JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Transferability of Various Molecular Property Tensors in Vibrational Spectroscopy

Shigeki Yamamoto,*[,†] Xiaojun Li,[‡] Kenneth Ruud,[‡] and Petr Bouř*[,†]

[†]Institute of Organic Chemistry and Biochemistry, Academy of Sciences, 166 10 Prague, Czech Republic

[‡]Centre for Theoretical and Computational Chemistry, Department of Chemistry, University of Tromsø, N-9037 Tromsø, Norway

ABSTRACT: Transfers of polarizability, dipole, and energy derivatives from smaller fragments enable extension of ab initio vibrational spectral simulations to larger molecules. The accuracy of the transfer, however, varies according to the molecule and the tensor type. Recent works (*J. Chem. Theory Comput.* **2011**, *7*, 1867 and *Coll. Czech. Chem. Commun.* **2011**, *76*, 567) questioned the accuracy of the tensor transfer for Raman optical activity (ROA) spectra. To estimate relative errors, the current study systematically compares transfer properties for four spectral types (ROA, Raman, vibrational circular dichroism (VCD), and infrared absorption (IR)). As test molecules, polyglycine and polyproline peptides in polyproline II and α-helical conformations were used. The results indicate that the discrepancies could be caused by variation of the transfer parameters but not by fundamental differences in the ROA tensor properties. In particular, when done consistently, the transferred electric and magnetic dipole derivatives, polarizability tensors, and harmonic force field provided a reasonable convergence of all spectra for transfers with increasing sizes of the fragments. The polarized spectroscopies (VCD and ROA), however, were found to be significantly more sensitive than the unpolarized (IR, Raman) ones. Optimization of the geometry and the character of the amino acid side chain was relatively unimportant for the transfer. On the other hand, intramolecular hydrogen bonds in the helix required larger fragments for the same accuracy than in an unfolded peptide. Nevertheless, with reasonable fragmentation, all the spectra (IR, VCD, Raman, and ROA) of the target structures could be reconstructed to an accuracy sufficient for a full interpretation of the experiment, with substantial savings of computer time as an added benefit.

INTRODUCTION

Vibrational spectroscopy provides precious information about molecular structure. In particular, chiral techniques exploring different absorption (vibrational circular dichroism, VCD) or scattering (Raman optical activity, ROA) of right- and left-circularly polarized light[1,2] are very sensitive not only to absolute configuration but also to finer differences in molecular conformation. Very typical applications of those methods are studies of biopolymers, such as peptides and proteins,[3] nucleic acids,[4] or saccharides.[5]

The spectra can be conveniently interpreted by quantum chemical methods, most often using density functional theory.[6,7] Efficient analytical implementations[6,8] in computer codes[9,10] allow for direct computations of systems with hundreds of atoms.[11] Suitably chosen basis sets can reduce computational costs significantly.[12,13] Nevertheless, because of the extensive demands of the computational methods on computer resources, the applicability of the direct approach is still severely limited.

In many cases, the Cartesian coordinate transfer (CCT) of molecular property tensors[14] may be a more economic alternative. The method was applied, for example, to proteins[15] and nucleic acids.[16] For infrared absorption (IR) and VCD it provided results very close to full quantum computations.[17] For Raman and ROA, despite successful applications reproducing the experimental spectra,[18,19] two recent studies questioned the accuracy and general applicability of CCT.[20,21] Typically, fragments comprising about four amino acid residues were used in such simulations. Thus a question arises if the agreement of CCT-generated spectra with experiment was not accidental. The ROA spectra obtained by the transfer method were found to be significantly less accurate than Raman intensities, and errors of up to 50 cm^{-1} in vibrational frequencies were reported.[20] Dependent on the fragmentation scheme, large errors were observed for some proteins, and a significant influence of mutual polarization of molecular parts on the transfer was suggested.[21]

As already pointed out,[20] these inconsistencies could not be easily reconciled in the past because of the various implementations and fine parametrization of the transfer[14,20] and because of different estimations of the success: sometimes the errors were based on the level of agreement with experiment and sometimes the transferred results were compared with a benchmark computation. Some problems in ref 21 were explained by an error in the computer implementation. Still, because of the many applications, there is a need to test the CCT method in more quantitative terms. In particular, a priori assessment or at least a rough estimation of the error associated with CCT is highly desirable for future applications.

To rectify the previous concerns, in the current study we investigate model systems where the transfer parameters could be controlled in a systematic way and essential aspects of the transfer thus could be studied more consistently. For the Raman and ROA intensity tensors, the through-space polarization influence on tensor derivatives is estimated for the water dimer. Such a model has previously proved to be useful for

Received: October 10, 2011
Published: January 25, 2012

estimating the static (nondifferentiated) polarizabilities.[21] We then study the convergence of the IR, VCD, Raman, and ROA spectra for model $(Gly)_{20}$ and $(Pro)_{20}$ oligopeptides with respect to the size of the source fragment. The results reveal, for example, the different behavior of the chiral (VCD and ROA) and nonpolarized (Raman and IR) spectra or the role of the force field in the transfer. Finally, the effects of computational aspects, such as the optimization, spectral region, and peptide type and conformation, on the quality of the results are discussed.

METHODS

Mutual Group Polarization in the Water Dimer. We can think of a molecule as a system of individual chromophores or groups i. Each group has a dipole moment μ_i, an electric dipole–electric dipole polarizability $\boldsymbol{\alpha}_i$, magnetic dipole–electric dipole polarizability $\mathbf{G}'_i$, and electric quadrupole–electric dipole polarizability $\mathbf{A}_i$, etc.[1] The dipole derivatives (atomic polar tensor, APT) and magnetic dipole derivatives (atomic axial tensor, AAT)[22] needed for IR and VCD do not formally depend on electromagnetic interactions between the groups.

However, for the tensors important for Raman and ROA spectroscopy ($\boldsymbol{\alpha}$, $\mathbf{A}$, and $\mathbf{G}'$), we must consider the mutual interaction between groups even in the independent group model.[1,21] To derive this dependence, we write molecular properties using a common origin, whereas the group properties are expressed in terms of local origins. The total tensors are then obtained as sums of individual group contributions, terms accounting for the origin dependence and terms due to the mutual polarization of the groups. For the mutual polarization terms, we only consider the electric dipolar interaction and obtain the following expressions for the optical activity tensors:[1]

$$\alpha_{\alpha\beta} = \sum_i \alpha_{i,\alpha\beta} + \sum_{j\neq i} \alpha_{i,\alpha\chi} T_{ij,\chi\varepsilon} \alpha_{j,\varepsilon\beta} \tag{1}$$

$$G'_{\alpha\beta} = \sum_i \left(G'_{i,\alpha\beta} - \frac{\omega}{2} \varepsilon_{\beta\gamma\delta} r_{i,\gamma} \alpha_{i,\delta\alpha} - \frac{\omega}{2} \varepsilon_{\beta\gamma\delta} \sum_{j\neq i} r_{i,\gamma} \alpha_{i,\delta a} T_{ij,ab} \alpha_{j,b\alpha} \right) \tag{2}$$

$$A_{\gamma,\alpha\beta} = \sum_i \left[A_{i\gamma,\alpha\beta} + \frac{3}{2}(r_{i,\alpha}\alpha_{i,\beta\gamma} + r_{i,\beta}\alpha_{i,\alpha\gamma}) - \delta_{\alpha\beta} r_{i,c} \alpha_{i,c\gamma} + \sum_{j\neq i} \left(\frac{3}{2}(r_{i,\alpha}\alpha_{i,\beta a} + r_{i,\beta}\alpha_{i,\alpha a}) - \delta_{\alpha\beta} r_{i,c} \alpha_{i,ca} \right) T_{ij,ab} \alpha_{j,b\gamma} \right] \tag{3}$$

where we have introduced the distance tensor $T_{ij,\alpha\beta} = (3r_{ij,\alpha}r_{ij,\beta} - \delta_{\alpha\beta}r_{ij}^2)(4\pi\varepsilon_0 r_{ij}^5)^{-1}$, $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$, $\mathbf{r}_i$ is the position of group i defined as

$$\mathbf{r}_i = N_i^{-1} \sum_{\lambda=1}^{N_i} \mathbf{r}^\lambda$$

and N_i is the number of atoms in group i.

