JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Analysis of the Cartesian Tensor Transfer Method for Calculating Vibrational Spectra of Polypeptides

Noah S. Bieler,[†] Moritz P. Haag,[†] Christoph R. Jacob,[*,‡] and Markus Reiher[*,†]

[†]ETH Zurich, Laboratorium für Physikalische Chemie, Wolfgang-Pauli-Strasse 10, 8093 Zurich, Switzerland

[‡]Karlsruhe Institute of Technology (KIT), Center for Functional Nanostructures, Wolfgang-Gaede-Str. 1a, 76131 Karlsruhe, Germany

Supporting Information

ABSTRACT: The Cartesian Tensor Transfer Method (CTTM) was proposed as an efficient way to calculate infrared, Raman, and Raman Optical Activity (ROA) spectra for large molecules from the Hessian matrix and property tensor derivatives calculated for smaller molecular fragments. Although this approach has been widely used, its reliability has not been analyzed in depth yet. Especially for ROA spectra, such an analysis became only recently possible because of methodological advances that allow for the calculation of full ROA spectra of fairly large molecules with large basis sets. In this work, we investigate an α-helical polypeptide of 20 alanine amino acids, for which we reported the full ROA spectra earlier, in order to study the CTTM for protein subunits. By comparing the full first-principles calculation of the vibrational spectra with spectra reconstructed with the CTTM from different fragment sizes, we find that infrared and Raman spectra are mostly well reproduced. However, this is not the case for the ROA spectrum. This might have implications for peptide and protein CTTM ROA spectra that have already been published in the literature.

1. INTRODUCTION

Vibrational spectra are a widely used tool to study biomolecules, in particular, polypeptides and proteins. Conventional infrared (IR) and Raman spectroscopy,[1−4] their chiral variants, vibrational circular dichroism (VCD)[5−8] and Raman optical activity (ROA),[9−14] as well as other specialized techniques such as resonance Raman spectroscopy[15−18] or multidimensional IR spectroscopy[19−21] can provide detailed insight into the structure and dynamics of biomolecules in solution. Biomolecules in the gas phase can be investigated using infrared multiple-photon dissociation (IR-MPD) spectroscopy[22−24] and IR-UV double resonance spectroscopy.[25,26]

However, the interpretation of experimental vibrational spectra is hampered by the difficulty of establishing a direct relationship between observed spectra and the molecular structure. Therefore, first-principles calculations of the vibrational spectra (for a review, see ref 27) are often necessary to unambiguously assign vibrational spectra. Such calculations can provide information that is not available from experimental results alone, such as the precise atomic displacements corresponding to each of the observed vibrational transitions. In combination with appropriate analysis tools,[28−30] this can be utilized to understand the relationship between molecular structure and vibrational spectra in detail.[31−34]

The computational cost for such first-principles calculations is very high because of the higher derivatives of the total electronic energy and of molecular property tensors that are required for the vibrational frequencies and intensities, respectively. Hence, more efficient approaches for the calculation of vibrational spectra have been developed. For instance, the mode-tracking[35−37] and intensity-tracking[38−41] methods allow for the direct calculation of particular normal modes or of only the intense modes. For the targeted modes, these approaches give results identical to those of a full calculation.

In addition, methods that introduce further approximations have been investigated.[42,43] An interesting and popular approximation is the so-called Cartesian Tensor Transfer Method (CTTM) proposed by Bouř et al. in 1997.[44] The CTTM constructs the second derivatives of the electronic energy (i.e., the Hessian matrix) and the intensities of a large molecule from the Hessian matrices and molecular property tensor derivatives calculated for smaller molecular fragments. Obviously, depending on the size of these fragments, this approximation neglects long-range interactions between atoms in the large molecule, which are further apart from one another so that they do not belong to the same fragment. Since information about these long-range interactions is not contained in the properties of the smaller fragments, it cannot be accounted for. In terms of the underlying Hessian matrix, this means that the far-off diagonal elements are set to zero. Similarly, the use of small fragments might significantly alter the electronic structure and in turn the property tensor derivatives determining the vibrational intensities.

The CTTM has been applied in numerous studies to calculate IR[45−49] and Raman[45,50] spectra of biomolecules. Also for the corresponding chiral analogues VCD[7,8,51−66] and ROA,[50,67−74] the CTTM has been applied extensively. An example is a study by Kapitan et al., who employed the CTTM for β-peptides,[70] where both the force field (i.e., the Hessian matrix) and the polarizability tensors' derivatives were constructed with the CTTM. In that work, the β-peptide under study was decomposed into fragments separately representing the backbone and the side

Received: February 28, 2011
Published: May 06, 2011

chain, and the backbone was further decomposed into two fragments. In another study,[67] Kapitan et al. calculated Raman and ROA spectra of poly-L-proline with the CTTM from HCO-(L-Pro)$_3$-NH$_2$ fragments.

