

Studying Collagen Architecture in Solution by Raman Optical Activity Spectroscopy

Jiří Kessler, Jaroslav Šebestík, Martin Šafařík, Radek Pelc, Petr Bouř, and Tao Wu*

Cite This: *Anal. Chem.* 2026, 98, 7481–7488

Read Online

ACCESS |



Metrics & More

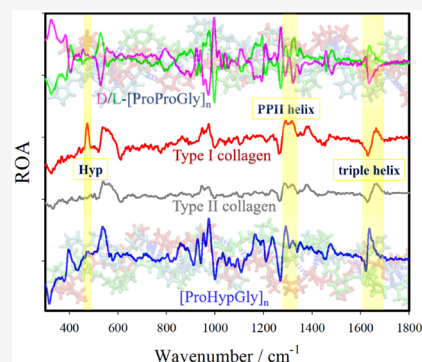


Article Recommendations



Supporting Information

ABSTRACT: Raman and Raman optical activity (ROA) spectroscopy provide a unique insight into the three-dimensional structure of biomacromolecules; however, it is often hampered by low sensitivity, low resolution, and the lack of theoretical models. To advance the methodology, we demonstrate that it can discriminate between two collagen proteins, types I and II. The data are interpreted on the basis of molecular modeling correlated with spectra of five synthetic collagen-type peptides serving as simple models. In the peptides, accurate density functional theory (DFT) calculations and correlation of the structure with the spectra are possible, allowing us to determine convenient marker bands linking spectral intensities to the molecular architecture. ROA spectra reflect the polyproline II (PPII) helical conformation of the peptide's main chain and indicate subtle concentration-dependent structural variations in type I collagen. Several vibrational bands originating from proline (Pro), hydroxyproline (Hyp), and the Pro–Hyp–Gly motifs can be related to the collagen triple helix core. ROA spectroscopy thus captures several aspects of collagen's chirality, enables the study of solvent effects and dynamics, and is expected to aid connective tissue studies.



INTRODUCTION

Collagen is the most abundant protein in the animal kingdom and is predominantly found in connective tissue, such as tendons, skin, cartilage, and bone. Its fine structure varies; fibril-forming collagen types I, II, III, V, and XI are the most important forms involved in aging and in many pathophysiological processes such as cancer and cystic fibrosis. Collagen fibers also play a key role in biomineralization by providing the scaffold for bone and dentin formation. Their highly ordered packing and hierarchical bundling lend collagen a remarkable mechanical resilience *in vivo* and even explain its long-term durability in archaeological findings.¹

A key structural feature of collagen is the triple helix, in which three polypeptide strands adopt polyproline II-type (PPII) geometry. Proline (Pro, P), hydroxyproline (Hyp, H), and glycine (Gly, G) are its most common amino acids, and the ProHypGly (PHG) triplet is instrumental in assembling the helical chains.² Despite decades of research, many aspects of collagen fibrillization remain incompletely understood.³ Even the most advanced protein-structure prediction algorithms such as AlphaFold2⁴ and AlphaFold3⁵ struggle to model higher-order collagen helices and fibrils, thus highlighting the need for more fundamental approaches that probe self-assembly across scales.⁶

The structure and physicochemical properties of collagen helices have been extensively studied.⁷ The triple-helical nature of collagen was established by diffraction studies in the 1950s.^{8,9} However, elucidation of its structure at atomic resolution appeared difficult, and an alternative approach employing collagen-type peptides as simple models was

introduced in the late 1960s.¹⁰ This led to the verification of the intramolecular hydrogen bonds between N–H(Gly) and O=C(Pro), stabilizing the triple helix using X-ray diffraction of collagen-type peptide crystals.¹¹ Numerous crystallography studies of collagen-type peptides were conducted thereafter.^{12–15}

Besides X-ray crystallography, other techniques used to study collagen proteins comprise scanning electron microscopy (SEM),¹⁶ atomic force microscopy (AFM),¹⁷ mass spectrometry (MS),¹⁸ and infrared (IR)^{19,20} and Raman²¹ spectroscopy. Unlike IR, Raman spectroscopy is very suitable for studying aqueous solutions. It is also label-free and nondestructive. As such, it has been widely used not only to characterize collagen in terms of its secondary structure but also to study the extracellular matrix, of which collagen is the main component.²¹

An even better structural sensitivity is achieved by Raman optical activity (ROA) spectroscopy, which measures a tiny intensity difference between scattered right- (I_R) and left- (I_L) circularly polarized light. It has been developed as an efficient analytical approach to study organic molecules and biomolecules²² as well as metal complexes.²³ Weak signal is ROA

Received: November 6, 2025

Revised: February 18, 2026

Accepted: February 26, 2026

Published: March 2, 2026



spectroscopy's inherent limitation. For example, the ROA/Raman ratio [a so-called normalized circular intensity difference, $CID = (I_R - I_L)/(I_R + I_L)$] is typically below 1×10^{-3} . Yet modern spectrometers do make it possible to detect the ROA signal even in samples of limited solubility, such as collagen. Another constraint is related to interpretation of the spectra, almost exclusively relying on *ab initio* modeling. Even here, however, the density functional theory (DFT), Cartesian tensor transfer, or similar fragment-based methods make the computations applicable to large molecules including biopolymers.

Previously, X-ray diffraction was used to distinguish between different collagen types,²⁴ often in conjunction with IR and Raman spectroscopy.^{25–28} However, these techniques are blind to molecular chirality, quite insensitive to minor conformational differences (IR, Raman) and “contraindicated” in solutions (X-ray diffraction). In this context, ROA spectroscopy is often a convenient complementary technique.

In the present study, we thus employ ROA spectroscopy jointly with molecular modeling to elucidate the structure of collagen proteins (types I and II) and five newly synthesized collagen-type peptides. Being highly sensitive to backbone conformation, the technique provides rich insight into the secondary and tertiary structure of these complex biomolecules in an aqueous milieu. The combination of experimental ROA spectra with theoretical modeling makes it possible to reliably assign characteristic spectral features to specific conformational motifs.

■ EXPERIMENTAL SECTION

Collagen type I from rat tail tendon and collagen type II from chicken sternal cartilage were purchased as powders from Sigma-Aldrich and dissolved in 0.1 M acetic acid. Although collagens I and II are both fibril-forming, collagen II has better solubility in 0.1 M acetic acid at room temperature. Three different concentrations (13, 23, and 50 mg/mL) were prepared for type I. The 50 mg/mL concentration could be used as well but required warming the sample to 50 °C, which led to a change in the native conformation.