We abbreviate derivatives with respect to the position of nucleus λ as $(\partial\alpha_{\alpha\beta})/\partial\mathbf{r}_\varepsilon^\lambda = \alpha_{\alpha\beta}^{(\lambda\varepsilon)}$, etc. The group containing the atom λ is denoted by the letter l. All derivatives are written using local origins located at the differentiated atoms. The leading order expansion terms are then obtained as

$$\alpha_{\alpha\beta}^{(\lambda\varepsilon)} \cong \alpha_{l,\alpha\beta}^{(\lambda\varepsilon)} + \sum_{j\neq l} (\alpha_{l,\alpha\chi}^{(\lambda\varepsilon)} T_{lj,\chi e} \alpha_{j,e\beta} + \alpha_{j,\alpha\chi} T_{lj,\chi e} \alpha_{l,e\beta}^{(\lambda\varepsilon)}) + \sum_{j\neq i} \alpha_{j,\alpha\chi} T_{ij,\chi e}^{(\lambda\varepsilon)} \alpha_{i,e\beta} \tag{1a}$$

$$G'^{(\lambda\varepsilon)}_{\alpha\beta} = G'^{(\lambda\varepsilon)}_{l,\alpha\beta} - \frac{\omega}{2} \sum_{j\neq l} \varepsilon_{\beta\gamma\delta} r_{i,\gamma} \alpha_{i,\delta a} T_{il,ab} \alpha_{l,b\alpha}^{(\lambda\varepsilon)} - \frac{\omega}{2} \varepsilon_{\beta\gamma\delta} \sum_{j\neq i} r_{i,\gamma} \alpha_{i,\delta a} T_{ij,ab}^{(\lambda\varepsilon)} \alpha_{j,b\alpha} \tag{2a}$$

$$A_{\gamma,\alpha\beta}^{(\lambda\varepsilon)} = A_{l\gamma,\alpha\beta}^{(\lambda\varepsilon)} + \alpha_{l,\gamma b}^{(\lambda\varepsilon)} \sum_{i\neq l} T_{il,ba} \left(\frac{3}{2}(r_{i,\alpha}\alpha_{i,\beta a} + r_{i,\beta}\alpha_{i,\alpha a}) - \delta_{\alpha\beta} r_{i,c} \alpha_{i,ca} \right) + \sum_{j\neq i} \left(\frac{3}{2}(r_{i,\alpha}\alpha_{i,\beta a} + r_{i,\beta}\alpha_{i,\alpha a}) - \delta_{\alpha\beta} r_{i,c} \alpha_{i,ca} \right) T_{ij,ab}^{(\lambda\varepsilon)} \alpha_{j,b\gamma} \tag{3a}$$

The distance tensor derivative can be obtained from a $\mathbf{t}$ tensor, $T_{ij,\alpha\beta}^{(\lambda\varepsilon)} = N_l^{-1}(\delta_{il} - \delta_{jl}) t_{ij,\alpha\beta\varepsilon}$, $t_{ij,\alpha\beta\varepsilon} = 3[(\delta_{\alpha\beta} r_{ij,\varepsilon} + \delta_{\alpha\varepsilon} r_{ij,\beta} + \delta_{\beta\varepsilon} r_{ij,\alpha}) r_{ij}^2 - 5 r_{ij,\alpha} r_{ij,\beta} r_{ij,\varepsilon}](4\pi\varepsilon_0 r_{ij}^7)^{-1}$. The correction terms quickly recede with the distance between the interacting chromophores, either as r^{-3} for the $\mathbf{T}$ tensor (referred to as the "T-correction") or as r^{-4} for its derivatives ("t-correction"). Resultant derivatives of $\mathbf{G}'$ and $\mathbf{A}$ expressed in the common origin ($\mathbf{0}$) can be obtained from the local origin ($\mathbf{R}$) with the usual transformations[1]

$$G'^{(\lambda\varepsilon)}_{\alpha\beta}(\mathbf{0}) = G'^{(\lambda\varepsilon)}_{\alpha\beta}(\mathbf{R}) - \frac{\omega}{2} \varepsilon_{\beta\gamma\delta} R_\gamma \alpha_{\delta\alpha}^{(\lambda\varepsilon)} \tag{4}$$

and

$$A_{\gamma,\alpha\beta}^{(\lambda\varepsilon)}(\mathbf{0}) = A_{\gamma,\alpha\beta}^{(\lambda\varepsilon)}(\mathbf{R}) + \frac{3}{2}(R_\alpha \alpha_{\beta\gamma}^{(\lambda\varepsilon)} + R_\beta \alpha_{\alpha\gamma}^{(\lambda\varepsilon)}) - \delta_{\alpha\beta} R_c \alpha_{c\gamma}^{(\lambda\varepsilon)} \tag{5}$$

To test the importance of the various terms in the above expressions (eqs 1a–3a), the dynamic polarizability tensors ($\boldsymbol{\alpha}$, $\mathbf{G}'$, and $\mathbf{A}$) of two water molecules (with parallel water planes, rotated by 45°, d_{OH} = 0.943 Å, ∠HOH = 106°) separated by 4, 4.5, 5.0, 5.5, 6.0, 7.0, 8.5, and 10 Å were calculated by Gaussian[10] at the HF/aug-cc-pVTZ level, using the excitation wavelength of 532 nm, which corresponds to a typical ROA experiment.[23] Our transfer CCT[24] and smaller programs were used for the tensor transfer and analyses.

In this work, the mutual polarization terms (eqs 1a–3a) were added only for the water test (Figure 3); otherwise, results based on the original implementation[14] are shown.

Amino Acid Dimers. To estimate the transfer error for noncovalently bound systems and to test the amino acid side chain influence, we also investigated amino acid dimers (Ala·Ala, Ala·Val, Tyr·Tyr, and Pro·Pro). The tensors $\boldsymbol{\alpha}$, $\mathbf{G}'$, and $\mathbf{A}$, were transferred from the monomer, and the exact and transferred Raman and ROA spectra were simulated at several

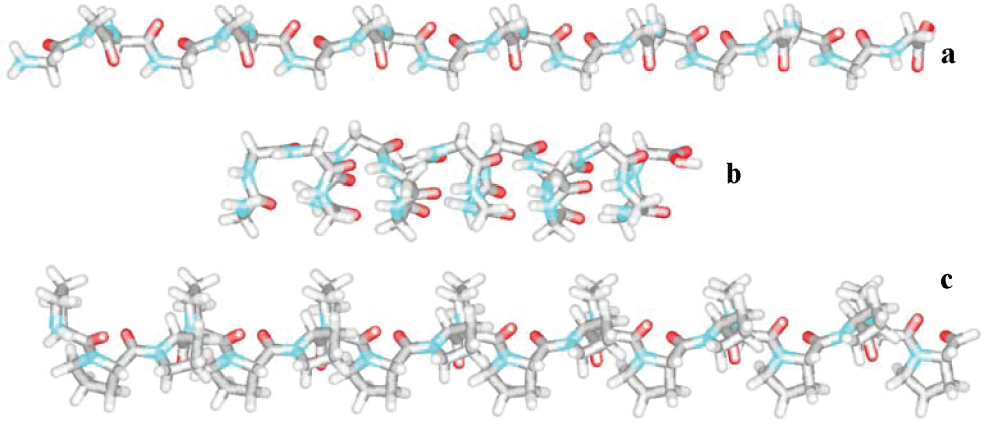

Figure 1. The $NH_2-(Gly)_{20}-COOH$ peptide in the PPII (a) and α-helical (b) conformations and polyproline conformation of $(Pro)_{20}-COH$ (c).