Despite its wide use, the CTTM has never been analyzed in detail for such large biomolecules by performing a comparison to full calculations. Mostly, this is because full calculations for such large systems have only become possible in recent years. In particular, for ROA spectroscopy, full calculations are now possible for rather large polypeptides[31,75,76] and even for small proteins[77] by using efficient density-fitting techniques for the calculation of the required polarizability tensors.[78] In the future, analytical derivative methods developed for Raman[79,80] and ROA spectroscopy[81−84] will push the limits even further.

In this work, we intend to close this gap by investigating the accuracy and reliability of the CTTM in detail. As a test system, we chose an α-helical polypeptide consisting of 20 alanine residues. To allow for a direct comparison, all calculations for the large molecules and the corresponding smaller fragments are carried out with the same basis set and exchange-correlation functional. Only the fragment size was varied in order to assess its influence. As a consequence, we should be able to determine the limits of this method and the reliability of the spectra of polypeptides calculated with the CTTM.

This work is organized as follows. In section 2, we review the theory of calculating vibrational spectra as well as the CTTM and outline our CTTM implementation, before explaining the computational details in section 3. This is followed in section 4 by two tests on small molecules in order to validate our implementation of the CTTM before we proceed to a detailed analysis of the CTTM for an α-helical polypeptide in section 5. Finally, conclusions are presented in section 6.

2. METHODOLOGY

Within the harmonic approximation, the vibrational frequencies ν_p and the normal modes $\mathbf{Q}_p$ can be obtained by diagonalizing the mass-weighted molecular Hessian matrix $\mathbf{H}^{(m)}$ with the elements

$$H^{(m)}_{i\alpha,j\beta} = \frac{1}{\sqrt{m_i m_j}}\left(\frac{\partial^2 E}{\partial R_{i\alpha}\partial R_{j\beta}}\right)_0 \quad (1)$$

where E is the total energy, $R_{i\alpha}$ is the $\alpha = x,y,z$ Cartesian component of the position of nucleus i, m_i is the atomic mass of nucleus i, and the subscript "0" indicates that the derivative is taken at the molecular equilibrium structure R_0. Here and in the following, we will use Greek indices $\alpha,\beta,\gamma,\delta$, etc. to refer to the Cartesian components x,y,z of a vector or tensor and indices i,j,k, etc. for nuclei.

The intensities for infrared, Raman, and ROA spectroscopy corresponding to each vibrational transition can be expressed as

$$I_p \propto \sum_l c_l \left(\frac{\partial P^{(i)}_l}{\partial \mathbf{Q}_p}\right)_0 \left(\frac{\partial P^{(i)}_l}{\partial \mathbf{Q}_p}\right)_0 \quad (2)$$

where c_l are coefficients depending on the type of spectroscopy as well as the experimental setup and $P^{(i)}_l$ are components of the appropriate property tensors. In the case of infrared spectroscopy, these are the components of the dipole moment μ, and for Raman spectroscopy, those of the electric-dipole−electric-dipole polarizability tensor α. For ROA spectroscopy, the electric-dipole−magnetic-dipole polarizability tensor $\mathbf{G}'$ and the electric-dipole−electric-quadrupole polarizability tensor $\mathbf{A}$ are also required. For explicit intensity expressions of these different types of vibrational spectroscopy, see, e.g., Appendix D in ref 29. The derivatives with respect to normal modes in eq 2 can be obtained from the derivatives of the property tensors with respect to Cartesian displacements $(\partial P^{(i)}_l/\partial R_{i\alpha})_0$.

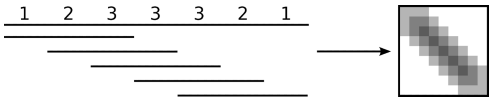

Figure 1. An abstract example for defining the overlaps: A large polymer (heptamer) is reconstructed from a fragment (trimer). The numbers above the bar show how many overlaps exist for one unit of the large polymer. An approximation of the Hessian matrix of the large molecule is then constructed with the ones of the fragments as illustrated on the right.