Peptide Synthesis

The synthesis of the five collagen-type peptides was performed using the Fmoc/tBu protocol (Fmoc = 9-Fluorenylmethoxycarbonyl)²⁹ on an automatic solid-phase synthesizer ABI 433A (Applied Biosystems) using the FastMoc 0.1 mmol program (SynthAssist version 3.1) with a single coupling: 10 equiv. of an excess of protected amino acids, the *o*-(Benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate coupling reagent, and 20 equiv. of an excess of diisopropylethylamine were used. H-Ala-HMPB-ChemMatrix and H-Gly-HMPB-ChemMatrix resins were used, with a loading of ca. 0.4 mmol/g. Hydroxyproline (Hyp) was incorporated by using Fmoc-Hyp(tBu)-OH protection. The cleavage of the peptides from the resin was carried out by a mixture of trifluoroacetic acid (4.5 mL), H₂O (150 mL), 1,2-ethanedithiol (150 mL), thioanisole (150 mL), and triisopropyl silane (50 mL) for 4 h. Peptide deprotection was carried out simultaneously. All of the peptides were obtained at >90% purity.

Chromatographic and MS Methods

Molecular weight of peptide fragments was determined using matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS). An instrument with a quaternary pump, thermostat, diode array detector, and reverse-phase C₁₈ columns was used for HPLC analysis. Purification of the peptides was carried out by semipreparative HPLC on the VYDAC 250 × 10 mm, 10 μm RP-18 column at a flow rate of 3 mL/min using a 0–100% ACN (acetonitrile) gradient in 0.05% aqueous TFA. HPLC analysis was performed on a Poroshell 120 SB-C18 2.7 μm, 3.0 × 50 mm column at a flow rate of 1 mL/min and diode array detection using a 5–5–100–

100% gradient of ACN in 0.05% aqueous TFA within 0–1–11–13 min. The injection peak was not integrated (Figure S1).

- (1) PXGA-triacontapeptide, (PHG)₉PHA–H-(Pro-Hyp-Gly)₉–Pro-Hyp-Ala-OH—HPLC: R_T 3.2 min. For C₁₂₁H₁₇₄N₃₀O₄₁ (2703.25), found MALDI-MS *m/z*: 2704.2 ([M + H]⁺) (Figure S2).
- (2) PPGA-triacontapeptide, (PPG)₉PPA–H-(Pro-Pro-Gly)₉–Pro-Pro-Ala-OH—HPLC: R_T 4.0 min. For C₁₂₁H₁₇₄N₃₀O₃₁ (2543.30), found MALDI-MS *m/z*: 2544.1 ([M + H]⁺) (Figure S3).
- (3) ppG-triacontapeptide, D-(PPG)₁₀–H-(D-Pro-D-Pro-Gly)₁₀–OH—HPLC: R_T 4.0 min. For C₁₂₀H₁₇₂N₃₀O₃₁ (2529.28), found MALDI-MS *m/z*: 2531.1 ([M + 2]⁺) (Figure S4).
- (4) XPGA-triacontapeptide, (HPG)₉HPA–H-(Hyp-Pro-Gly)₉–Hyp-Pro-Ala-OH—HPLC: R_T 3.1 min. For C₁₂₁H₁₇₄N₃₀O₄₁ (2703.25), found MALDI-MS *m/z*: 2704.4 ([M + H]⁺) (Figure S5).
- (5) GXPA-hentriacontapeptide, (GHP)₁₀A–H-(Gly-Hyp-Pro)₁₀–Ala-OH—HPLC: R_T 3.1 min. For C₁₂₃H₁₇₇N₃₁O₄₂ (2760.27), found MALDI-MS *m/z*: 2761.3 ([M + H]⁺) (Figure S6).

Raman/ROA Spectroscopy

Raman and backscattered circular polarization (SCP) ROA spectra were acquired at room temperature on a BioTools ROA spectrometer (laser excitation: 532 nm, resolution: 7 cm^{−1}) in a cuvette of ~120 μL volume made of fused silica. The laser power was within the range of 200–1000 mW, and accumulation times were 8 to 72 h. To quench the fluorescence, all samples were irradiated at 200 mW for several hours before measurements, Raman intensities normalized to the water peak at 1650 cm^{−1}, and the remaining fluorescence background was corrected by polynomial baseline.

■ THEORETICAL CALCULATIONS

Initial geometry of the (PHG)₉PHA and (PPG)₉PPA peptides was extracted from the crystal structure of collagen-like peptides.^{11,13} Optimization and subsequent calculation of partial atomic charges by the Merz–Kollman method were carried out at the DFT level using the B3PW91 functional with the 6-31++G** basis set, CPCM water solvent model, and empirical GD3 dispersion, using the Gaussian 16 program.³⁰ The partial atomic charges were refined by the RESP procedure³¹ using the Antechamber program from the Amber 18 package.³² Each peptide was placed in a periodic box (120 × 50 × 50 Å) filled with water molecules using the Packmol program.³³

The systems were minimized, equilibrated, and then the final 10 ns molecular dynamics (MD) run was carried out using the GAFF force field,³⁴ periodic boundary conditions, a canonical ensemble (NVT), a temperature of 300 K, and a 1 fs time step.

The entire collagen-type peptide was cut into a set of partially overlapping fragments containing 4 or 7 amide units and offset by one amide. For example, the first 4-peptide fragment included amides 1 to 4, the second one included amides 2 to 5, etc.³⁵ The fragments were partially optimized using the normal mode coordinates,³⁶ and Raman and ROA tensors and force fields were calculated at the B3PW91/6-31++G**/CPCM(water) level. Then, the tensors were transferred back to the original molecule and to another 1000 snapshots to account for the MD dispersion of the geometry. Backscattered Raman and ROA intensities were calculated at the harmonic level and averaged. The intensities were convoluted with Lorentzian peaks (full width at half maximum, FWHM 10 cm^{−1}) while taking into account the Boltzmann temperature factor at 300 K.

RESULTS AND DISCUSSION

Experimentally, ROA spectra of collagen required extended acquisition times due to their inherently weak signals. Although laser light may theoretically affect the sample, no laser-induced degradation was detected in the Raman spectra. However, they exhibit a concentration dependence. Raman spectra of collagen I (23 mg/mL) and collagen II (25 mg/mL) are almost identical except for the low wavenumber (200–500 cm^{-1}) bands (Figure 1).

