surface complex as a primary response to the molecular stress imposed by the c-MYC binding.

Construction of chemoreceptors based on short peptides rather than more common antibodies offer superior potential in terms of structural fidelity, stability, a controlled organization at the metal surface, and cost. Furthermore, phage display libraries with billions of peptides are available for screening to select amino acid sequences with a specific affinity toward a multitude of targets (e.g., nucleic acids, proteins, polysaccharides, and cells) [32]. With such an approach, Zhu et al. [14] recently selected two short peptides for the discrimination of amyloid-β proteins, Aβ42, in the monomeric and fiber form in buffered solution. 5,5′-Dithiobis succinimidyl-2-nitrobenzoate (DNSB) was used as a SERS transducer. A dense layer of DNSB was first deposited onto the silver film and subsequently modified by covalent attachment with peptides P1 or P2, yielding the corresponding DNSB-P1 and DNSB-P2 surface chemoreceptors (Fig. 19.5A). The sensing platform was then incubated with a solution of Aβ42 monomers and the SERS spectra monitored over time. In particular, the shift of the intense NO_2 symmetric stretching band at 1334 cm^{-1} was selected as the spectral marker of the Aβ42 binding. In the case of DNSB-P1, the nitro band redshifts up to approximately 0.7 cm^{-1} after 60 minutes of incubation, before progressively returning to its initial frequency position at 300 minutes time (Fig. 19.5B). For DNSB-P2, the frequency shift occurs much slower, with a maximum frequency shift (approximately 1.1 cm^{-1}) acquired for 5-hour incubation time (Fig. 19.5C). The authors interpreted the results as depicted in Fig. 19.5D. At $t = 0$, only Aβ42 monomers are present in the solution and diffuse to the metallic surface where they preferably interact with P1 via $\pi-\pi$ stacking. As the captured Aβ42 monomers progressively fold and aggregate into small oligomeric species, they decrease their affinity for P1 until they are