For large molecular systems, such as polypeptides, the calculation of the second energy derivatives in eq 1 and of the property tensor derivatives becomes a very expensive step, in particular for calculations of ROA spectra. Therefore, in the CTTM of Bouř et al., these derivatives are approximated by those calculated for smaller fragments. In the following, we briefly outline the CTTM as well as the details of our implementation.

In the simplest case, one considers a molecule constituted of several identical monomers such as, for instance, an alanine polypeptide. In this case, a smaller fragment of only a few monomers is used to perform the actual calculations of the Hessian matrix and the property tensor derivatives, and these results are then transferred to the larger molecule. A schematic example is illustrated in Figure 1. The large molecule is in this case a heptamer, and as a small fragment a trimer is used. The small fragment can then be moved along the large fragment, and for each position, a mapping between atoms of the small fragment and atoms of the large molecule can be defined (following the terminology of ref 44, this mapping will be called an *overlap*). As shown in the figure, the atoms in the central part of the large molecule can be mapped to the smaller fragment in several different ways; i.e., different overlaps are possible. Of course, by using several different small fragments, such a procedure can be generalized to polymers consisting of different monomers and to entirely general large molecules.

The Hessian matrix is then calculated in the following way: For each pair of atoms i and j, one finds those overlaps that map both atoms to the same small fragment. If no such overlap is found, the corresponding elements of the Hessian are set to zero. If one or more overlaps are found, the corresponding elements of the Hessian of the small fragment are transferred to the large molecule (see below for details). This results in a band-diagonal Hessian matrix, as is illustrated in Figure 1. Initially, Bouř et al. suggested to fill the gaps where the Hessian has been set to zero with entries of a semiempirical or empirical force field.[44] However, since in most of the published papers so far this possibility was not used, we will only investigate the generic method here as a true *ab initio* method without any empirical contributions.

To transfer the Hessian matrix elements from the small fragment to the large molecule, one has to determine a rotation matrix $\mathbf{U}^{(ij)}$ that maps the relevant atoms of the small fragment to the corresponding atoms of the large molecule. For this, one considers the atoms i and j as well as their next neighbors (as defined by a connectivity table). If this results in less than three atoms, another set of next neighbors is included. Then, both sets of atoms $\{\mathbf{R}^{(\mathrm{large})}_k\}$ and $\{\mathbf{R}^{(\mathrm{small})}_k\}$ are translated to their

geometric centers ($\overline{\mathbf{R}}^{(ij)}_{\text{large}} = 1/N\ \Sigma_k \mathbf{R}^{(\text{large})}_k$ and $\overline{\mathbf{R}}^{(ij)}_{\text{small}} = 1/N\ \Sigma_k \mathbf{R}^{(\text{small})}_k$, respectively, where the index k runs over the N relevant atoms). Subsequently, the rotation matrix $\mathbf{U}^{(ij)}$ is determined such that the mean square error

$$\delta(\mathbf{U}^{(ij)}) = \sum_k |(\mathbf{R}^{(\text{large})}_k - \overline{\mathbf{R}}^{(ij)}_{\text{large}}) - \mathbf{U}^{(ij)}(\mathbf{R}^{(\text{small})}_k - \overline{\mathbf{R}}^{(ij)}_{\text{small}})|^2 \quad (3)$$

is minimized, where the index k runs over the relevant atoms. For this minimization, one can apply different methods. Some methods employ a Euler angle parametrization, which can raise different problems such as the Gimbal lock;[85,86] others employ a quaternion parametrization.[85,86] The latter has the advantage that it allows one to apply algebraic methods to solve the minimization problem, which is then turned into an eigenvalue problem, rather than to deal with trigonometric functions. For these reasons, an algorithm based on quaternions, similar to the one described in ref 87, is employed in this work.

It is important to note that a different rotation matrix $\mathbf{U}$ is determined for each pair of atoms i and j, which is indicated by the superscript index "(ij)". Alternatively, it would also be possible to apply a common rotation matrix for each overlap of a small fragment with the larger molecule. However, by readjusting the rotation such that the agreement is optimal for the atoms i and j and their neighbors, a more accurate Hessian (and property tensor derivatives) should be obtained. Finally, if there are several overlaps that contain the atom pair i and j, one has to decide which of them (i.e., which small fragment) is used. The papers of Bouř and co-workers mention that there are several possibilities for handling such cases[44] (e.g., always choosing the fragment for which the geometry matches best or performing a, possibly weighted, average). However, they are not very clear in explaining which option has actually been used in their calculations. On the basis of our own tests, we chose to always use the overlap for which the smallest error $\delta(\mathbf{U}^{(ij)})$ is obtained.