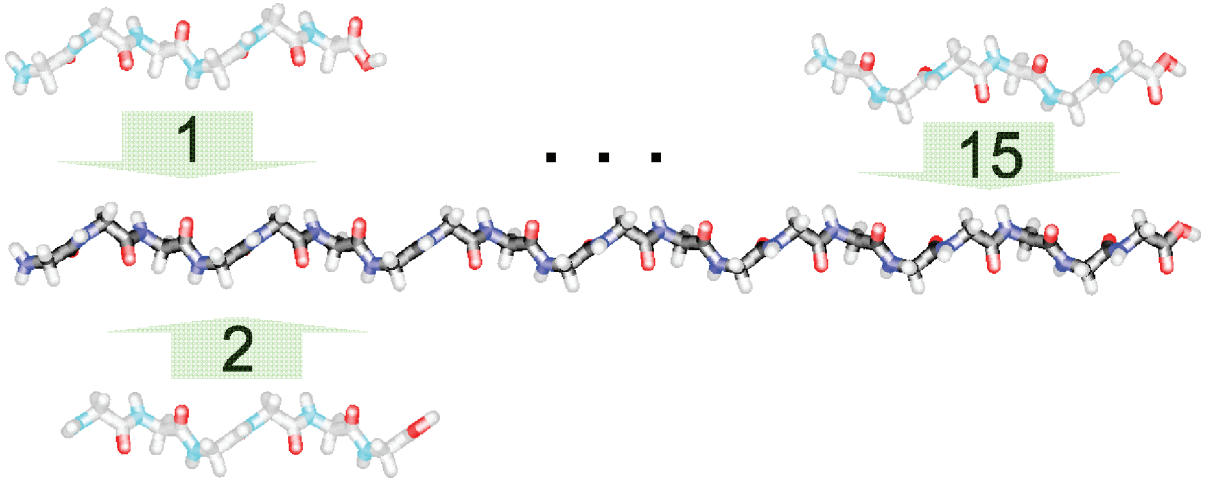

Figure 2. Transfer scheme of the tensor from the $NH_2-(Gly)_6-COOH$ fragment to the peptide $NH_2-(Gly)_{20}-COOH$. The fragment slid along the longer chain, making altogether 15 overlaps. For each atom or atomic pair in the peptide the closest fragment was used.

separations given by the *C−*C chiral carbon distances. The monomers were optimized at the B3PW91/6-31G** level, whereas the HF/rDPS[12] basis was used for the Raman and ROA tensors. The force field was not transferred but calculated for the dimer at the B3PW91/6-31G** level. Other basis sets and levels provided nearly the same results and are not shown.

Transfer on Oligopeptides. The error associated with the transfer of the ROA, AAT, APT tensors and the molecular force field (FF) and the effect on the spectra of systems that represent realistic application studies of ROA and VCD was investigated with $(Gly)_{20}$ (in polyproline II (PPII) and α-helical conformations) and $(Pro)_{20}$ peptides (Figure 1). For $(Pro)_{20}$, the COH (aldehyde) C-terminal group was used for numerical stability of the ROA tensor computations. The PPII conformation is a faithful model of so-called random peptide conformation and represents an unfolded peptide without internal hydrogen bonds.[25] The α-helix serves as a model of a folded peptide with internal hydrogen bonds.[26] The comparison of Pro and Gly enabled us to investigate the role of the amino acid side chain.

The geometries of $(X)_{20}$ (X = Gly, Pro) and the smaller analogues $(X)_N$, N = 1, 2, ..., 19 were created by the MCM software. Standard PPII (ω = 180°, φ = 78°, ψ = 149°) and α-helical (ω = 180°, φ = −57°, ψ = −47°) torsional angles were used.[26] Then the Gaussian program was used for geometry optimization with all torsion angles constrained to the initial values, with the B3LYP/6-31G** computational model. Alternatively, the structure of $(Pro)_{20}$ in the PPII conformation was fully optimized. Because of the similarity of the results for fully and partially optimized structures, the fully optimized geometry is used only in the example in Figure 10. The Raman, ROA, IR, and VCD tensors and spectral intensities within the harmonic approximation were calculated using Gaussian at the same level as the geometries.

Within the tensor transfer approach,[14] the smaller fragments were slid along the longer chain with the increment of one amino acid (Figure 2). Note that one (intensity tensors) or two (FF) atomic tensor components need to be rotated within CCT and that the rotation matrices are obtained by a best overlap of corresponding atoms.[14,20] In this work only the best overlap was used; the transfer program (CCT)[24] makes it possible to weight the overlaps, which, however, produces nearly the same results as the best overlap option.[21]

The difference between the full ab initio spectral curves and spectra obtained using the CCT method was calculated using the dimensionless quantity

$$\delta = \int_{\omega_{min}}^{\omega_{max}} |S_N(\omega) - S_{20}(\omega)|d\omega / \int_{\omega_{min}}^{\omega_{max}} |S_{20}(\omega)|d\omega \quad (6)$$

dx.doi.org/10.1021/ct200714h | J. Chem. Theory Comput. 2012, 8, 977−985

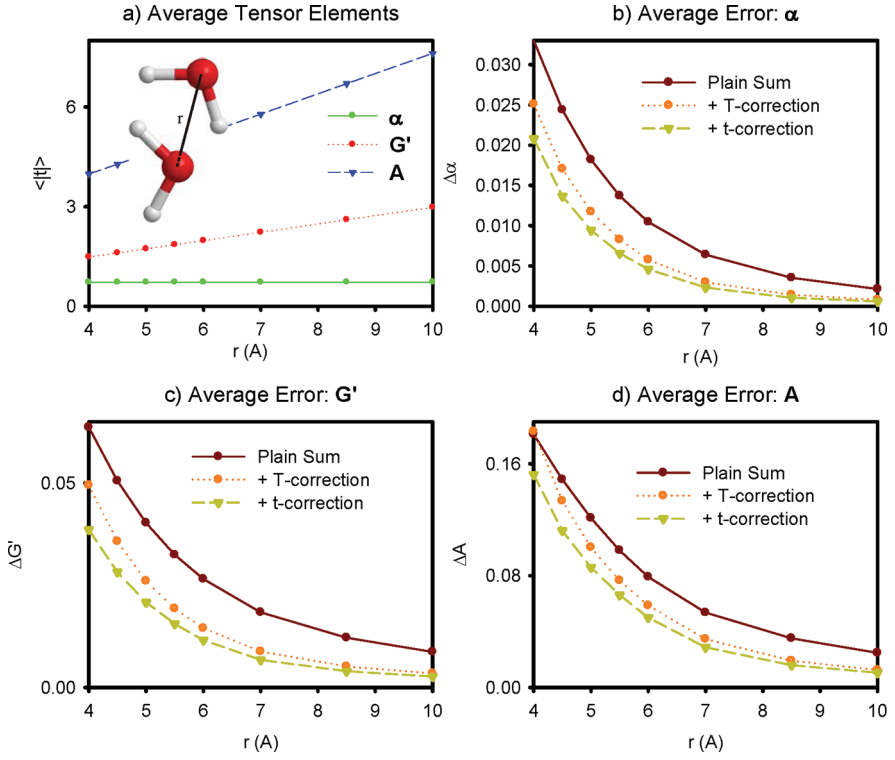

Figure 3. (a) Water dimer, average values of the tensor derivatives, and average errors of the (b) $\boldsymbol{\alpha}$, (c) $\mathbf{G}'$, and (d) $\mathbf{A}$ tensor derivatives as dependent on the distance.