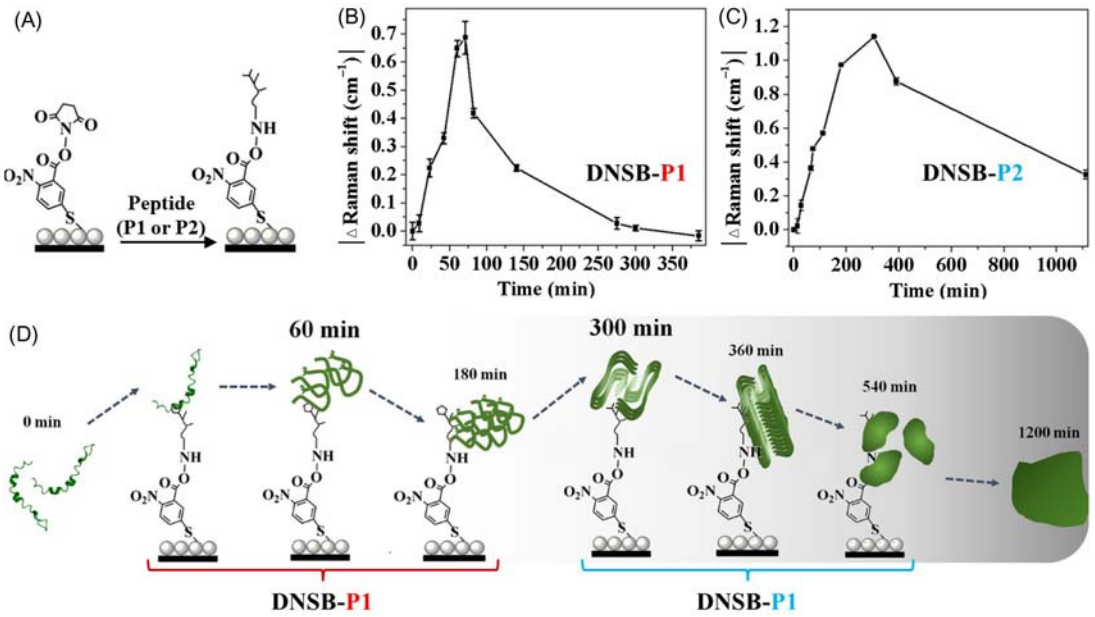


Figure 19.5

(A) Silver nanoparticle film is functionalized with a monolayer of 5,5′-dithiobis succinimidyl-2-nitrobenzoate (DSNB). Subsequently, peptides P1 or P2 are covalently attached to the SERS transducer. (B, C) Raman shifts of the nitro stretching band at ca. 1334 cm^{-1} as a function of time upon incubation with Aβ42 for DSNB-P1 and DSNB-P2, respectively. (D) Proposed mechanism of the chemoreceptor response to Aβ protein conformational changes in solution. Source: *Adapted with permission from W. Zhu, et al., In situ monitoring the aggregation dynamics of amyloid-β protein Aβ42 in physiological media via a Raman-based frequency shift method, ACS Appl. Bio Mater. 1 (3) (2018) 814−824. Copyright 2018, American Chemical Society.*

finally released in the solution. After approximately 300 minutes, larger mature fibers begin to form in solution, becoming then available to specifically interact with P2. Further incubation beyond 300 minutes results in the formation of a network structure of connected fibers which finally detach from P2 into the solution. In addition to the recognition of different Aβ42 forms, this work demonstrates the capability of indirect SERS chemosensing to monitor the corresponding dynamic aggregation process.

19.3 Antibody-based chemosensing

The use of antibody (Ab) receptors conjugated with SERS transducers for the detection of protein was first devised by Olivo's group [12,13,16,33]. In their first seminal work [13], a self-assembled monolayer of 4-aminothiophenol (ATP) was generated on an Au/Ag bimetallic nanopatterned surface obtained via deep-UV lithography (Fig. 19.6A). The anti-influenza-H1 Ab was then covalently linked to the amino group of ATP surface