Given the rotation matrix $\mathbf{U}$, the entries of the Hessian matrix for the atom pair i and j can be transformed with[44]

$$\frac{\partial^2 E^{(\text{large})}}{\partial R_{i\alpha}\partial R_{j\beta}} = U^{(ij)}_{\alpha\gamma} U^{(ij)}_{\beta\delta} \frac{\partial^2 E^{(\text{small})}}{\partial R_{i\gamma}\partial R_{j\delta}} \quad (4)$$

The usual sum convention is used throughout (i.e., a Greek index occurring more than once on the right-hand side is summed over all Cartesian components) if no explicit sum sign is used.

For the spectral intensities, one has to obtain the different derivatives of electric property tensors (i.e., the electric-dipole polarizability μ, the electric-dipole−electric-dipole polarizability tensor α, the electric-quadrupole−electric-dipole polarizability $\mathbf{A}$, and the electric-dipole−magnetic-dipole polarizability $\mathbf{G}'$) with respect to Cartesian coordinates. These are also obtained by transferring the property tensor derivatives from the small fragments to the large molecule. However, since these derivatives only depend on one atom i (in contrast to the elements of the Hessian, which depend on two atoms), the rotation matrices $\mathbf{U}^{(ii)}$ found for the diagonal elements of the Hessian can be used.

The dipole moment μ and the electric-dipole−electric-dipole polarizability α of an uncharged molecule are not origin-dependent and transform as[44]

$$\frac{\partial}{\partial R_{i\varepsilon}} \mu^{(\text{large})}_\alpha = U^{(ii)}_{\varepsilon\eta} U^{(ii)}_{\alpha\pi} \frac{\partial}{\partial R_{i\eta}} \mu^{(\text{small})}_\pi \quad (5)$$

$$\frac{\partial}{\partial R_{i\varepsilon}} \alpha^{(\text{large})}_{\alpha\beta} = U^{(ii)}_{\varepsilon\eta} U^{(ii)}_{\alpha\pi} U^{(ii)}_{\beta\rho} \frac{\partial}{\partial R_{i\eta}} \alpha^{(\text{small})}_{\pi\rho} \quad (6)$$

The electric-dipole−magnetic-dipole polarizability $\mathbf{G}'$ and the electric-quadrupole−electric-dipole polarizability $\mathbf{A}$ change under a gauge transformation (i.e., a shift of the origin).[9,78] Therefore, for transferring their derivatives from the small fragments to the large molecule, it is not sufficient to transform them with the rotation matrix $\mathbf{U}^{(ii)}$, but the translation by $-\overline{\mathbf{R}}^{(ii)}_{\text{small}}$ applied before the rotation and the translation by $\overline{\mathbf{R}}^{(ii)}_{\text{large}}$ applied after the rotation have also to be considered. This way, it is ensured that the property tensor derivatives for the large molecule all refer to the same common origin. This is only relevant for the calculation of ROA spectra, but (as long as neutral molecules are studied) not for IR or Raman spectra.

The first step is a translation of the small fragment to its geometric center, i.e., by the vector $-\overline{\mathbf{R}}^{(ii)}_{\text{small}}$. This corresponds to shifting the origin $\mathbf{O}$ to $\mathbf{O} + \overline{\mathbf{R}}^{(ii)}_{\text{small}}$. Under this transformation, the Cartesian derivatives of $\mathbf{G}'$ and $\mathbf{A}$ change as[9,44,78]

$$\frac{\partial}{\partial R_{i\varepsilon}} G'^{(\text{small})}_{\alpha\beta}(\mathbf{O} + \overline{\mathbf{R}}^{(ii)}_{(\text{small})}) = \frac{\partial}{\partial R_{i\varepsilon}} G'^{(\text{small})}_{\alpha\beta}(\mathbf{O}) + \frac{1}{2}\omega\varepsilon_{\beta\gamma\delta}\overline{R}^{(ii)}_{\text{small},\gamma} \frac{\partial}{\partial R_{i\varepsilon}} \alpha^{(\text{small})}_{\alpha\delta} \quad (7)$$