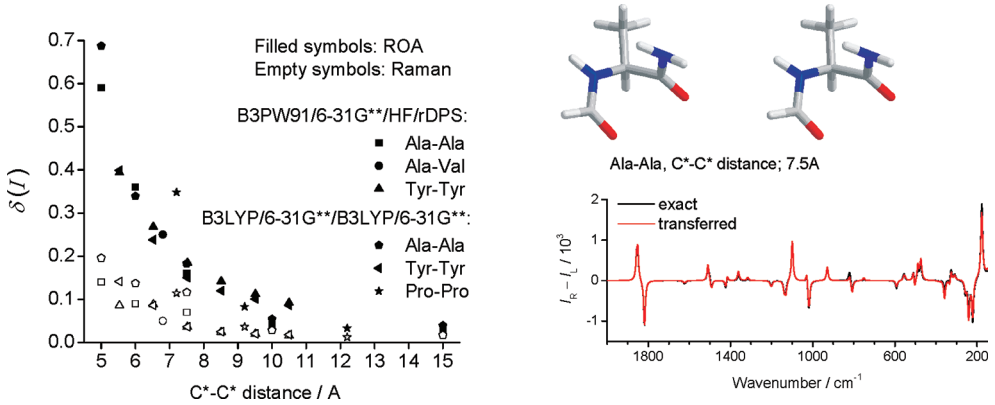

Figure 4. (left) The dependence of the errors of the ROA and Raman spectra of selected amino acid dimers on the C*−C* distance, and (right) example of the ROA spectra for the Ala·Ala dimer. The computational levels are indicated separately, respectively, for FF (not transferred) and the ROA tensors.

where $S_N(\omega)$ is the spectrum obtained by the transfer from a fragment containing N amino acids, and $S_{20}(\omega)$ is the reference spectrum. By default we constrain ourselves to the spectroscopically most important frequency region from ω_{min} = 200 to ω_{max} = 2000 cm^{-1}, although no qualitative differences in the results were observed for the frequencies above 2000 cm^{-1}. For calculated spectra, Lorentzian bands with full widths at half height (fwhh) of 10 cm^{-1} were used.

Note that although the dimensionless quantity δ can conveniently be expressed in %, it can exceed 100%. The parameter is very sensitive to even small differences in the spectra, depends on fwhh, and bears in itself a frequency shift of vibrational bands via the convolution with the Lorentzian function. In particular, even very similar spectral shapes can results in large δ error. On the other hand, for no frequency shifts, a 10% intensity error would correspond to δ = 10%, etc.

RESULTS AND DISCUSSION

Mutual Polarization Error for the ROA and Raman Tensors. The results from the water dimer test are summarized in Figure 3. From panel (a) in the figure we can see that average absolute values of tensor $\mathbf{G}'$ and $\mathbf{A}$ derivatives grow because of the origin-dependent terms (4) and (5). On the other hand, the $\boldsymbol{\alpha}$ tensor derivatives are not dependent on the gauge origin and change only slightly with increasing separation of the two water molecules. The average error of $\boldsymbol{\alpha}$ (panel b) is relatively small (~3%) already for the shortest distance of 4 Å, and further decreases with r. The error can be significantly reduced (up to 50%) by accounting for the T- and t-corrections. The latter correction becomes especially important for shorter distances; nevertheless its contribution is always limited to <30% of the total error.

dx.doi.org/10.1021/ct200714h | J. Chem. Theory Comput. 2012, 8, 977−985

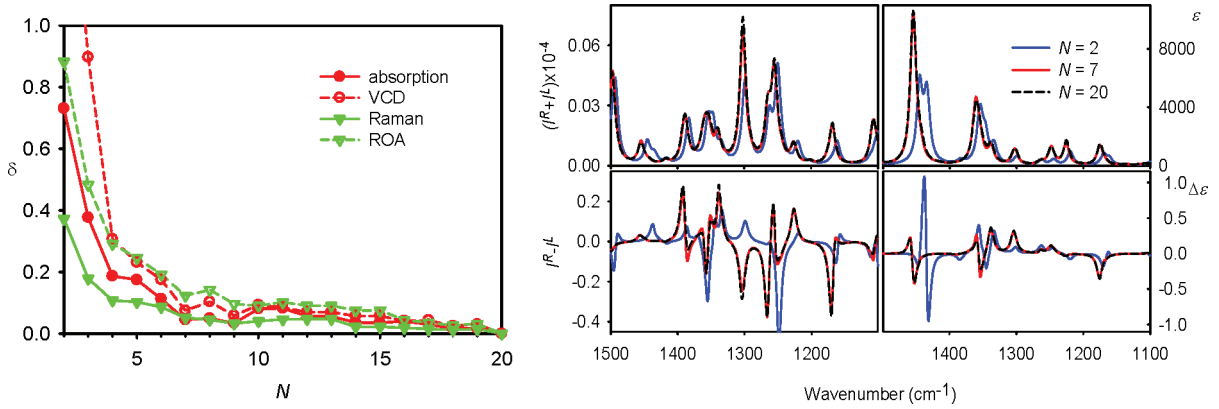

Figure 5. (Left) The dependence of the spectral errors of $(Pro)_{20}$ (PPII conformation) on the length of the $(Pro)_N$ fragments used for the CCT method. All tensors (FF, APT, AAT, $\boldsymbol{\alpha}$, $\mathbf{G}'$, and $\mathbf{A}$) were transferred. (Right) Example of the spectra, for Raman ($I^R + I^L$), ROA ($I^R - I^L$), IR (ε), and VCD ($\Delta\varepsilon$).

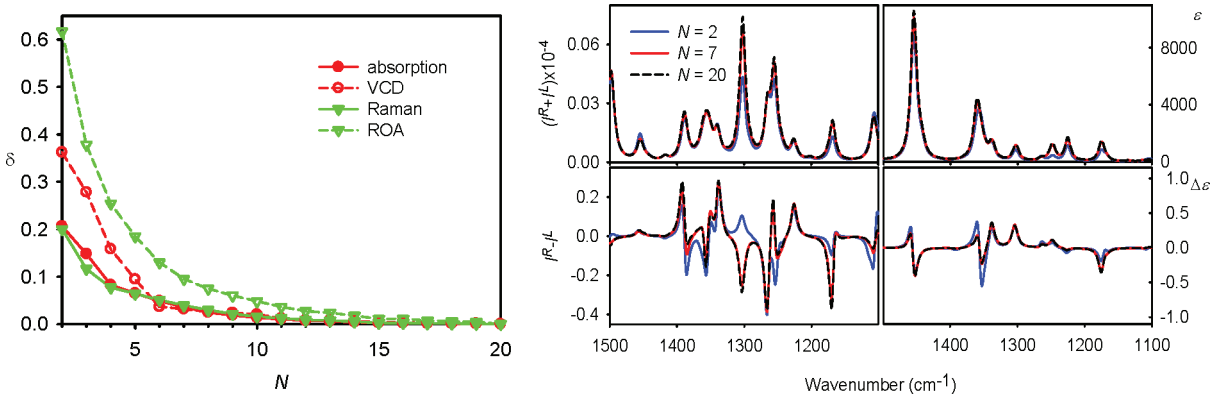

Figure 6. (Left) The dependence of the spectral errors of $(Pro)_{20}$ (PPII) on the length of the $(Pro)_N$ fragments used for the CCT method. Only intensity tensors (APT, AAT, $\boldsymbol{\alpha}$, $\mathbf{G}'$, and $\mathbf{A}$) were transferred, for the same FF. (Right) Examples of Raman ($I^R + I^L$), ROA ($I^R - I^L$), IR (ε), and VCD ($\Delta\varepsilon$) spectra.