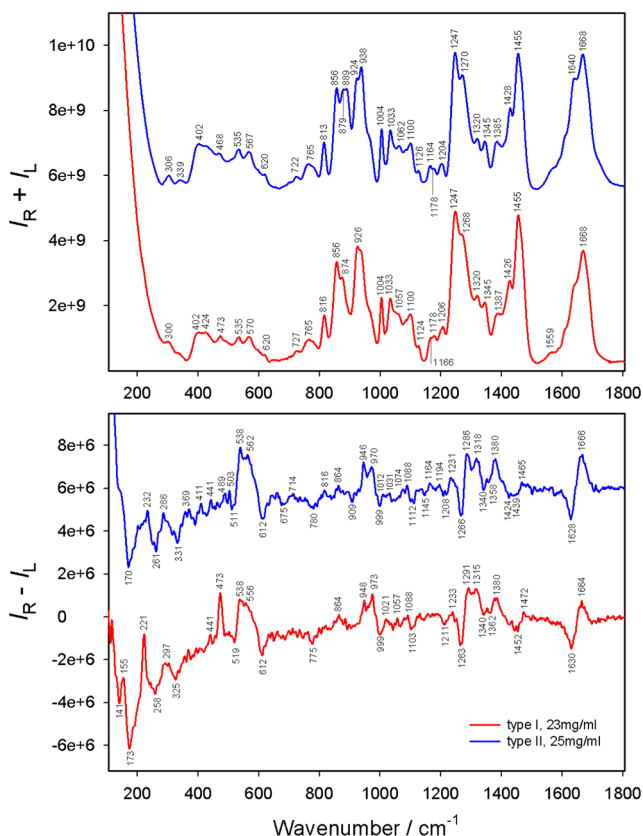


Figure 1. Raman ($I_R + I_L$) and ROA ($I_R - I_L$) spectra of collagen I (23 mg/mL) and collagen II (25 mg/mL) measured in 0.1 M acetic acid.

Raman spectrum of collagen I solution was slightly affected by local concentration variations due to viscosity. At a low concentration (13 mg/mL) of type I, there are no obvious Raman bands below 400 cm^{-1} (Figure 2). As the concentration increases to 23 mg/mL, the solution becomes more viscous, and a weak band appears at 300 cm^{-1} , most likely due to the delocalized chain deformation, possibly coupled with vibrations of the aqueous solvent.³⁷ A previous study has shown that irreversible denaturation occurs at the melting point of collagen (40 °C).³⁸ At 50 °C, the solubility of collagen I can be as high as 50 mg/mL, this being the highest concentration of collagen I we used. Upon cooling such a high-concentration solution to room temperature (at which the measurements were carried out), a gel was formed in the spectroscopic cuvette. The weak band at 300 cm^{-1} becomes hardly visible, and a new band appears at 221 cm^{-1} (corresponding to delocalized peptide chain deformation^{37,39}); it was not detected at 23 or 13 mg/mL. These spectral variations indicate that the structure of collagen I undergoes

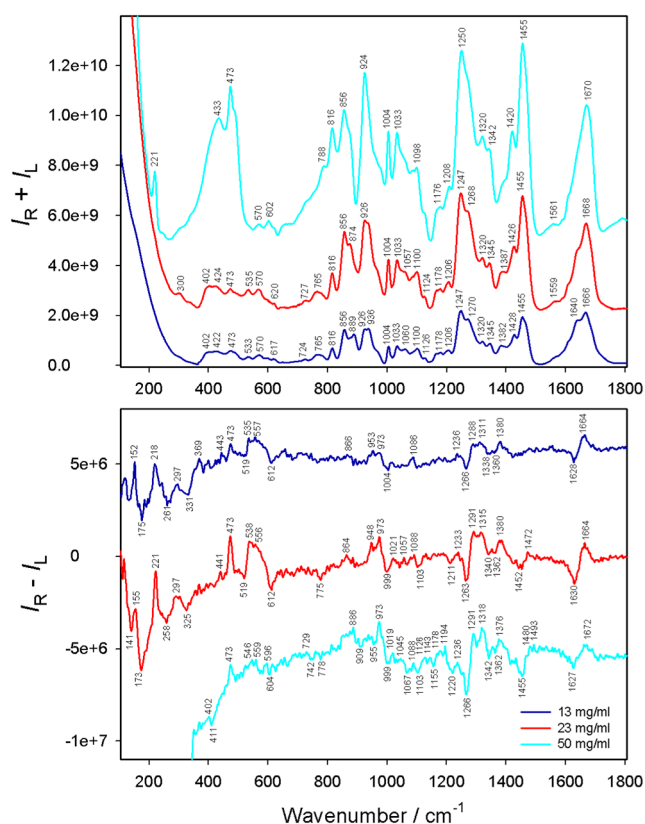


Figure 2. Raman ($I_R + I_L$) and ROA ($I_R - I_L$) spectra of collagen I (13, 23, 50 mg/mL) in 0.1 M acetic acid. Close-to-highest possible concentrations were used, and 50 mg/mL was obtained by briefly heating the sample beyond the denaturation temperature.

irreversible changes at a high temperature. Two small bands at 306 and 339 cm^{-1} are encountered in type II only and thus represent a spectral fingerprint (type II vs I). Major Raman and ROA bands of collagen proteins are given in Table 1.

For collagen I, concentration-dependent Raman features include amide I vibrations: at 50 mg/mL a broad band appeared at 1670 cm^{-1} , which was shifted to 1668 cm^{-1} at 23 mg/mL, and further split into two bands at 1666 and 1640 cm^{-1} at 13 mg/mL (Figure 2). Similar variations were observed for the amide III vibrations. While the gel-like sample (50 mg/mL) exhibited only a single band at 1250 cm^{-1} , there were two peaks at the two lower concentrations (e.g., 1270, 1247 cm^{-1} at 13 mg/mL and 1268, 1247 cm^{-1} at 23 mg/mL).

The Raman intensity of the 473 cm^{-1} band (assigned to the Pro ring by other authors^{21,41}) is much stronger at 50 mg/mL compared to 23 or 13 mg/mL. A weak Raman band was also discernible in collagen II at 468 cm^{-1} . Although the Raman intensity is strongest in the gel-state sample (50 mg/mL), the corresponding ROA intensity is the weakest, i.e., the CID value is the smallest, suggesting the triple helical structure was strongly affected by the high temperature (50 °C) required to dissolve the sample (i.e., above the denaturation temperature of 40 °C, as discussed above).

ROA spectral patterns of collagen I (13 and 23 mg/mL) and collagen II (25 mg/mL) were quite similar to each other, except in the low wavenumber region (100–500 cm^{-1}). The vibration modes of collagen I in this region are difficult to assign. ROA spectral shapes also slightly vary with concentration. Nevertheless, an intense band was recorded at 473

Table 1. Summary of the Major Raman and ROA Bands Observed for Collagen Types I and II in 0.1 M Acetic Acid

Wavenumber/cm ⁻¹				
Raman		ROA		Band assignment ^{22,40}
Type I ^a	Type II ^b	Type I ^a	Type II ^b	
1668	1668	1664 (+)	1666 (+)	Amide I
	1640	1630 (–)	1628 (–)	$\nu(\text{C=O})$
1455	1455	1472 (+)	1465 (+)	δCH_3 , δCH_2
1426	1428	1452 (–)	1439 (–)	
1345	1345	1340 (–)	1340 (–)	$\text{C}_\alpha\text{--H}$, CH_2 wagging
1320	1320	1315 (+)	1318 (+)	
1268	1270	1291 (+)	1286 (+)	Amide III, $\delta\text{N--H}$
1247	1247	1263 (–)	1266 (–)	
1033	1033	1021 (+)	1031 (+)	Phe; Hyp/Pro ring
		999 (–)	1012 (+)	
			999 (–)	
		973 (+)	970 (+)	Hyp/Pro ring deformation
926	938			$\nu(\text{C--C})$ of protein backbone
	924			$\nu(\text{C--C})$ of Pro ring
874	889	864 (+)	864 (+)	$\nu(\text{C--C})$ of Hyp/Pro ring
856	856			
765	765	775 (–)	780 (–)	Amide IV
570	567	556 (+)	562 (+)	S–S vibrations; Hyp/Pro ring deformation
535	535	538 (+)	538 (+)	
473	468	473 (+)		Hyp ring deformation
300	339	325 (–)	331 (–)	Delocalized chain deformation
	306			

^a23 mg/mL. ^b25 mg/mL.

cm⁻¹ in type I at all three concentrations (Figure 2). As it is missing in the collagen II spectrum (Figure 1), it can serve as an ROA fingerprint of collagen I.