$$\frac{\partial}{\partial R_{i\varepsilon}} A^{(\text{small})}_{\alpha,\beta\gamma}(\mathbf{O} + \overline{\mathbf{R}}^{(ii)}_{(\text{small})}) = \frac{\partial}{\partial R_{i\varepsilon}} A^{(\text{small})}_{\alpha,\beta\gamma}(\mathbf{O}) - \frac{3}{2}\overline{R}^{(ii)}_{\text{small},\beta} \frac{\partial}{\partial R_{i\varepsilon}} \alpha^{(\text{small})}_{\alpha\gamma} - \frac{3}{2}\overline{R}^{(ii)}_{\text{small},\gamma} \frac{\partial}{\partial R_{i\varepsilon}} \alpha^{(\text{small})}_{\alpha\beta} + \delta_{\beta\gamma}\overline{R}^{(ii)}_{\text{small},\delta} \frac{\partial}{\partial R_{i\varepsilon}} \alpha^{(\text{small})}_{\alpha\delta} \quad (8)$$

where ω is the angular frequency of the incident light, $\delta_{\beta\gamma}$ is the Kronecker delta, and $\varepsilon_{\beta\gamma\delta}$ is the Levi−Civita symbol. After this transformation, the rotation $\mathbf{U}^{(ii)}$ has to be applied:

$$\frac{\partial}{\partial R_{i\varepsilon}} G'^{(\text{large})}_{\alpha\beta}(\mathbf{O} + \overline{\mathbf{R}}^{(ii)}_{(\text{large})}) = U^{(ii)}_{\varepsilon\eta} U^{(ii)}_{\alpha\pi} U^{(ii)}_{\beta\rho} \frac{\partial}{\partial R_{i\eta}} G'^{(\text{small})}_{\pi\rho}(\mathbf{O} + \overline{\mathbf{R}}^{(ii)}_{(\text{small})}) \quad (9)$$

$$\frac{\partial}{\partial R_{i\varepsilon}} A^{(\text{large})}_{\alpha,\beta\gamma}(\mathbf{O} + \overline{\mathbf{R}}^{(ii)}_{(\text{large})}) = U^{(ii)}_{\varepsilon\eta} U^{(ii)}_{\alpha\pi} U^{(ii)}_{\beta\rho} U^{(ii)}_{\gamma\sigma} \frac{\partial}{\partial R_{i\eta}} A^{(\text{small})}_{\pi,\rho\sigma}(\mathbf{O} + \overline{\mathbf{R}}^{(ii)}_{(\text{small})}) \quad (10)$$

Following this rotation around the origin, the polarizability tensor derivatives refer to the large fragment that has been shifted such that its geometric center is at the origin, i.e., by $-\overline{\mathbf{R}}^{(ii)}_{\text{large}}$. Thus, the $\mathbf{G}'$ tensor and the $\mathbf{A}$ tensor have to be translated back to the original position of the large molecule. This corresponds to shifting the origin from $\mathbf{O} + \overline{\mathbf{R}}^{(ii)}_{(\text{large})}$ to $\mathbf{O}$ and therefore,

$$\frac{\partial}{\partial R_{i\varepsilon}} G'^{(\text{large})}_{\alpha\beta}(\mathbf{O}) = \frac{\partial}{\partial R_{i\varepsilon}} G'^{(\text{large})}_{\alpha\beta}(\mathbf{O} + \overline{\mathbf{R}}^{(ii)}_{(\text{large})}) - \frac{1}{2}\omega\varepsilon_{\beta\gamma\delta}\overline{R}^{(ii)}_{\text{large},\gamma} \frac{\partial}{\partial R_{i\varepsilon}} \alpha^{(\text{large})}_{\alpha\delta} \quad (11)$$

$$\frac{\partial}{\partial R_{i\varepsilon}} A^{(\text{large})}_{\alpha,\beta\gamma}(\mathbf{O}) = \frac{\partial}{\partial R_{i\varepsilon}} A^{(\text{large})}_{\alpha,\beta\gamma}(\mathbf{O} + \overline{\mathbf{R}}^{(ii)}_{(\text{large})}) + \frac{3}{2}\overline{R}^{(ii)}_{\text{large},\beta} \frac{\partial}{\partial R_{i\varepsilon}} \alpha^{(\text{large})}_{\alpha\gamma} + \frac{3}{2}\overline{R}^{(ii)}_{\text{large},\gamma} \frac{\partial}{\partial R_{i\varepsilon}} \alpha^{(\text{small})}_{\alpha\beta} - \delta_{\beta\gamma}\overline{R}^{(ii)}_{\text{large},\delta} \frac{\partial}{\partial R_{i\varepsilon}} \alpha^{(\text{large})}_{\alpha\delta} \quad (12)$$

```
calculate Hessians and property tensor derivatives for all small fragments

for all atoms i in large molecule:
    for all atoms j in large molecule:

        find the relevant atoms
        [i, j, and their neighbors, at least three atoms]

        for all overlaps (i.e., small fragments):

            if atoms i and j are mapped to this small fragment:

                find relevant atoms in small fragment
                determine translation vectors R̄^(ij)_large and R̄^(ij)_small
                determine rotation matrix U^(ii)
                calculate mean−square error δ(U^(ii))

        find best overlap [with smallest δ(U^(ii))]

        if no overlap found :
            set H^(large)_(iα,jβ) = 0
        else :
            calculate H^(large)_ij with eq 4
            if i == j:
                calculate dipole derivatives ∂μ/∂R_iα with eq 5
                calculate polarizability derivatives ∂α/∂R_iα with eq 6

                calculate G−tensor derivatives ∂G′/∂R_iα with eqs 7, 9, 11
                calculate A−tensor derivatives ∂A/∂R_iα with eqs 8, 10, 12

calculate vib. frequencies and normal modes of large molecule from H^(large)
calculate IR, Raman, and ROA intensities from property tensor derivatives
```