As can be seen in panels c and d in Figure 3, the error of the $\mathbf{G}'$ and $\mathbf{A}'$ derivatives recedes more slowly than for $\boldsymbol{\alpha}$ derivatives (panel b). They also seem to recede slightly more slowly with distance than for the static polarizabilities investigated previously.[21] Nevertheless, the relative error is also small, at most ~4% on average. This can be further reduced by accounting for the polarization corrections. Interestingly, the t-correction of $\mathbf{A}$ becomes much more important for shorter separations (r < 5 Å) than the T-correction.

To verify the water dimer results and to further investigate the effect of mutual polarization on actual ROA spectra, we performed the transfer on the Ala·Ala, Ala·Val, Tyr·Tyr, and Pro·Pro dimers. For simplicity, here we use as the benchmark that includes the mutual polarization only the exact dimer computation. Indeed, as can be seen in Figure 4, the spectral error is rather small and quickly diminishes below 0.2 with amino acid separation larger than 7.5 Å. Neither the difference in the amino acid side chain nor the calculation method seems to have a large impact on the dependence. On the other hand, the Raman spectral errors are significantly smaller than for ROA. For larger distances, the error for both spectral kinds quickly becomes negligible (cf. the exact and transferred ROA spectra at the right-hand side of Figure 4 simulated for a separation of 7.5 Å). Variations in the computational method and choice of basis set do not have any significant effect on the general accuracy trends exhibited by the transfer.

We can therefore conclude that the effect of mutual polarization may be important for benchmark computations and that it can be significantly reduced by including the mutual polarization terms given in eqs 1a−3a. In practical simulations of spectra, however, it can be neglected in most cases, as it is smaller than other computational errors. For the transfers in larger molecules involving molecular fragments with overlapping regions (e.g., that in Figure 2, discussed below), an easy inclusion of the polarization correction terms is not possible. Nevertheless, in this case the general precision of the transfer can be increased more easily by an enlargement of the fragment size.

Comparison of the IR, VCD, Raman, and ROA Tensor Transfer Properties. In Figure 5 (left), the error (δ, eq 6) obtained for the $(Pro)_N \rightarrow (Pro)_{20}$ transfer is plotted as a function of fragment size for the IR, VCD, Raman, and ROA spectra. Both the intensity tensors and FF were transferred. Immediately, we see that the errors converge to zero as the fragment size increases and that, for a given fragment size N, the error is significantly larger for the chiral spectrum (based on circularly polarized light, VCD, and ROA) than for the unpolarized one (IR and Raman), as was the case for the noncovalent systems (Figure 4).

For N = 2, large errors (δ = 0.4−1.76, according to the spectral type) are obtained, and the spectra shapes deviate significantly from the benchmark values, as documented in the example spectra on the right-hand side of Figure 5. However,

dx.doi.org/10.1021/ct200714h | J. Chem. Theory Comput. 2012, 8, 977−985

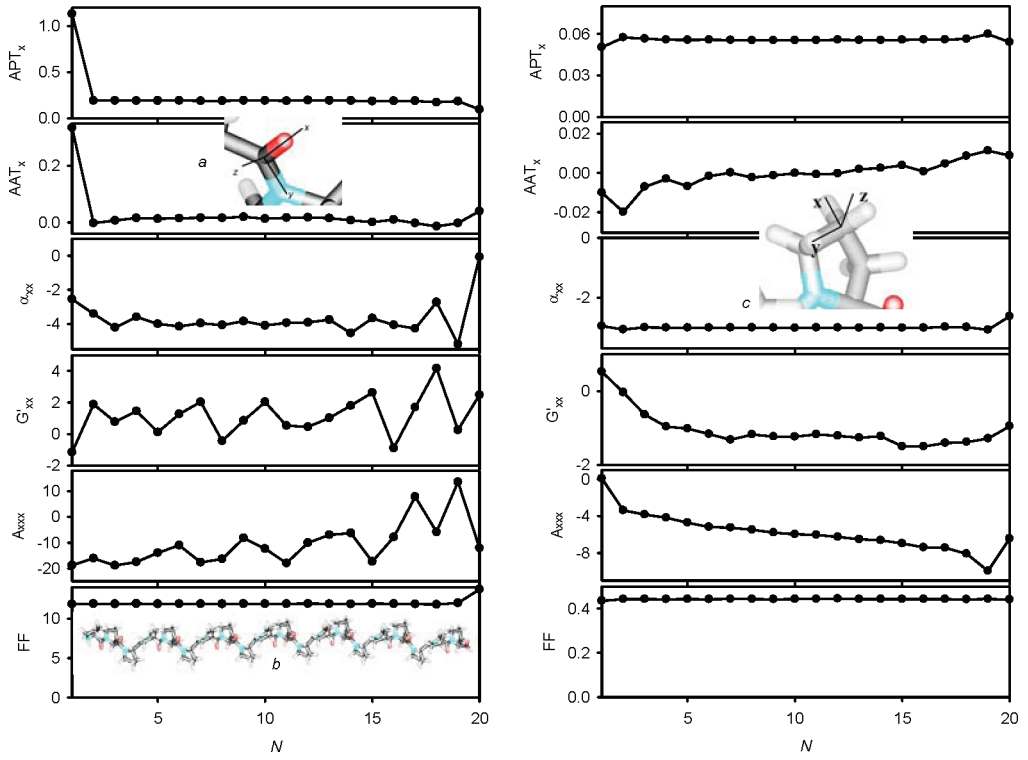

Figure 7. Local tensor components in $(Pro)_{20}$ in atomic units, as calculated at the B3LYP/6-31G** level, for (left) carbonyl and (right) proline ring carbon. From top to bottom: APT$_x$, AAT$_x$, α_{xx}, G′$_x$, and A$_{xxx}$ derivatives and FF (C═O stretching constant, left, and f_{xx} component, right). Insets (a) and (c) schematically show positions of the local coordinate systems, and the whole peptide is in inset (b).

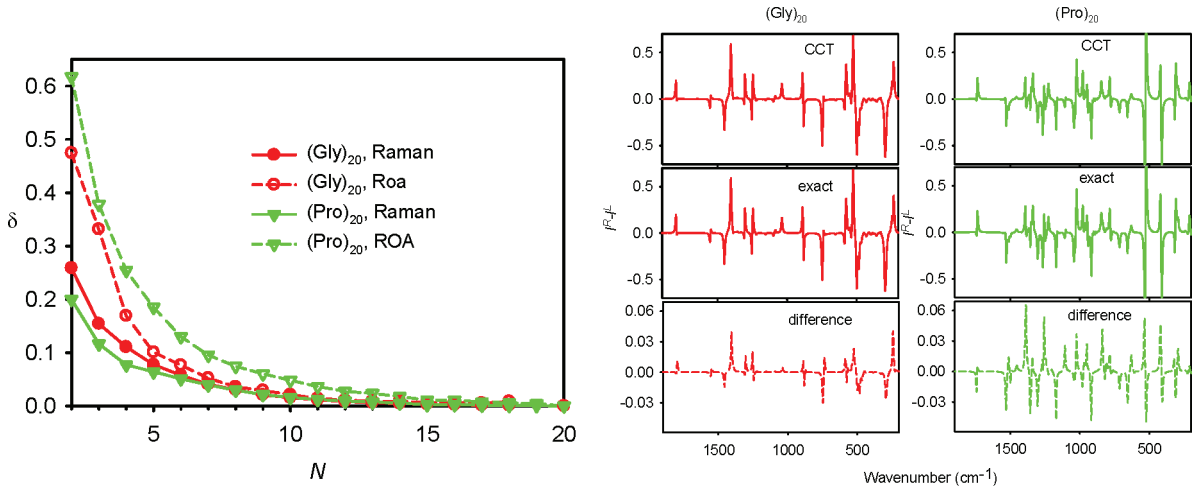

Figure 8. Amino acid chain type influence. (Left) The dependence of Raman and ROA spectral errors of $(Pro)_{20}$ and $(Gly)_{20}$ (PPII). Only intensity tensors were transferred. (Right) transferred (N = 7) and exact ROA spectra and the difference.

the error is quickly reduced with increasing N. For $N > 7$, δ becomes smaller than ~10%, so that it is difficult to visually distinguish the transferred and exact spectral shapes. All spectral types (IR, VCD, Raman, and ROA) behave similarly, and similar accuracy is achieved over the whole frequency region.