Figure 3 shows the Raman and ROA spectra of the collagen-type peptides (PHG)₉PHA, (PPG)₉PPA, and D-(PPG)₁₀, along with the calculated spectra for the first two. Raman and ROA spectra for all five synthesized collagen-type peptides are presented in Figure S7. Comparison of Raman and ROA bands for collagen proteins and the five collagen-type peptides is provided in Table S1 and Table S2, respectively. Apparently, the Raman spectra of the peptides and proteins (collagens) are more alike than the ROA spectra. The ROA spectrum of the (PHG)₉PHA peptide is most similar to those of both collagen proteins (I or II). This includes couplets at ~1610–1670 cm⁻¹, ~1260–1295 cm⁻¹, ~1080–1100 cm⁻¹, and ~970–1000 cm⁻¹; positive bands at ~1310–1320 cm⁻¹, ~945–955 cm⁻¹, and 538 cm⁻¹; and a negative band at ~310–340 cm⁻¹.

ROA spectra of the (PPG)₉PPA and D-(PPG)₁₀ peptides are almost perfect “mirror images” despite a minor composition difference (L-Ala instead of Gly). The amide I vibration mode of (PPG)₉PPA exhibits a strong/weak positive/negative band at high/low wavenumber (1634/1619 cm⁻¹). This ROA couplet is more easily discernible when compared with D-(PPG)₁₀ than when measured alone and is well reproduced in the simulated spectrum (Figure 3). A similar couplet is often encountered in α -helical proteins.⁴² The amide I ROA bands of (PPG)₉PPA strongly depend on the concentration. A very slight decrease from 25 to 20 mg/mL results in the ROA couplet (+/–) of amide I vibration being transformed into a positive-only band; the rest of the ROA bands remain mostly unaffected (Figure S8). The ROA spectral pattern is similar to the PPII form of polyproline,³⁷

indicating that at low concentrations (PPG)₉PPA preferably adopts a PPII conformation. This is consistent with the crystal structures of similar collagen-type peptides reported elsewhere.^{11,13}

Among the five collagen-type peptides, the prominent Raman band found in collagen I at 473 cm⁻¹ (Figure 2) was detected only in (PHG)₉PHA (Figure 3 and Figure S7). It was noticeably weaker in the other two Hyp-containing peptides, (HPG)₉HPA and (GHP)₁₀A at 476 cm⁻¹ and 479 cm⁻¹, respectively. In the Hyp-free peptides, i.e., (PPG)₉PPA and D-(PPG)₁₀, it was absent altogether (as in other polyproline peptides³⁷), and 452 cm⁻¹ was its nearest equivalent. We may thus conclude^{21,41} that it is the Hyp rather than the Pro ring in the Hyp-Pro-Gly fragment that contributes to the above-mentioned 473 cm⁻¹ band.

Raman bands at 1033 and 1004 cm⁻¹ are typically assigned to phenylalanine (Phe).⁴³ We have nevertheless detected both of them in collagen I and II (Figure 1), similarly to other authors.⁴⁴ We also recorded very similar Raman bands at 1033 cm⁻¹ and 1007 cm⁻¹ in one of the Hyp-containing peptides (PHG)₉PHA, but not in the Hyp-free peptides (PPG)₉PPA and D-(PPG)₁₀ or even the Hyp-containing peptides (HPG)₉HPA and (GHP)₁₀A, in which the amino acid sequence in the Pro-Hyp-Gly domain has been modified. This indicates that the Hyp residue may contribute to the 1033/1004 cm⁻¹ bands of collagen proteins.

A Raman peak at 954 cm⁻¹, previously assigned to Pro and used to differentiate between collagen types,²⁸ was not discernible in collagen proteins and peptides measured in the present study. However, ROA bands could be clearly detected at ~940–980 cm⁻¹. This indicates that even collagen structures of the same type may differ depending on their source.

A strong Raman band at 623 cm⁻¹ was recorded in the (HPG)₉HPA peptide only (Figure S7). Incidentally, the peptide has a slightly lower solubility in acetic acid than (PHG)₉PHA, (PPG)₉PPA, and D-(PPG)₁₀. Many ROA spectral features of collagen were also absent in the (HPG)₉HPA and (GHP)₁₀A peptides, such as the amide I couplet (1610–1670 cm⁻¹) found in collagen I and II (Figure 1) and the (PHG)₉PHA peptide (Figure 3). This may be due to the low concentration similar to (PPG)₉PPA (20 versus 25 mM) (Figure S8), but another ROA couplet observed in the collagen proteins within the region of 1080–1100 cm⁻¹ is also missing. In addition, a ROA negative band at 481 cm⁻¹ was observed in (HPG)₉HPA and 479 cm⁻¹ in (GHP)₁₀A, as opposed to the positive pattern at 473 cm⁻¹ in collagen I and 479 cm⁻¹ in (PHG)₉PHA. These spectral patterns make the presence of Hyp-Pro-Gly and Gly-Hyp-Pro domains in the collagen structure rather unlikely. The Pro-Hyp-Gly and Pro-Pro-Gly domains are the most important components in collagen proteins, the former one being predominant.