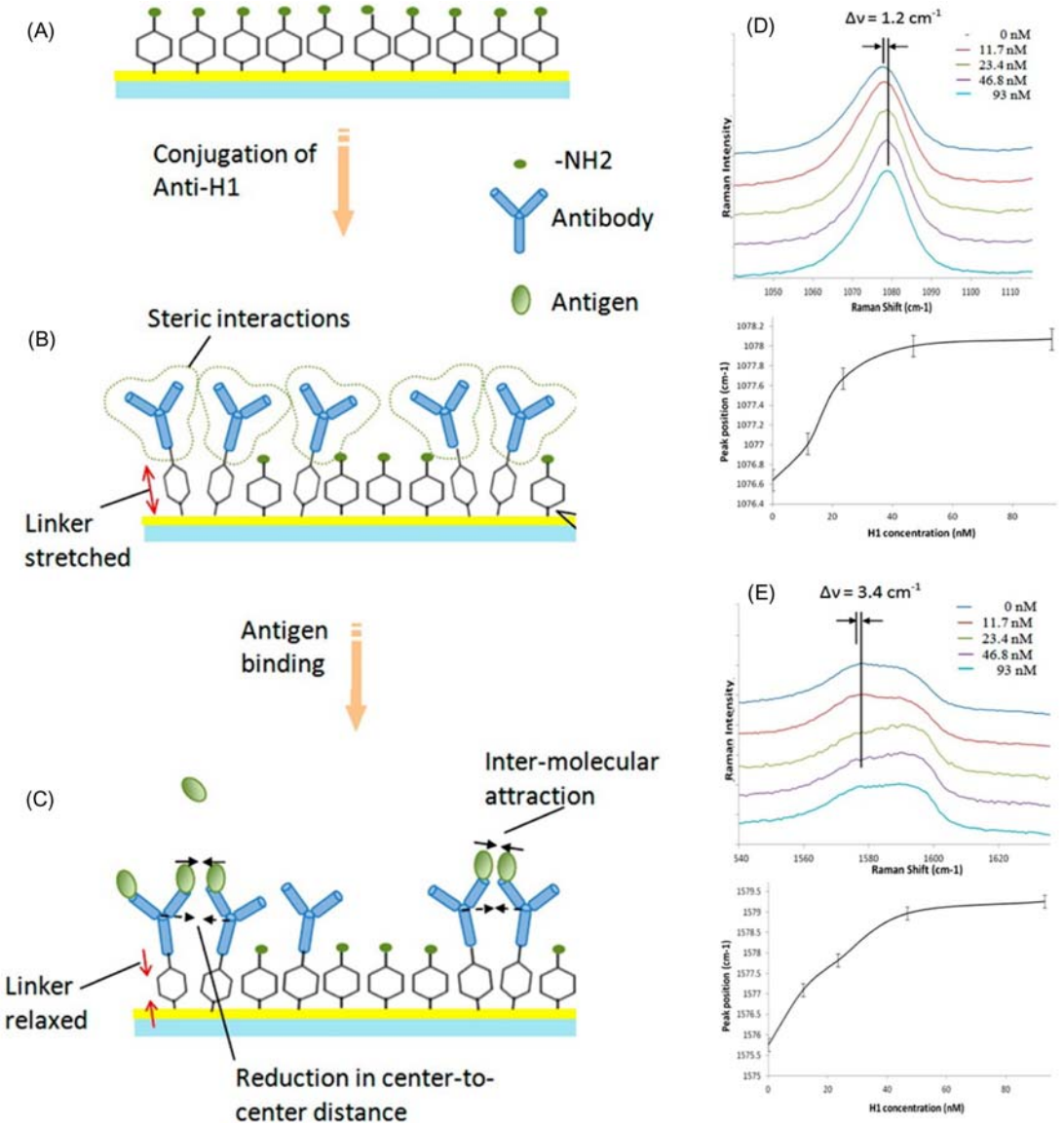


Figure 19.6

(A) A self-assembled monolayer of 4-aminothiophenol (ATP) is formed on the Au/Ag bimetallic nanostructured surface. (B) Conjugation of the antibody anti-H1 via 1-ethyl-3-(3 dimethylaminopropyl)carbodiimide (EDC)/*N*-hydroxysuccinimide coupling leading to mechanical stretching of the surface-enhanced Raman scattering (SERS) transducer. (C) Binding of the H1 antigen promotes the reduction of the center-to-center distance between surface Abs which, in turn, allows for the mechanical relaxation of ATP. (D, E) Raman shifts of the 1080 and 1580 cm^{-1} bands in response to the antigen binding at increasing H1 concentration. Source*: Adapted with permission from K.W. Kho, et al., Frequency shifts in SERS for biosensing, ACS Nano 6 (6) (2012) 4892–4902. Copyright 2012, American Chemical Society.*

molecules upon activation of the carboxyl terminal of the crystal segment via ethyl-3-(3-dimethylaminopropyl)carbodiimide/*N*-hydroxysucinimide (EDC/NHS) coupling (Fig. 19.6B). As a result, the Ab preferably adopts a vertical position which favors the surface complexation of the target H1 antigen. The peak positions of the intense 1580 and 1080 cm^{-1} bands, corresponding to ν(C = C) modes and ring breathing coupled with ν(C−S) modes, respectively, were selected as spectral markers to quantify the H1 binding in buffered solution (Fig. 19.6D and E). Differently to what was observed for c-Jun detection (see Fig. 19.3C and D), a redshift is observed for both peak positions when the substrate is exposed to increasing concentrations of H1. A mechanical model has been proposed to explain the SERS shifts. Here, ATP undergoes a tensile deformation upon anti-H1 conjugation due to the extensive intermolecular repulsion between the sterically bulky antibodies (Fig. 19.6B). Such steric repulsion is subsequently decreased via H1 binding, due to attractive hydrophobic interactions between the surface-bound antigens, leading to structural relaxation of the SERS transducer (Fig. 19.6C), as revealed by the redshift of the marker bands. A detection limit of 2.2 nM has been reported.