Figure 2. Pseudocode for the calculation of vibrational spectra of large molecules according to the CTTM.

Once all (pairs of) atoms of the large molecule have been considered, the IR, Raman, and ROA spectra can be computed from the final Hessian matrix and property tensor derivatives using the usual procedure.

3. COMPUTATIONAL DETAILS

All calculations of Hessians and property tensor derivatives have been performed using the SNF program.[88,89] The Hessian matrix is calculated by numerical differentiation of analytical gradients calculated with density-functional theory in the Turbomole program package.[90] For the calculation of the property tensor derivatives, the dipole moment as well as the electric-dipole−electric-dipole, electric-dipole−magnetic-dipole, and the electric-dipole−electric-quadrupole polarizability tensors calculated with Turbomole are differentiated numerically. The polarizability tensors are obtained with time-dependent density-functional theory from a modified version of Turbomole's escf program.[78] The electric-dipole−electric-dipole polarizability α is calculated both in the length and in the velocity representation, the electric-dipole−magnetic-dipole polarizability $\mathbf{G}'$ is calculated in the velocity representation, and the electric-dipole−electric-quadrupole polarizability $\mathbf{A}$ is calculated in the length representation. All Turbomole calculations employ the BP86 exchange-correlation functional[91,92] and Ahlrichs' valence triple-ζ basis with one set of polarization functions (def-TZVP)[93,94] and the corresponding auxiliary basis sets.[95,96]

The CTTM method as described in the previous section has been implemented in an add-on package to SNF written in the Python programming language. A pseudocode representation of the calculation of vibrational spectra with our implementation of the CTTM is shown in Figure 2. This add-on reads the Hessian matrices and property tensor derivatives calculated by SNF for the small fragments and assembles the Hessian matrix and property tensor derivatives of the large molecule. For determining the rotation matrix $\mathbf{U}$, a quaternion-based algorithm[86,87] is used. For this step, our program makes use of the routines provided by the PyVib2 library of Fedorovsky.[97]

The Hessian matrix and the property tensor derivatives constructed using the CTTM are then read back into the SNF program, so that the usual routines can be employed for calculating the vibrational frequencies and normal modes, as well as IR, Raman, and ROA intensities. When calculating the ROA intensity differences, the $\beta(\mathbf{G}')^2$ invariant is calculated in the velocity representation to ensure gauge invariance, whereas the $\beta(\mathbf{A})^2$ invariant, which is always gauge invariant, is calculated in the length representation.[78] All Raman scattering factors are calculated for linearly polarized incident light and for the scattered light detected at 90°, ROA intensity differences are for 180° backscattering. Both the Raman and the ROA spectra use an excitation wavelength of 799 nm. In all plotted spectra, the calculated transitions have been broadened using a Lorentzian line shape with a full-width at half-maximum of 15 cm^{-1}. If included in the plots, the line spectrum has been scaled by 0.05 compared to the broadened spectrum.

4. VERIFYING THE CTTM IMPLEMENTATION

First, we carefully tested that our implementation of the CTTM is correct. As a simple test case to verify that the various transformation steps have been implemented correctly, we study the L-alanine molecule. The IR, Raman, and ROA spectra calculated for L-alanine are shown in Figure 3 as "Original" at the bottom of each plot. The molecule is then rotated by 120° in the xy plane and afterward translated by 1.0 bohr in the x and y directions. The spectra calculated for this rotated and translated molecule are shown in Figure 3 as "Rotated & Translated" at the top of each plot. The Hessian and the property tensor derivatives calculated for the rotated and translated L-alanine molecule are then transformed back to the coordinates of the original molecule. The resulting spectra are shown in Figure 3 in the middle of each plot as "Transferred". The IR, Raman, and ROA spectra are identical in all three cases, which demonstrates that the implementation is working correctly. For a more detailed comparison, the IR intensities as well as the Raman and ROA invariants for all normal modes are listed in the Supporting Information. From the calculation for the translated and rotated molecule and from the CTTM identical values are obtained. For the original L-alanine molecule, very small deviations are observed. These are due to the numerical integration grid in the DFT calculations, which depends on the orientation of the molecule.