Previous reports^{44,45} have attributed the Raman band at 535 cm⁻¹ to S–S vibrations. In the denatured form of collagen I (Figure 2), many of the otherwise present Raman/ROA bands were not detected, most likely due to the effect of heating on the S–S bond. However, the Raman band at 535 cm⁻¹ and a corresponding ROA band at 538 cm⁻¹ in the (PHG)₉PHA peptide (Figure 3) are almost identical to those encountered in collagen I and II proteins, and similar Raman and ROA bands were also recorded in the other four peptides (Figure S7). This indicates³⁷ that the $\nu(\alpha\text{C--N})$ vibration mode of the Hyp/Pro

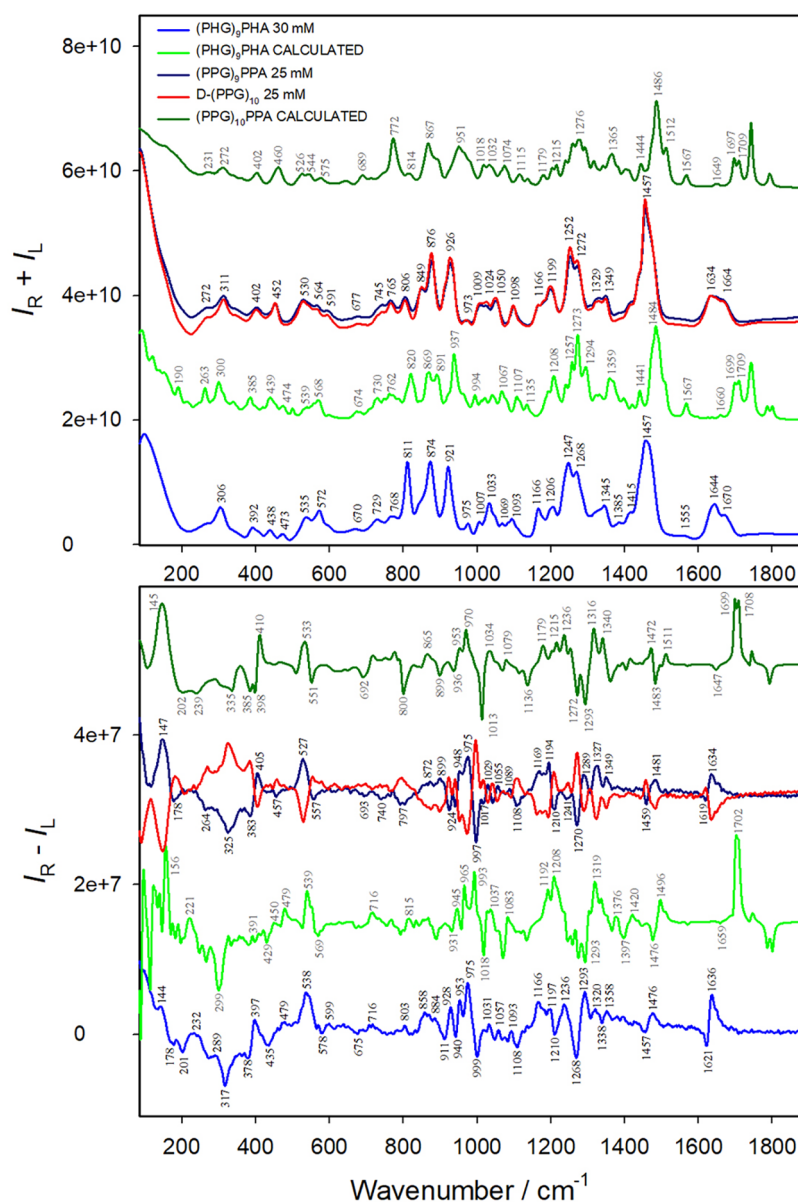


Figure 3. Raman ($I_R + I_L$) and ROA ($I_R - I_L$) spectra of three collagen-type peptides measured in 0.1 M acetic acid and calculated spectra of two of them.

ring deformation strongly contributes to the Raman band at 535 cm^{-1} .

Three parallel polypeptide strands in the left-handed PPII helical conformation form a right-handed triple collagen superhelix. One may expect to detect specific spectral patterns of both the PPII helix and the superhelix. Indeed, some studies have already characterized the PPII structure, and a strong positive ROA band in the range of $1314\text{--}1325\text{ cm}^{-1}$ has been detected.^{22,46} Positive ROA bands at 1315 cm^{-1} (intact collagen I at 23 mg/mL) and 1318 cm^{-1} (denatured collagen I at 50 mg/mL) are mostly due to $C_\alpha\text{--H}$ and CH_2 wagging (Figure 2). We recorded a similar band at 1318 cm^{-1} for collagen II (Figure 1), at 1320 cm^{-1} for the $(\text{PHG})_9\text{PHA}$ peptide, at 1327 cm^{-1} for the $(\text{PPG})_9\text{PPA}$ peptide, and at 1324 cm^{-1} for $(\text{HPG})_9\text{HPA}$ and $(\text{GHP})_{10}\text{A}$ peptides (Figure S7).

MD-simulated geometries of $(\text{PHG})_9\text{PHA}$ and $(\text{PPG})_9\text{PPA}$ in a water box (Figure S9) are close to the crystal structure of other collagen-type peptides.^{11,13} The histograms of segment

lengths in Pro-Hyp-Gly in $(\text{PHG})_9\text{PHA}$ and Pro-Pro-Gly in $(\text{PPG})_9\text{PPA}$ are comparable, as revealed by MD simulations (Figure S10). Differences between the Ramachandran plots of Gly-Pro in the $(\text{PHG})_9\text{PHA}$ peptide (Figure 4A) and $(\text{PPG})_9\text{PPA}$ peptide (Figure 4D) are minimal. A replacement of Hyp for Pro in $(\text{PHG})_9\text{PHA}$ translates to changes in the Ramachandran plot: Pro-Hyp (Figure 4B) versus Pro-Pro (Figure 4E) and Hyp-Gly (Figure 4C) versus Pro-Gly (Figure 4F). The left-handed PPII helical conformation of the three parallel strands in the two polypeptides is quite stable, and indeed, most of the observed ROA spectral patterns correspond to the PPII structure.

Major ROA spectral patterns of collagen proteins and peptides are summarized in Table S2.^{22,40} The most important one is the ROA couplet of amide I vibrations (centered at $\sim 1650\text{ cm}^{-1}$, with a negative/positive part at low/high wavenumber). The amide I couplet, arising mainly from the C=O stretching mode specific to the triple collagen