Such a sensing approach was also extended to the simultaneous detection of two target proteins (p53 and epidermal growth factor receptor) in buffer solution by using ATP and 6-mercaptopurine as distinct SERS transducers [13,16]. Subsequently, the sensitivity for H1 quantification was improved to the 10 pM level by boosting the enhancing performance of the plasmonic substrate [12]. Further improvements to the biosensing assay were made by Tang et al. [18] who demonstrated the detection of two liver cancer biomarkers α-fetoprotein (AFP) and glypican-3 down to subpicomolar levels in a model physiological medium (fetal calf serum).

In recent work, He and coworkers [15] extended the application of this frequency shift−based chemosensing by integrating it with SEPs for indirect SERS determination of the relative content of L3, one of the three glycoforms of the α-fetoprotein (AFP-L3), over the total AFPs. The corresponding ratio (AFP-L3%) is a tumor marker for screening chronic liver disease patients for hepatocellular carcinoma (HCC) [34]. In this dual-mode sensor, silver films were functionalized with a compact layer of 4-mercaptobenzoic acid-antibody (MBA-Ab) chemoreceptors, where the Ab consisted of anti-AFP which selectively binds AFPs regardless of their isoform (Fig. 19.7A). Once again, the frequency shift of the MB spectral markers (in this case, the 1075 cm^{-1} band ascribed to ring breathing coupled to C−S stretching) were used to quantify the total AFP content in the sample at subnanomolar level (Fig. 19.7B). Subsequently, SEPs comprising a gold core, DSNB as a SERS label, and anti-AFP-L3 as a biorecognition element (Fig. 19.7, inset) were captured at the metal surface in a sandwich immunoassay, as revealed by the appearance of DSNB bands in the SERS spectra, such as the intense symmetric NO_2 stretching band at 1331 cm^{-1} (Fig. 19.7C). The peak height ratio between the nitro band of DSNB and the 1071 cm^{-1} feature of MBA as an internal standard (I_{1331}/I_{1071}) was selected as the spectral marker for AFP-L3 quantification (Fig. 19.7D). Clinical validation of the sensor was performed by analyzing 32 real samples from the serum