However, since the origin dependence of the $\mathbf{G}'$ tensor and the $\mathbf{A}$ tensor drops out when the ROA invariants are calculated, one still has to verify that the translation of the molecule (eqs 7−8, 11, and 12) is performed correctly. Therefore, in the Supporting Information, also the elements of derivatives of the $\mathbf{G}'$ tensor and the $\mathbf{A}$ tensor with respect to some of the nuclei are given. Again, a close agreement between all three calculations is found. This validates our implementation of the CTTM.

As a first test of the CTTM itself, we consider a simple case where the CTTM gives correct results. To this end, we chose two L-alanine molecules which are located far apart from each other, separated by approximately 10.5 Å. The IR, Raman, and ROA spectra calculated for these two L-alanine molecules in a full calculation are shown in Figure 4 as "Original" at the bottom of

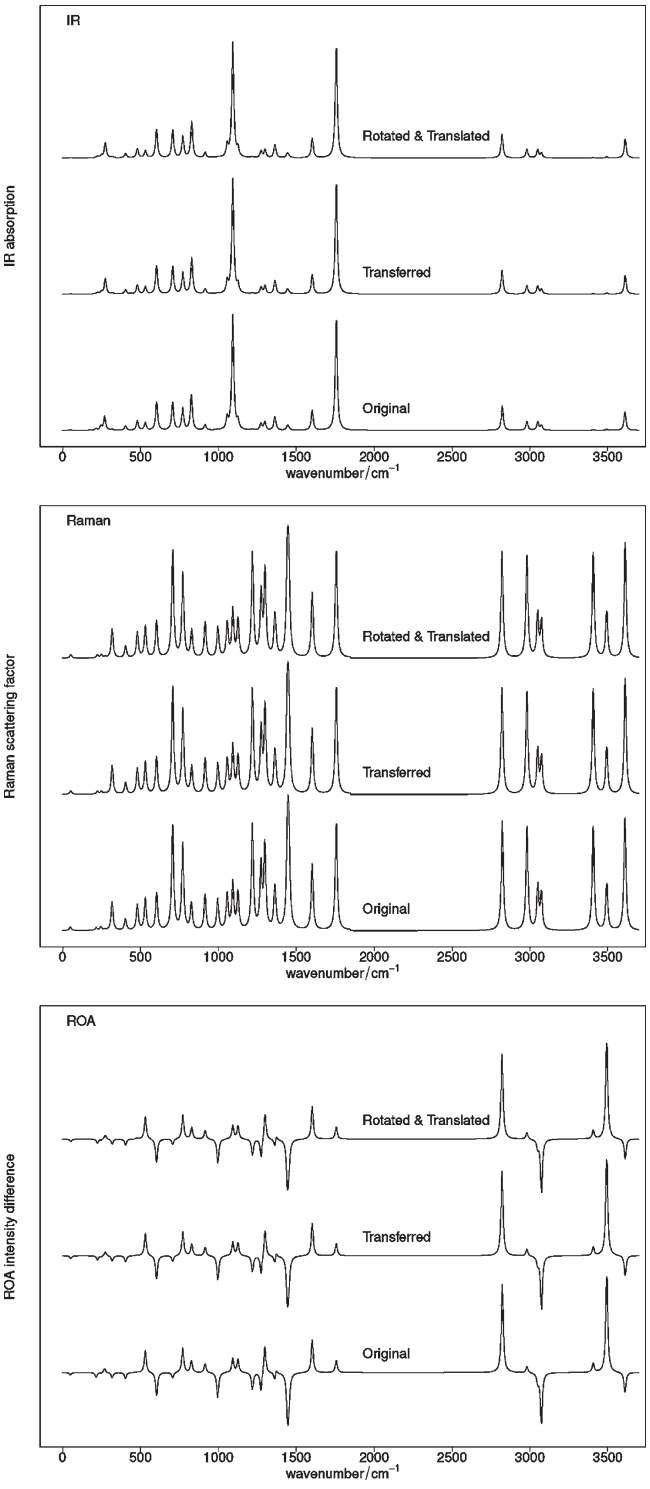

Figure 3. Test case for the implementation of the Cartesian Tensor Transfer Method: Reconstruction of the spectrum of L-alanine from an L-alanine molecule translated and rotated in space.