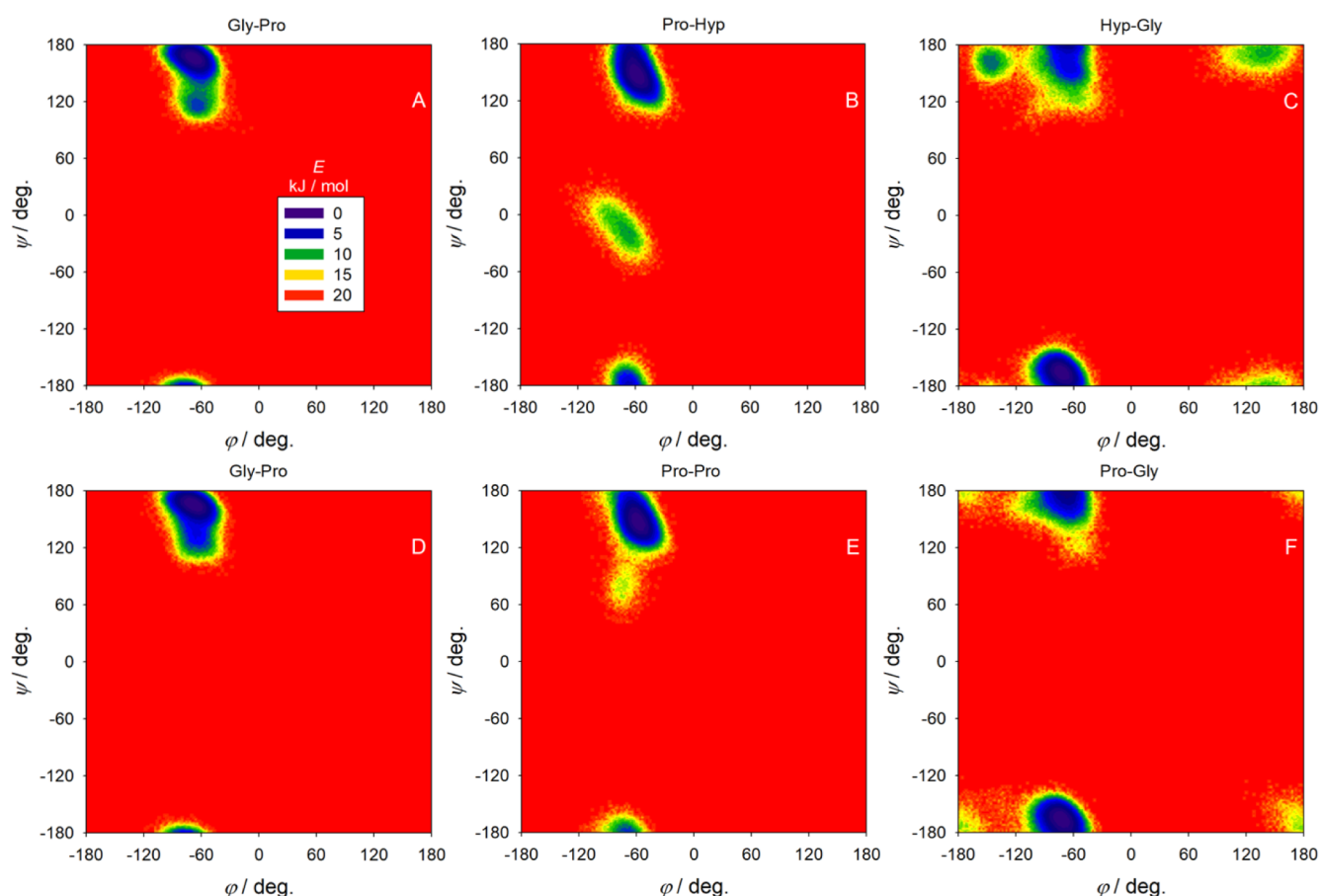


Figure 4. Dependence of the free energy on the (ϕ , ψ) angles, as obtained by MD in the Gly-Pro-Hyp-Gly sequence (A, B, C) in the (PHG)₉PHA peptide, and the Gly-Pro-Pro-Gly sequence (D, E, F) in the (PPG)₉PPA peptide.

superhelix, is also summarized separately in Table S3. In particular, we compared the CID value at the high and low wavenumber components (CID₁ vs CID₂). The greatest CID value ($\sim 9 \times 10^{-4}$) was recorded in the (PHG)₉PHA peptide at high wavenumber (1636 cm⁻¹) and smallest ($\sim 9 \times 10^{-5}$) in the denatured collagen I also at high wavenumber (1672 cm⁻¹). The difference of the two CID values ($\Delta_{\text{CID}} = \text{CID}_1 - \text{CID}_2$) followed a similar pattern, with the peptide (PHG)₉PHA featuring the highest ($\sim 1.68 \times 10^{-3}$) and denatured collagen I the smallest ($\sim 3.5 \times 10^{-4}$) values. The concentration dependence of the Δ_{CID} value for collagen I itself is prominent; the low/medium (13/23 mg/mL) concentrations yield the highest/moderate Δ_{CID} . This indicates that the low concentration (less viscous state) favors stabilization of the triple helical structure. The Δ_{CID} value of the (PHG)₉PHA peptide is slightly higher compared to either (PPG)₉PPA or D-(PPG)₁₀, indicating that the presence of Hyp contributes more strongly than Pro to the formation and stability of helical conformations.

Another ROA couplet (−/+) centered at ~ 1100 cm⁻¹ is also frequently attributed to the α -helix structure.⁴⁰ We observed it as inverted (+/−) in collagens I and II (Figure 1), and the (PHG)₉PHA and (PPG)₉PPA peptides (Figure 3); it is hardly visible in (HPG)₉HPA and (GHP)₁₀A (Figure S7). Other spectral signatures previously observed^{40,46} for the α -helices included positive ROA bands at ~ 1340 – 1345 cm⁻¹, ~ 1297 – 1312 cm⁻¹, and ~ 870 – 950 cm⁻¹. While we did not record any positive ROA band at ~ 1340 – 1345 cm⁻¹ and ~ 1297 – 1312

cm⁻¹ (either for the collagen proteins or the peptides), we noticed a positive ROA band at 948 cm⁻¹ in collagen I (at 23 mg/mL). In collagen II, we recorded a similar positive band at 946 cm⁻¹, plus a negative band at 909 cm⁻¹ (Figure 1). All five collagen-type peptides also exhibited both positive and negative ROA bands in this region (Figure S7). These spectral signatures are also seen in α -helices; thus, they cannot be used to identify collagen proteins and peptides.⁴⁰

CONCLUSIONS

We measured Raman and Raman optical activity (ROA) spectra of natural collagen as well as model peptides to study their structure in solution and used a multiscale computational protocol to interpret the experimental data. The spectroscopy confirmed that both type I and type II collagen predominantly adopt the polyproline II (PPII) helical conformation and that a characteristic amide I signature is indicative of the triple helical arrangement. These features are also found in the spectra of synthesized collagen-type peptides. Molecular dynamics (MD) and density functional theory (DFT) calculations provided a reasonable agreement between computed and experimental frequencies and intensities, thus constituting a reliable basis to interpret the experimental patterns in the peptide models and native collagen proteins. This combined experimental-and-theoretical approach makes it easier to identify spectral peaks in Raman and ROA spectra of collagen proteins and assign them to structural motifs. This is of particular relevance as no collagen crystals have been grown yet.

Raman and ROA spectra sensitively reflect the vibrational modes associated with proline (Pro) and hydroxyproline (Hyp) and confirm their cooperative role in stabilizing the Pro–Hyp–Gly motif within the collagen triple helix. For example, the Hyp-derived ROA marker band at 473 cm^{−1} makes it possible to distinguish type I from type II collagen, thus complementing earlier studies, in which this Raman band was empirically assigned to proline. The chiral Hyp fingerprint in the ROA spectra was confirmed by molecular modeling.

In the future, improvements in instrument sensitivity and molecular modeling accuracy are expected to further refine the assignment of spectral features to details of molecular architecture, dynamics, and interactions with the environment. The results obtained so far document the potential of Raman and ROA spectroscopy, when combined with molecular modeling, to relate molecular chirality to peptide folding and, to some extent, connective tissue architecture.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.analchem.5c07017>.