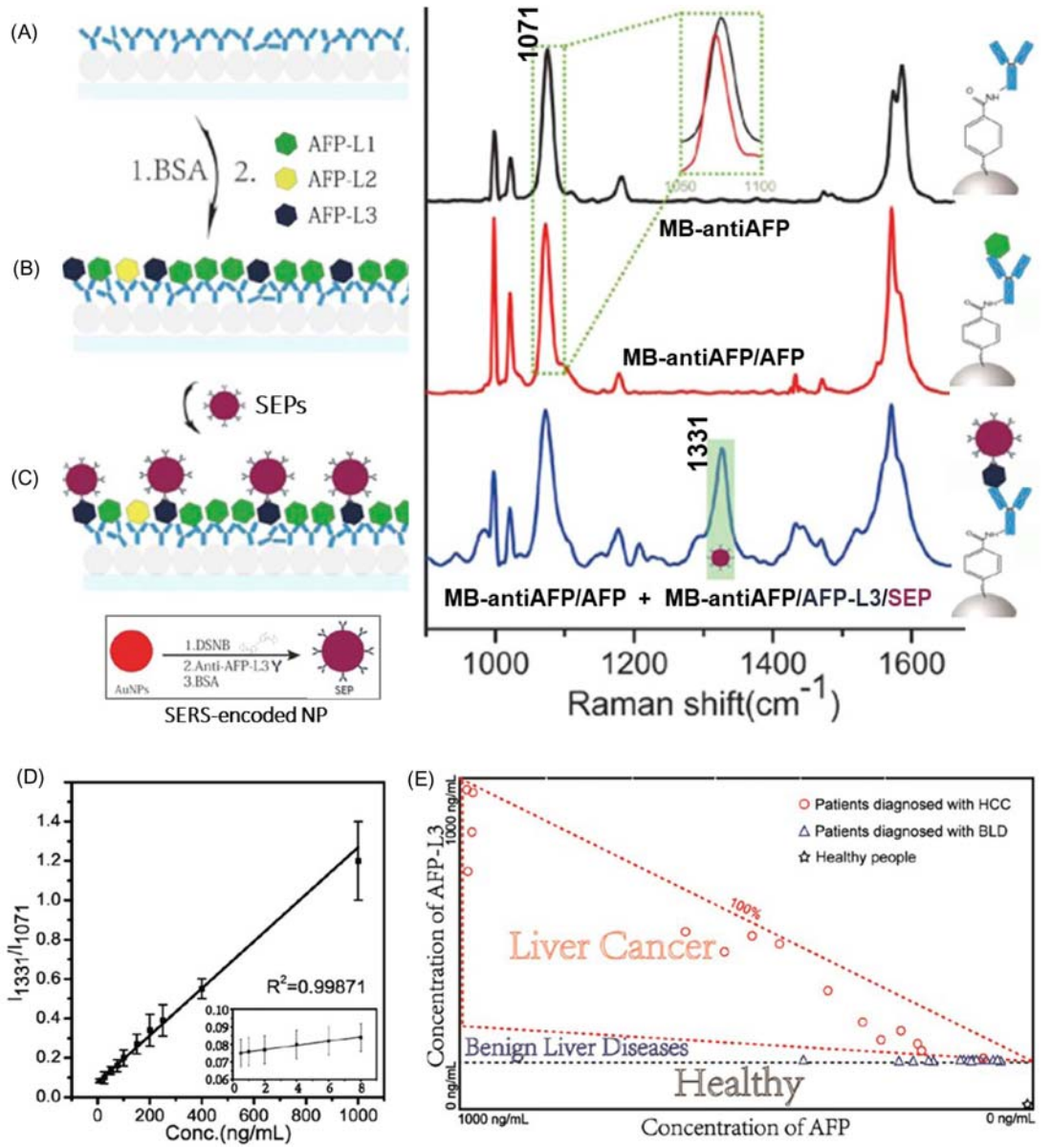


Figure 19.7

(A) 4-Mercaptobenzoic acid (MBA) is self-assembled onto the surface of the silver film. The antibody α-fetoprotein (anti-AFP) is covalently bound to the MBA transducer via ethyl-3-(3-dimethylaminopropyl)carbodiimide/*N*-hydroxysucinimide (EDC/NHS) coupling. (B) Bovine serum albumin (BSA) is used as a blocking agent to prevent nonspecific binding of antigens to the silver surface. After removing unbound BSA, the chip is immersed into a AFP solution. Frequency shifts of MBA marker bands are detected in the SERS spectra and quantitatively correlated with the

(*Continued*)

of healthy donors and diagnosed patients (Fig. 19.7E), demonstrating the viability of the proposed SERS-based sensor for the early diagnosis of HCC.

19.4 Bioreceptor-free chemosensing

A few examples of bioreceptor-free strategies have been also described in the literature. For instance, our group functionalized gold nanoparticles assembled onto polystyrene beads (3 μm diameter) with an MBA-based metal−organic chemoreceptor. In this case, the carboxyl group of MBA bound to the plasmonic surface was coordinated with Al^{3+}, in place of the traditional peptide or Ab element [11]. Aluminum ions have been reported to exert a notable complexation activity against β-amyloid proteins. In the presence of the 91-residue small protein HypF-N, a model system for amyloid-like aggregates, the SERS profile of the MB-Al^{3+} chemoreceptor undergoes a set of unique spectral changes, enabling the identification of the oligomeric species in buffer solution at the submicromolar level.

More recently, He and coworkers [22] designed a chemosensing approach where perylenetetra carboxylic acid (PTCA) anchored onto silver film and activated by ECD/NHS is exposed to different solutions of ten proteins with different molecular weight (bovine serum albumin, pepsin, egg white albumin, trypsin, hemoglobin, horseradish peroxidase, myoglobin, cytochrome C, glucose oxidase, and lysozyme). The covalent binding of the protein with PTCA results in a set of spectral changes of the SERS active molecule that are protein-dependent and have been ascribed to the perturbation of the Raman polarizability. All proteins were successfully discriminated with the aid of hierarchical cluster analysis. The authors suggested that the diverse structural response of the SERS linker to the protein conjugation could spring from the diverse chemical environment of the protein binding sites.