Transfer of the Force Field. Independently of whether the transfer is applied only to the intensity tensors (APT, AAT, $\boldsymbol{\alpha}$, **G′**, and **A**) or also to the FF, the convergence behavior of the error due to the transfer is the same (Figure 6). However, the overall error is reduced by about a factor or two when the FF of the 20-mer is used compared to the case when also the FF is transferred (cf. Figure 5). In other words, the FF is responsible for roughly half of the error associated with the complete transfer.

We also note another important effect of using the 20-mer FF: If only the intensity tensors are transferred (Figure 6), the error in the ROA spectra is about 2 or 3 times larger than for all the other (IR, Raman, and VCD) kinds of spectra. For N = 2, a band around 1320 cm^{-1} (right-hand side of Figure 6) even changes sign. This is most probably caused by a combination of the end effects (tensors in smaller fragments deviate from those in the longer peptide chain) and the error stemming from the mutual polarization of molecular parts (Figures 3 and 4). This partially confirms the previous observations about the exceptional properties of ROA.[20,21] Still, also the transferred ROA

dx.doi.org/10.1021/ct200714h | J. Chem. Theory Comput. 2012, 8, 977–985

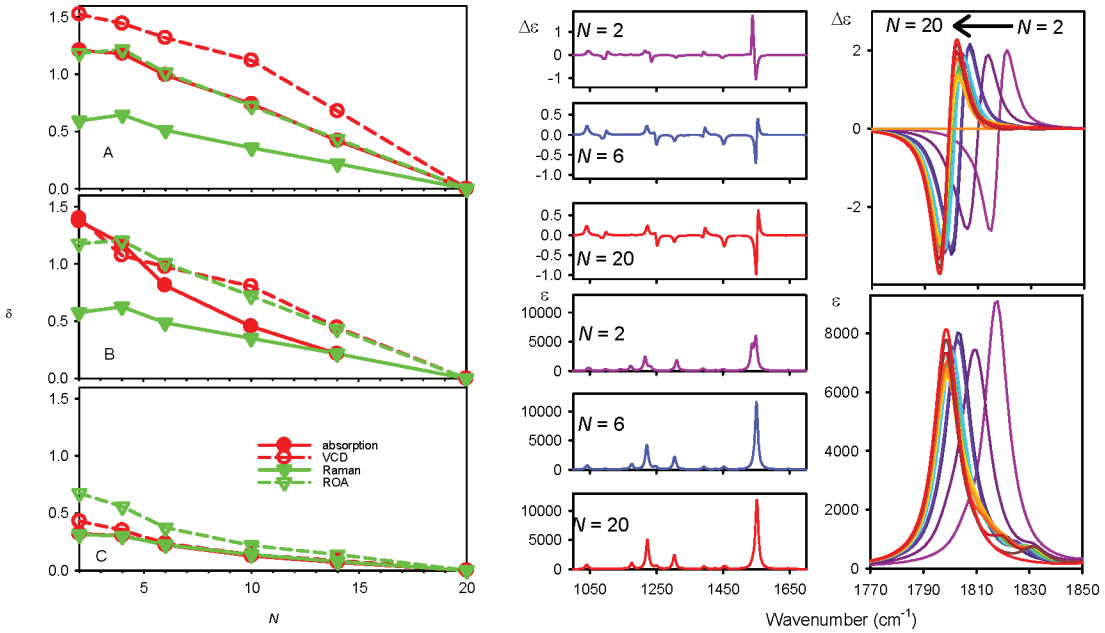

Figure 9. On the left, the dependence of the spectral errors of $(Gly)_{20}$ in the α-helical conformation on the length of the $(Gly)_N$ fragments used for the CCT method is plotted. All (A, B) and intensity only (C) tensors were transferred. For (A, C), the error was estimated in the whole frequency region and (B) shows the error within 1000−1600 cm^{-1}. On the right, examples of absorption (ε) and VCD ($\Delta\varepsilon$) are given within the 1000−1700 and 1770−1850 cm^{-1} spectral regions.

spectra quickly converge to the exact shape (right-hand side of Figure 5). The accuracy evaluation is consistent with previous studies[19] where four amino acid fragments provided satisfying spectra. Another interesting feature is that for $N > 5$, VCD can be simulated with the same precision as IR.

The high sensitivity of the ROA **G′** and **A** tensors to the local molecular neighborhood is documented in Figure 7 for the $(Pro)_{20}$ peptide. The tensors were calculated by the B3LYP/6-31G** method for partially optimized geometry at the same level. The APT, AAT, $\boldsymbol{\alpha}$, **G′**, **A**, and FF tensor components are displayed for the carboxyl group carbon (left) and carbon atoms in the proline ring (right). The intensity tensors were transferred to the carbon atoms and rotated to the local coordinate systems, so they could be directly compared. Some tensors (APT, $\boldsymbol{\alpha}$, and FF) are relatively independent of the positions in the peptide. This holds only partially for AAT, which is quite small on average but relatively large for the first carbonyl; for the proline carbon, it slowly grows as N increases, except for the terminal residues. However, **G′** and **A** components depend more on the position in the peptide chain. For carbonyl (Figure 7, left) the largest deviations from the average occur at the C terminus of the peptide ($N = 20$). For most of the other atoms the **G′** and **A** tensors exhibit a more regular behavior than for the carbonyl, for example, the proline ring atom (Figure 7, right) tensor components almost monotonically decrease with N. The carbonyl dependence is extreme because of the nonchiral amide environment (chiral tensor components are induced by interactions with other molecular parts) and carbonyl polarity. The sensitivity of the **G′** and **A** tensors to the position in the peptide chain is consistent with the larger error of ROA obtained by CCT when compared to Raman spectra. It should nevertheless be noted that the effect of the variations in the **G′** and **A** local components on the ROA spectra is rather limited, as the intensity is predominantly dependent on the nonlocal components coming from the $\boldsymbol{\alpha}$ tensor only.[1,18,27]

The Role of the Side Chain and Conformation. The error of the transferred Raman and ROA spectra for $(Gly)_{20}$ and $(Pro)_{20}$ is compared in Figure 8. The error is significantly larger for proline, for all fragment sizes. This can be partially explained by the interaction of the proline side chains. However, not only normal modes localized on the proline rings are affected. As follows from the comparison of the difference spectra (right-hand side of Figure 8), the error is larger for proline for the entire spectral range.

So far we have been concerned with the PPII conformation. However, although the dependence is similar for the more folded α-helical conformer of $(Gly)_{20}$ (Figure 9), the stronger intermolecular interactions in the α-helix between amino acid residues clearly require larger fragments to be used in order to achieve the same precision as for PPII. The errors in the absorption, and in particular VCD, spectra are especially large for the complete transfer (top left of Figure 9), as the spectra are dominated by the signal of the strongly interacting carbonyl groups. Indeed, the error is lower for the 1000−1600 cm^{-1} region not so affected by the long-range polar and hydrogen-bond interactions (middle left panel) and can be further significantly reduced by using the exact FF (bottom left). Still, the error of the CCT transfer for the α-helical conformation is larger than for the PPII conformer.