HPLC chromatograms of synthetic peptides, their MALDI, Raman, and ROA spectra, MD-simulated structure of two peptides, MD-simulated distribution of selected peptide segment distances, and a list of major Raman and ROA bands of collagen proteins and peptides (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Tao Wu – Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Prague 16 000, Czech Republic; orcid.org/0000-0002-0244-3046; Email: tao.wu@uochb.cas.cz

Authors

Jiří Kessler – Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Prague 16 000, Czech Republic; orcid.org/0000-0001-6307-4339

Jaroslav Šebestík – Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Prague 16 000, Czech Republic; orcid.org/0000-0002-0614-2064

Martin Šafařík – Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Prague 16 000, Czech Republic

Radek Pelc – Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Prague 16 000, Czech Republic; orcid.org/0000-0002-6651-056X

Petr Bouř – Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Prague 16 000, Czech Republic; orcid.org/0000-0001-8469-1686

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.analchem.5c07017>

Author Contributions

The manuscript was written with the contributions of all authors, and they approved its final version.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors are grateful to Dr. Jaroslav Kozák for kindly providing a sample of collagen I isolated at the Department of Biochemical Pharmacology and used in preliminary experiments. The present study was supported by the Czech Science Foundation (25-15726S to P.B.).

■ ABBREVIATIONS

ACN, acetonitrile; CID, circular intensity difference; DIPEA, diisopropylethylamine; DMF, *N,N*-dimethylmethanamide; EDT, 1,2-ethanedithiol; Fmoc, 9-fluorenylmethoxycarbonyl; HBTU, *o*-(benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate; HOBt, *N*-hydroxybenzotriazole; HPLC, high-performance liquid chromatography; MALDI, matrix-assisted laser desorption/ionization; MS, mass spectrometry; ROA, Raman optical activity; *t*Bu, *tert*-butyl; TFA, trifluoroacetic acid; TOF, time of flight (detection); TIS, triisopropyl silane; VCD, vibrational circular dichroism

■ REFERENCES

- (1) Wilke, C. The Hunt for Proteins in Dinosaur Fossils. *ACS Cent. Sci.* **2025**, *11* (7), 1020–1023.
- (2) Shoulders, M. D.; Raines, R. T. Collagen Structure and Stability. *Annu. Rev. Biochem.* **2009**, *78* (1), 929–958.
- (3) Roth, J.; Hoop, C.; Williams, J. K.; Nanda, V.; Baum, J. Real-time single-molecule observation of incipient collagen fibrillogenesis and remodeling. *Proc. Natl. Acad. Sci. U. S. A.* **2024**, *121* (33), No. e2401133121.
- (4) Akdel, M.; Pires, D. E. V.; Pardo, E. P.; Jänes, J.; Zalevsky, A. O.; Mészáros, B.; Bryant, P.; Good, L. L.; Laskowski, R. A.; Pozzati, G.; Shenoy, A.; Zhu, W. S.; Kundrotas, P.; Serra, V. R.; Rodrigues, C. H. M.; Dunham, A. S.; Burke, D.; Borkakoti, N.; Velankar, S.; Frost, A.; Basquin, J.; Lindorff-Larsen, K.; Bateman, A.; Kajava, A. V.; Valencia, A.; Ovchinnikov, S.; Durairaj, J.; Ascher, D. B.; Thornton, J. M.; Davey, N. E.; Stein, A.; Elofsson, A.; Croll, T. I.; Beltrao, P. A structural biology community assessment of AlphaFold2 applications. *Nat. Struct. Mol. Biol.* **2022**, *29* (11), 1056–1067.
- (5) Kreutzberger, M. A. B.; Yu, L. T.; Bui, T. H.; Hancu, M. C.; Purdy, M. D.; Osinski, T.; Kasson, P. M.; Egelman, E. H.; Hartgerink, J. D. A Collagen Triple Helix without the Superhelical Twist. *ACS Cent. Sci.* **2025**, *11* (2), 331–345.
- (6) Pelc, R.; Kessler, J.; Šebestík, J.; Šafařík, M.; Bouř, P.; Wu, T. Enantioselective Lanthanide Binding Modulates Collagen Self-Assembly. *Aggregate* **2026**, *7* (2), No. e70293.
- (7) Bella, J.; Berman, H. M. Crystallographic Evidence for Cα–H...O = C Hydrogen Bonds in a Collagen Triple Helix. *J. Mol. Biol.* **1996**, *264* (4), 734–742.
- (8) Rich, A.; Crick, F. H. C. The Structure of Collagen. *Nature* **1955**, *176* (4489), 915–916.
- (9) Cowan, P. M.; McGavin, S.; North, A. C. T. The Polypeptide Chain Configuration of Collagen. *Nature* **1955**, *176* (4492), 1062–1064.
- (10) Fields, G. B.; Prockop, D. J. Perspectives on the synthesis and application of triple-helical, collagen-model peptides. *Biopolymers* **1996**, *40* (4), 345–357.
- (11) Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M. Crystal and Molecular Structure of a Collagen-Like Peptide at 1.9 Å Resolution. *Science* **1994**, *266* (5182), 75–81.
- (12) Nagarajan, V.; Kamitori, S.; Okuyama, K. Crystal structure analysis of collagen model peptide (Pro-Pro-Gly)₁₀. *J. Biochem.* **1998**, *124* (6), 1117–1123.
- (13) Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari, A. Crystal structure of the collagen triple helix model [(Pro-Pro-Gly)₁₀]₃. *Protein Sci.* **2002**, *11* (2), 262–270.
- (14) Schumacher, M. A.; Mizuno, K.; Bächinger, H. P. The crystal structure of a collagen-like polypeptide with 3(S)-hydroxyproline

residues in the Xaa position forms a standard 7/2 collagen triple helix. *J. Biol. Chem.* **2006**, 281 (37), 27566–27574.

(15) Okuyama, K.; Miyama, K.; Mizuno, K.; Bächinger, H. P. Crystal structure of (Gly-Pro-Hyp)₉: Implications for the collagen molecular model. *Biopolymers* **2012**, 97 (8), 607–616.

(16) Tampieri, A.; Celotti, G.; Landi, E.; Sandri, M.; Roveri, N.; Falini, G. Biologically inspired synthesis of bone-like composite: Self-assembled collagen fibers/hydroxyapatite nanocrystals. *J. Biomed. Mater. Res., Part A* **2003**, 67A (2), 618–625.

(17) Habelitz, S.; Balooch, M.; Marshall, S. J.; Balooch, G.; Marshall, G. W. In situ atomic force microscopy of partially demineralized human dentin collagen fibrils. *J. Struct. Biol.* **2002**, 138 (3), 227–236.