19.5 Conclusions

SERS-based technologies provide a novel emerging chance for the quantification and characterization of proteins, speeding up their structural and functional exploration, while developing novel non-labeled medical diagnosis technologies. Nanostructures such as gold

◀ total AFP content in the 0.5−1000 ng/mL range. (C) SERS-encoded nanoparticles (SEPs) selectively bind AFP-L3 antigens onto the surface, yielding SERS signal from the SERS label 5,5′-dithiobis succinimidyl-2-nitrobenzoate (DSNB) (see the intense band at 1331 cm^{-1}, ascribed to nitro stretching) that overlaps the SERS from MBA transducer. Inset: outline of the SEP major features and components. (D) Intensity ratio (I_{1331}/I_{1071}) as a function of AFP-L3 concentration in the 8−1000 ng/mL range. (E) Clinical sample test results from real human samples compared with the diagnosis results. Source: *Adapted with permission from H. Ma, et al., Multiplex immunochips for high-accuracy detection of AFP-L3% based on surface-enhanced Raman scattering: implications for early liver cancer diagnosis, Anal. Chem. 89 (17) (2017) 8877−8883. Copyright 2017, American Chemical Society.*

or silver colloids, 2D nanosheets, etc. can markedly enhance the Raman scattering signal of proteins, down to the single molecule detection level in some cases. The potential molecular mechanism is associated with LSPRs and the formation of hotspots. The potential of indirect SERS detection in applications such as SERS quantification of protein biomarkers has been positively explored. To date, much of the existing literature still demonstrates concepts in lab samples with few reports describing the detection of proteins in biological conditions. However, although the single and multiplex detection of relevant proteins in real biofluids and solid samples (tissues) is still challenging, the field is advancing at a high velocity, with many new discoveries occurring almost on a daily basis. In summary, although SERS sensing for protein quantification and characterization faces many challenges, SERS combined with nanotechnology has exhibited attractive technological prospects. We foresee that this approach will dramatically advance our ability to understand the chemical biology of proteins while developing novel series of SERS-based technology, which will eventually help to promote the application of SERS in other fields.

Acknowledgments

This work was funded by the Spanish Ministerio de Economía, Industria y Competitividad (CTQ2017-88648R, RYC-2016-20331), the Generalitat de Cataluña (2017SGR883), the Universitat Rovira i Virgili (2018PFRURV-B2-02), and the Universitat Rovira i Virgili and Banco Santander (2017EXIT-08).

References

[1] B. Pelaz, et al., Diverse applications of nanomedicine, ACS Nano 11 (3) (2017) 2313−2381.
[2] L.A. Lane, X.M. Qian, S.M. Nie, SERS nanoparticles in medicine: from label-free detection to spectroscopic tagging, Chem. Rev. 115 (19) (2015) 10489−10529.
[3] S. Schlücker, Surface-enhanced Raman spectroscopy: concepts and chemical applications, Angew. Chem. Int. Ed. 53 (19) (2014) 4756−4795.
[4] L. Guerrini, et al., Cancer characterization and diagnosis with SERS-encoded particles, Cancer Nanotechnol. 8 (1) (2017) 5.
[5] Y.Q. Wang, B. Yan, L.X. Chen, SERS tags: novel optical nanoprobes for bioanalysis, Chem. Rev. 113 (3) (2013) 1391−1428.
[6] L. Fabris, SERS tags: the next promising tool for personalized cancer detection? ChemNanoMat 2 (4) (2016) 249−258.
[7] L. Guerrini, R.A. Alvarez-Puebla, Surface-enhanced Raman spectroscopy in cancer diagnosis, prognosis and monitoring, Cancers 11 (6) (2019) 748.
[8] B. Shan, et al., Novel SERS labels: rational design, functional integration and biomedical applications, Coord. Chem. Rev. 371 (2018) 11−37.
[9] N. Feliu, et al., SERS quantification and characterization of proteins and other biomolecules, Langmuir 33 (38) (2017) 9711−9730.
[10] M. Moskovits, J.S. Suh, Surface selection rules for surface-enhanced Raman spectroscopy: calculations and application to the surface-enhanced Raman spectrum of phthalazine on silver, J. Phys. Chem. 88 (23) (1984) 5526−5530.