In spite of the error, reasonable spectral shapes are obtained with longer fragments, and the transferred spectrum is usable for a rough interpretation of an experiment from $N \sim 6$ (examples of IR and VCD spectra are plotted on the right-hand side of Figure 9; both the FF and intensity tensors were transferred). Note that the evaluation of the δ error based on eq 6 depends on the band shape and emphasizes large frequency shifts, occurring especially for the C=O stretching bands around 1800 cm^{-1}. Similar performance of the CCT transfer is exhibited for the Raman and ROA spectra (Figure 10), except for the lowest wavenumber region (below ~400 cm^{-1}) affected by the nonoptimized vibrational normal modes.

dx.doi.org/10.1021/ct200714h | J. Chem. Theory Comput. 2012, 8, 977−985

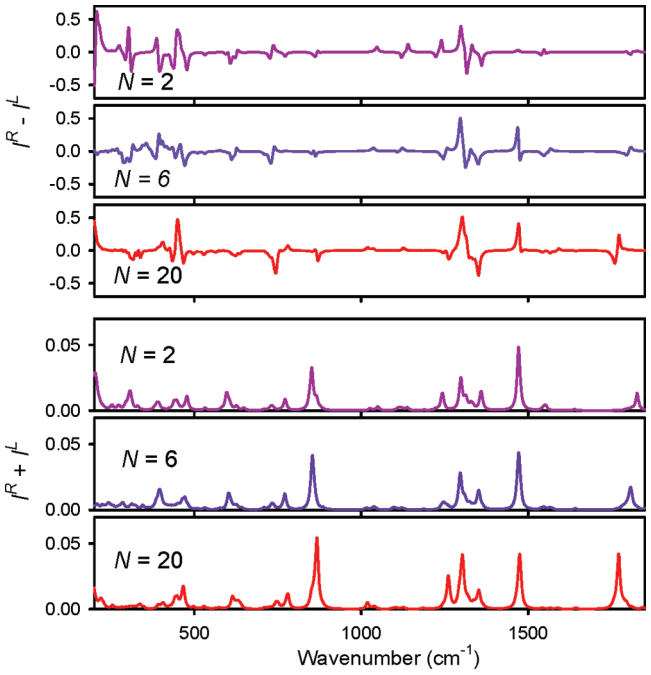

Figure 10. $(Gly)_{20}$ in the α-helical conformation, Raman ($I^R + I^L$) and ROA ($I^R - I^L$) spectra within 200−1850 cm^{-1} as simulated with the tensors from the $(Gly)_2$, $(Gly)_6$, and $(Gly)_{20}$ fragments (see also Figure 9 for the errors).

For a visual comparison of the simulations with experiment, relatively short fragments provide satisfying spectra. The example of the $(Pro)_4 \rightarrow (Pro)_{20}$ transfer (fully optimized geometry, Figure 11) documents that resultant spectral shapes

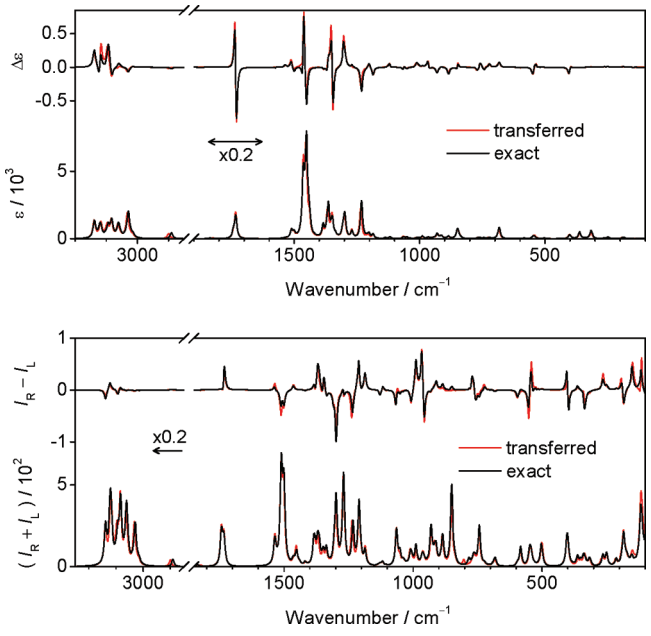

Figure 11. IR and VCD (top) and Raman and ROA (bottom) spectra of $(Pro)_{20}$ obtained by the tensor transfer from $(Pro)_4$ fragments (red, all property tensors FF, $\boldsymbol{\alpha}$, $\mathbf{G}'$, and $\mathbf{A}$ were transferred) and by the exact calculation (black) in the entire region of frequencies. The B3LYP/6-31G** level was used, for fully optimized geometry.

(IR, VCD, Raman, and ROA) are nearly identical to the exact results in the whole frequency region, at least when judged by the naked eye. The largest relative intensity deviations appear for the ROA peaks at ∼1500, ∼900, ∼550, and below 150 cm^{-1}; the ROA spectrum appears only slightly more sensitive to the transfer compared to the other spectroscopies, in agreement with previous assessments.[20,21] Interestingly, the constrained normal mode optimization of the fragments[28] which we also tried did not change the trends observed for the transfer and the agreement with the ab initio benchmarks, although they obviously make the spectra more realistic when comparing to experiment.

To summarize the results, we note that the present study did confirm some particularities of the Raman and ROA tensor transfer, namely the relatively larger error of the CCT transfer and mutual polarization if compared to IR and VCD. However, we believe that this may not be a principal obstacle for applications in large molecules, such as peptides and proteins. The contribution of the mutual polarization even for very close systems (cf. Figures 3 and 4) is quite small if compared to the total Raman and ROA signals. The error stemming from the regular covalent hydrogen-bond network represents a more serious problem, as relatively large fragments are needed for accurate results (Figures 9 and 10). Nevertheless, even in proteins with a high α-helical content, the helices are seldom regular to support the longer distance exciton effects observed for the polyglycine model. Thus the choice of moderate fragments (approximately four amino acids) used in previous works may be sufficient for approximate interpretation of experimental results. In other words, the detailed performance of CCT depends on the side chain type and secondary structure of the peptide, and the accuracy can always be improved by a choice of suitable fragmentation.

CONCLUSIONS

We have quantitatively evaluated the performance of the transfer of molecular property tensors needed to calculate IR, VCD, Raman, and ROA spectra. The evaluation confirmed that with suitably chosen fragments, the CCT method can provide faithful spectra of very large molecules with nearly ab initio quality. Additionally, the evaluation revealed finer differences in the locality of the tensors; the chiral spectroscopies (VCD, ROA) were slightly more sensitive to the transfer parameters than the unpolarized (IR, Raman) techniques. The error stemming from the FF transfer accounted for about half of the total error associated with the transfer.

An optimization of the structure did not have significant effect on transfer properties. The glycine and proline oligopeptides behaved slightly differently for ROA; however, this could not be explained by a simple through-space interaction of the proline amino acid side chains. As expected, the intramolecular hydrogen-bond interactions in the oligoglycine α-helix deteriorated the locality, in particular for the FF, and for the same accuracy, much larger fragments had to be used than for the more extended PPII conformer. A similar accuracy was observed within the entire range of wavenumbers, although individual vibrational modes showed different sensitivity to the transfer.

The analysis thus confirmed previous hypotheses that ROA spectra are slightly more sensitive to the transfer parameters than the other spectroscopies. However, our results suggest that there are no fundamental problems with the transfer technique. Indeed, our results demonstrate that with suitably chosen parameters, significant savings of computer time can be achieved. For the $(Pro)_4 \rightarrow (Pro)_{20}$ transfer, for example, the

dx.doi.org/10.1021/ct200714h | J. Chem. Theory Comput. 2012, 8, 977−985

computational time needed for the ab initio tensor calculation (1080 h, 2 GHz CPU) could be reduced by a factor of 277.