(18) Brandt, J.; Krogh, T. N.; Jensen, C. H.; Frederiksen, J. K.; Teisner, B. Thermal instability of the trimeric structure of the N-terminal propeptide of human procollagen type I in relation to assay technology. *Clin. Chem.* **1999**, 45 (1), 47–53.

(19) Talari, A. C. S.; Martinez, M. A. G.; Movasaghi, Z.; Rehman, S.; Rehman, I. U. Advances in Fourier transform infrared (FTIR) spectroscopy of biological tissues. *Appl. Spectrosc. Rev.* **2017**, 52 (5), 456–506.

(20) Mukherjee, S.; Gopinath, A.; Madhan, B.; Shanmugam, G. Vibrational circular dichroism spectroscopy as a probe for the detection of collagen fibril and fibrillation in solution. *Biosens. Bioelectron. X* **2022**, 10, 100108.

(21) Martinez, M. G.; Bullock, A. J.; MacNeil, S.; Rehman, I. U. Characterisation of structural changes in collagen with Raman spectroscopy. *Appl. Spectrosc. Rev.* **2019**, 54 (6), 509–542.

(22) Barron, L. D. The development of biomolecular Raman optical activity spectroscopy. *Biomed. Spectrosc. Imaging* **2015**, 4, 223–253.

(23) Wu, T.; Pelc, R.; Bouř, P. Molecular Properties of 3d and 4f Coordination Compounds Deciphered by Raman Optical Activity Spectroscopy. *ChemPlusChem* **2023**, 88 (9), No. e202300385.

(24) Grynpas, M. D.; Eyre, D. R.; Kirschner, D. A. Collagen type II differs from type I in native molecular packing. *Biochim. Biophys. Acta* **1980**, 626 (2), 346–355.

(25) Belbachir, K.; Noreen, R.; Gouspillou, G.; Petibois, C. Collagen types analysis and differentiation by FTIR spectroscopy. *Anal. Bioanal. Chem.* **2009**, 395 (3), 829–837.

(26) Hanifi, A.; McCarthy, H.; Roberts, S.; Pleshko, N. Fourier Transform Infrared Imaging and Infrared Fiber Optic Probe Spectroscopy Identify Collagen Type in Connective Tissues. *PLoS One* **2013**, 8 (5), No. e64822.

(27) Van Gulick, L.; Saby, C.; Morjani, H.; Beljebbar, A. Age-related changes in molecular organization of type I collagen in tendon as probed by polarized SHG and Raman microspectroscopy. *Sci. Rep.* **2019**, 9 (1), 7280.

(28) Ingle, I.; Kerns, J. G.; Shepherd, R. F. Multivariate analysis of Raman spectra for discriminating human collagens: In vitro identification of extracellular matrix collagens produced by an osteosarcoma cell line. *Spectrochim. Acta, Part A* **2025**, 328, 125434.

(29) Fields, G. B.; Noble, R. L. Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *Int. J. Pept. Protein Res.* **1990**, 35 (3), 161–214.

(30) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H., et al. *Gaussian 16, Revision C.01*; Gaussian, Inc.: Wallingford CT, 2016.

(31) Wang, J.; Cieplak, P.; Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **2000**, 21 (12), 1049–1074.

(32) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; Debolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput. Phys. Commun.* **1995**, 91 (1–3), 1–41.

(33) Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. PACKMOL: A package for building initial configurations for

molecular dynamics simulations. *J. Comput. Chem.* **2009**, 30 (13), 2157–2164.

(34) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, 25 (9), 1157–1174.

(35) Kessler, J.; Kapitán, J.; Bouř, P. First-Principles Predictions of Vibrational Raman Optical Activity of Globular Proteins. *J. Phys. Chem. Lett.* **2015**, 6 (16), 3314–3319.

(36) Bouř, P.; Keiderling, T. A. Partial optimization of molecular geometry in normal coordinates and use as a tool for simulation of vibrational spectra. *J. Chem. Phys.* **2002**, 117 (9), 4126–4132.

(37) Profant, V.; Baumruk, V.; Li, X.; Safárik, M.; Bouř, P. Tracking of the Polyproline Folding by Density Functional Computations and Raman Optical Activity Spectra. *J. Phys. Chem. B* **2011**, 115 (50), 15079–15089.

(38) Mrevlishvili, G. M.; Metreveli, N. O.; Razmadze, G. Z.; Mdzinarashvili, T. D.; Kakabadze, G. R.; Khvedelidze, M. M. Partial heat capacity change—Fundamental characteristic of the process of thermal denaturation of biological macromolecules (proteins and nucleic acids). *Thermochim. Acta* **1998**, 308 (1), 41–48.

(39) Tsereteli, G. I.; Belopolskaya, T. V.; Grunina, N. A.; Vavelioug, O. L. Calorimetric Study of the Glass Transition Process in Humid Proteins and DNA. *J. Therm. Anal. Calorim.* **2000**, 62 (1), 89–99.

(40) Barron, L. D.; Hecht, L.; Blanch, E. W.; Bell, A. F. Solution structure and dynamics of biomolecules from Raman optical activity. *Prog. Biophys. Mol. Biol.* **2000**, 73 (1), 1–49.

(41) Guilbert, M.; Said, G.; Happillon, T.; Untereiner, V.; Garnotel, R.; Jeannesson, P.; Sockalingum, G. D. Probing non-enzymatic glycation of type I collagen: A novel approach using Raman and infrared biophotonic methods. *Biochim. Biophys. Acta, Gen. Subj.* **2013**, 1830 (6), 3525–3531.

(42) Barron, L. D. Structure and behaviour of biomolecules from Raman optical activity. *Curr. Opin. Chem. Biol.* **2006**, 16 (5), 638–643.

(43) Hernández, B.; Pflüger, F.; Kruglik, S. G.; Ghomi, M. Characteristic Raman lines of phenylalanine analyzed by a multi-conformational approach. *J. Raman Spectrosc.* **2013**, 44 (6), 827–833.

(44) Nguyen, T. T.; Gobinet, C.; Feru, J.; Brassart-Pasco, S.; Manfait, M.; Piot, O. Characterization of Type I and IV Collagens by Raman Microspectroscopy: Identification of Spectral Markers of the Dermo-Epidermal Junction. *J. Spectrosc.* **2012**, 27, 421–427.

(45) Rygula, A.; Majzner, K.; Marzec, K. M.; Kaczor, A.; Pilarczyk, M.; Baranska, M. Raman spectroscopy of proteins: a review. *J. Raman Spectrosc.* **2013**, 44 (8), 1061–1076.

(46) McColl, I. H.; Blanch, E. W.; Gill, A. C.; Rhie, A. G. O.; Ritchie, M. A.; Hecht, L.; Nielsen, K.; Barron, L. D. A New Perspective on β -Sheet Structures Using Vibrational Raman Optical Activity: From Poly(L-lysine) to the Prion Protein. *J. Am. Chem. Soc.* **2003**, 125 (33), 10019–10026.