[11] L. Guerrini, et al., SERS detection of amyloid oligomers on metallorganic-decorated plasmonic beads, ACS Appl. Mater. Interfaces 7 (18) (2015) 9420−9428.
[12] J. Perumal, et al., Design and fabrication of random silver films as substrate for SERS based nano-stress sensing of proteins, RSC Adv. 4 (25) (2014) 12995−13000.
[13] K.W. Kho, et al., Frequency shifts in SERS for biosensing, ACS Nano 6 (6) (2012) 4892−4902.
[14] W. Zhu, et al., In situ monitoring the aggregation dynamics of amyloid-β protein Aβ42 in physiological media via a Raman-based frequency shift method, ACS Appl. Bio Mater. 1 (3) (2018) 814−824.
[15] H. Ma, et al., Multiplex immunochips for high-accuracy detection of AFP-L3% based on surface-enhanced Raman scattering: implications for early liver cancer diagnosis, Anal. Chem. 89 (17) (2017) 8877−8883.
[16] P. Owens, et al., Sensing of p53 and EGFR biomarkers using high efficiency SERS substrates, Biosensors 5 (4) (2015) 664−677.
[17] S.K. Srivastava, et al., SERS biosensor using metallic nano-sculptured thin films for the detection of endocrine disrupting compound biomarker vitellogenin, Small 10 (17) (2014) 3579−3587.
[18] B. Tang, et al., Ultrasensitive, multiplex Raman frequency shift immunoassay of liver cancer biomarkers in physiological media, ACS Nano 10 (1) (2016) 871−879.
[19] L. Guerrini, et al., Highly sensitive SERS quantification of the oncogenic protein c-Jun in cellular extracts, J. Am. Chem. Soc. 135 (28) (2013) 10314−10317.
[20] E. Pazos, et al., Surface-enhanced Raman scattering surface selection rules for the proteomic liquid biopsy in real samples: efficient detection of the oncoprotein c-MYC, J. Am. Chem. Soc. 138 (43) (2016) 14206−14209.
[21] F. Sun, et al., Multi-functional, thiophenol-based surface chemistry for surface-enhanced Raman spectroscopy, Chem. Commun. 53 (33) (2017) 4550−4561.
[22] H. Ma, et al., Antibody-free discrimination of protein biomarkers in human serum based on surface-enhanced Raman spectroscopy, Anal. Chem. 90 (21) (2018) 12342−12346.
[23] H. Ko, S. Singamaneni, V.V. Tsukruk, Nanostructured surfaces and assemblies as SERS media, Small 4 (10) (2008) 1576−1599.
[24] L. Guerrini, D. Graham, Molecularly-mediated assemblies of plasmonic nanoparticles for surface-enhanced Raman spectroscopy applications, Chem. Soc. Rev. 41 (21) (2012) 7085−7107.
[25] A.X. Wang, X.M. Kong, Review of recent progress of plasmonic materials and nano-structures for surface-enhanced Raman scattering, Materials 8 (6) (2015) 3024−3052.
[26] M.F. Cardinal, et al., Expanding applications of SERS through versatile nanomaterials engineering, Chem. Soc. Rev. 46 (13) (2017) 3886−3903.
[27] H.K. Lee, et al., Designing surface-enhanced Raman scattering (SERS) platforms beyond hotspot engineering: emerging opportunities in analyte manipulations and hybrid materials, Chem. Soc. Rev. 48 (3) (2019) 731−756.
[28] P.K. Vogt, Jun, the oncoprotein, Oncogene 20 (19) (2001) 2365−2377.
[29] T.H. Joo, M.S. Kim, K. Kim, Surface-enhanced Raman scattering of benzenethiol in silver sol, J. Raman Spectrosc. 18 (1) (1987) 57−60.
[30] A.T. Zayak, et al., Chemical Raman enhancement of organic adsorbates on metal surfaces, Phys. Rev. Lett. 106 (8) (2011).
[31] S.K. Saikin, et al., On the chemical bonding effects in the Raman response: Benzenethiol adsorbed on silver clusters, Phys. Chem. Chem. Phys 11 (41) (2009) 9401−9411.
[32] J. Pande, M.M. Szewczyk, A.K. Grover, Phage display: concept, innovations, applications and future, Biotechnol. Adv. 28 (6) (2010) 849−858.
[33] K.V. Kong, et al., A rapid and label-free SERS detection method for biomarkers in clinical biofluids, Small 10 (24) (2014) 5030−5034.
[34] F.A. Durazo, et al., Des-gamma-carboxyprothrombin, alpha-fetoprotein and AFP-L3 in patients with chronic hepatitis, cirrhosis and hepatocellular carcinoma, J. Gastroenterol. Hepatol. 23 (10) (2008) 1541−1548.