AUTHOR INFORMATION

Corresponding Author

*E-mail: aporoa@gmail.com; bour@uochb.cas.cz.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This study was performed with the support from the Academy of Sciences (M200550902), the KONTAKT II MSMT program (LH11033), Grant Agency of the Czech Republic (P208/11/0105), Luna, Metacentrum, and University of Tromsø computer facilities. X.L. and K.R. have been supported by the Research Council of Norway through a Centre of Excellence Grant (grant no. 179568/V30) and a research grant (grant no. 191251).

REFERENCES

(1) Barron, L. D. *Molecular Light Scattering and Optical Activity*; Cambridge University Press: Cambridge, U.K., 2004.

(2) Kubelka, J.; Bouř, P.; Keiderling, T. A. Quantum Mechanical Calculations of Peptide Vibrational Force Fields and Spectral Intensities. In *Advances in Biomedical Spectroscopy, Biological and Biomedical Infrared Spectroscopy*, Barth, A., Haris, P. I., Eds.; IOS Press: Amsterdam, The Netherlands, 2009; Vol. *2*, pp 178−223.

(3) Abdali, S.; Jalkanen, K. J.; Cao, X.; Nafie, L. A.; Bohr, H. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2434−2439. Silva, R. A. G. D.; Kubelka, J.; Decatur, S. M.; Bouř, P.; Keiderling, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8318−8323.

(4) Zhong, W.; Gulotta, M.; Goss, D. J.; Diem, M. *Biochemistry* **1990**, *29*, 7485−7491. Blanch, E. W.; Hecht, L.; Barron, L. D. *Methods* **2003**, *29*, 196−209.

(5) Bell, A. F.; Hecht, L.; Barron, L. D. *Chem.—Eur. J.* **1997**, *3*, 1292−1298.

(6) Stephens, P. J.; Ashvar, C. S.; Devlin, F. J.; Cheeseman, J. R.; Frisch, M. J. *Mol. Phys.* **1996**, *89*, 579−594.

(7) Ruud, K.; Helgaker, T.; Bouř, P. *J. Phys. Chem. A* **2002**, *106*, 7448−7455. Haesler, J.; Schindelholz, I.; Riguet, E.; Bochet, C. G.; Hug, W. *Nature* **2007**, *446*, 526−529.

(8) Ruud, K.; Thorvaldsen, J. *Chirality* **2009**, *21*, E54−E67. Nicu, V. P.; Autschbach, J.; Baerends, E. J. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1526−1538.

(9) Ahlrichs, R.; Bar, M.; Baron, H.-P.; Bauernschmitt, R.; Bocker, S.; Ehrig, M.; Eichkorn, K.; Elliot, S.; Furche, F.; Haase, F.; Haser, M.; Horn, H.; Huber, C.; Huniar, U.; Kattannek, M.; Kolmel, C.; Koolwitz, M.; May, K.; Ochsenfeld, C.; Ohm, H.; Schafer, A.; Schneider, U.; Treutler, O.; von Arnim, M.; Weigend, F.; Weis, P.; Weiss, H. *Turbomole*, version 6; Quantum Chemistry Group, University of Karlsruhe: Karlsruhe, Germany, 1998; Angeli, C.; Bak, K. L.; Bakken, V.; Christiansen, O.; Cimiraglia, R.; Coriani, S.; Dahle, P.; Dalskov, E. K.; Enevoldsen, T.; Fernandez, B.; Haettig, C.; Hald, K.; Halkier, A.; Heiberg, H.; Helgaker, T.; Hettema, H.; Jensen, H. J. A.; Jonsson, D.; Joergensen, P.; Kirpekar, S.; Klopper, W.; Kobayashi, R.; Koch, H.; Lutnaes, O. B.; Mikkelsen, K. V.; Norman, P.; Olsen, J.; Packer, M. J.; Pedersen, T. B.; Rinkevicius, Z.; Rudberg, E.; Ruden, T. A.; Ruud, K.; Salek, P.; Sanchez de Meras, A.; Saue, T.; Sauer, S. P. A.; Schimmelpfennig, B.; Sylvester-Hvid, K. O.; Taylor, P. R.; Vahtras, O.; Wilson, D. J.; Agren, H. *Dalton, a molecular electronic structure program*, release 2.0; University of Oslo: Oslo, Norway, 2005−2009.

(10) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J., J., A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision A.02; Gaussian, Inc.: Wallingford, CT, 2009.

(11) Luber, S.; Reiher, M. *J.Phys. Chem. B* **2010**, *114*, 1057−1063.

(12) Zuber, G.; Hug, W. *J. Phys. Chem. A* **2004**, *108*, 2108−2118.

(13) Cheeseman, J. R.; Frisch, M. J. *J. Chem. Theory Comput.* **2011**, *7*, 3323−3334.

(14) Bouř, P.; Sopková, J.; Bednárová, L.; Maloň, P.; Keiderling, T. A. *J. Comput. Chem.* **1997**, *18*, 646−659.

(15) Kubelka, J.; Keiderling, T. A. *J. Am. Chem. Soc.* **2001**, *123*, 12048−12058.

(16) Andrushchenko, V.; Wieser, H.; Bouř, P. *J. Phys. Chem. B* **2002**, *106*, 12623−12634.

(17) Bouř, P.; Keiderling, T. A. *J. Phys. Chem. B* **2005**, *109*, 23687−23697. Kubelka, J. IR and VCD spectroscopy of model peptides. *Theory and Experiment*. Ph.D. Thesis; University of Illinois at Chicago: Chicago, IL, 2002.

(18) Kapitán, J.; Baumruk, V.; Kopecký, V. Jr.; Bouř, P. *J. Am. Chem. Soc.* **2006**, *128*, 2438−2443.

(19) Yamamoto, S.; Straka, M.; Watarai, H.; Bouř, P. *Phys. Chem. Chem. Phys.* **2010**, *12*, 11021−11032. Yamamoto, S.; Watarai, H.; Bouř, P. *ChemPhysChem* **2011**, *12*, 1509−1518.

(20) Bieler, N. S.; Haag, M. P.; Jacob, C. R.; Reiher, M. *J. Chem. Theory Comput.* **2011**, *7*, 1867−1881.

(21) Yamamoto, S.; Bouř, P. *Collect. Czech. Chem. Commun.* **2011**, *76*, 567−583.

(22) Stephens, P. J.; Lowe, M. A. *Annu. Rev. Phys. Chem.* **1985**, *36*, 213−241.

(23) Kapitán, J.; Johannessen, C.; Bouř, P.; Hecht, L.; Barron, L. D. *Chirality* **2009**, *21*, E4−E12.

(24) Bouř, P. *CCT*, program for Cartesian tensor transfer; Academy of Sciences: Prague, Czech Republic, 2009.

(25) Dukor, R. K.; Keiderling, T. A. *Biopolymers* **1991**, *31*, 1747−1761. Dukor, R. K.; Keiderling, T. A.; Gut, V. *Int. J. Pept. Protein Res.* **1991**, *38*, 198−203.

(26) Creighton, T. E. *Proteins: Structures and Molecular Properties*, 2nd ed.; W. H. Freeman and Co.: New York, 1993.

(27) Bouř, P.; Baumruk, V.; Hanzlíková, J. *Collect. Czech. Chem. Commun.* **1997**, *62*, 1384−1395.

(28) Bouř, P.; Keiderling, T. A. *J. Chem. Phys.* **2002**, *117*, 4126−4132. Bouř, P. *Collect. Czech. Chem. Commun.* **2005**, *70*, 1315−1340.