This page intentionally left blank

Index

Note: Page numbers followed by "*f*" and "*t*" refer to figures and tables, respectively.

D

E

F

G

H

I

J

K

L

M

N

Q

R

S

T

VIBRATIONAL SPECTROSCOPY IN PROTEIN RESEARCH

From Purified Proteins to Aggregates and Assemblies

Edited by

Yukihiro Ozaki, Professor Emeritus, School of Science and Technology, Kwansei Gakuin University, Hyogo, Japan

Malgorzata Baranska, Professor, Faculty of Chemistry & Jagiellonian Centre for Experimental Therapeutics (JCET), Jagiellonian University, Krakow, Poland

Igor K. Lednev, Professor, Department of Chemistry, University at Albany, State University of New York, Albany, NY, United States

Bayden R. Wood, Professor, Director, Centre for Biospectroscopy, School of Chemistry, Monash University, Clayton, VIC, Australia

In recent years, a range of new, state-of-the-art vibrational spectroscopic techniques have been developed, including Raman imaging, Raman optical activity (ROA), SERS, TERS, nonlinear Raman spectroscopy, AFM-IR, cascade laser IR spectroscopy, IR imaging, NIR imaging, Terahertz spectroscopy, and time-resolved vibrational spectroscopy. In addition, spectral analysis methods, such as chemometrics and quantum chemical calculation, are now seeing wider use by researchers across the biomedical and life sciences.

Vibrational Spectroscopy in Protein Research: From Purified Proteins to Aggregates and Assemblies offers a thorough discussion of vibrational spectroscopy in protein research, providing researchers with clear, practical guidance on methods employed, areas of application, and modes of analysis. With chapter contributions from international leaders in the field, the book addresses basic principles of vibrational spectroscopy in protein research, instrumentation and technologies available, sampling methods, quantitative analysis, origin of group frequencies, and qualitative interpretation. In addition to discussing vibrational spectroscopy for analysis of purified proteins, chapter authors also examine use in studying complex protein systems, including protein aggregates, fibrous proteins, membrane proteins, and protein assemblies.

Emphasis throughout is placed on applications in human tissue, cell development, and disease analysis, with chapters dedicated to studies of molecular changes that occur during disease progression, as well as identifying changes in tissues and cells in disease studies.

Key Features

- Provides thorough guidance in implementing cutting-edge vibrational spectroscopic methods from international leaders in the field
- Emphasizes in vivo, in situ, and noninvasive analysis of proteins in biomedical and life science research more broadly
- Chapters address vibrational spectroscopy for the study of not only simple purified proteins but also protein aggregates, fibrous proteins, membrane proteins, and protein assemblies

ACADEMIC PRESS

An imprint of Elsevier

elsevier.com/books-and-journals