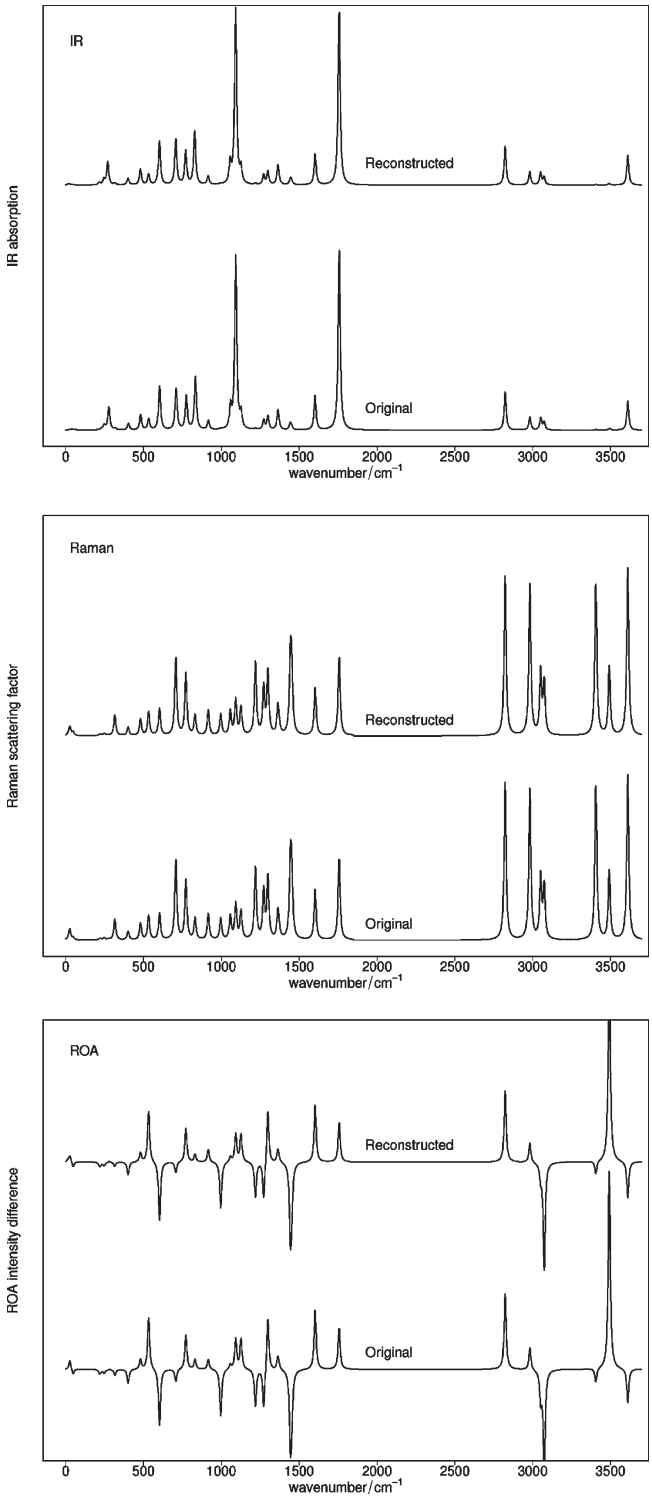

Figure 4. Test case for our implementation of the Cartesian Tensor Transfer Method: Reconstruction of the spectrum of two distant L-alanine molecules from two separately calculated L-alanine molecules.

each plot. The "Reconstructed" spectra are obtained by merging two separately calculated L-alanine molecules with exactly the same structure as in the full calculation with the CTTM. The IR, Raman, and ROA spectra show no differences between the result of the full calculation and of the CTTM, which also demonstrates that our implementation is correct.

In an additional test, we consider a N-methyl-acetamide (NMA) trimer. A similar test case was already used by Bouř et al.[44] The structure of the trimer is fully optimized, and as a reference, the full IR, Raman, and ROA spectra are calculated. From the trimer structure, we then constructed two NMA dimers, one by taking the two N-terminal NMA units and

dx.doi.org/10.1021/ct2001478 |J. Chem. Theory Comput. 2011, 7, 1867–1881

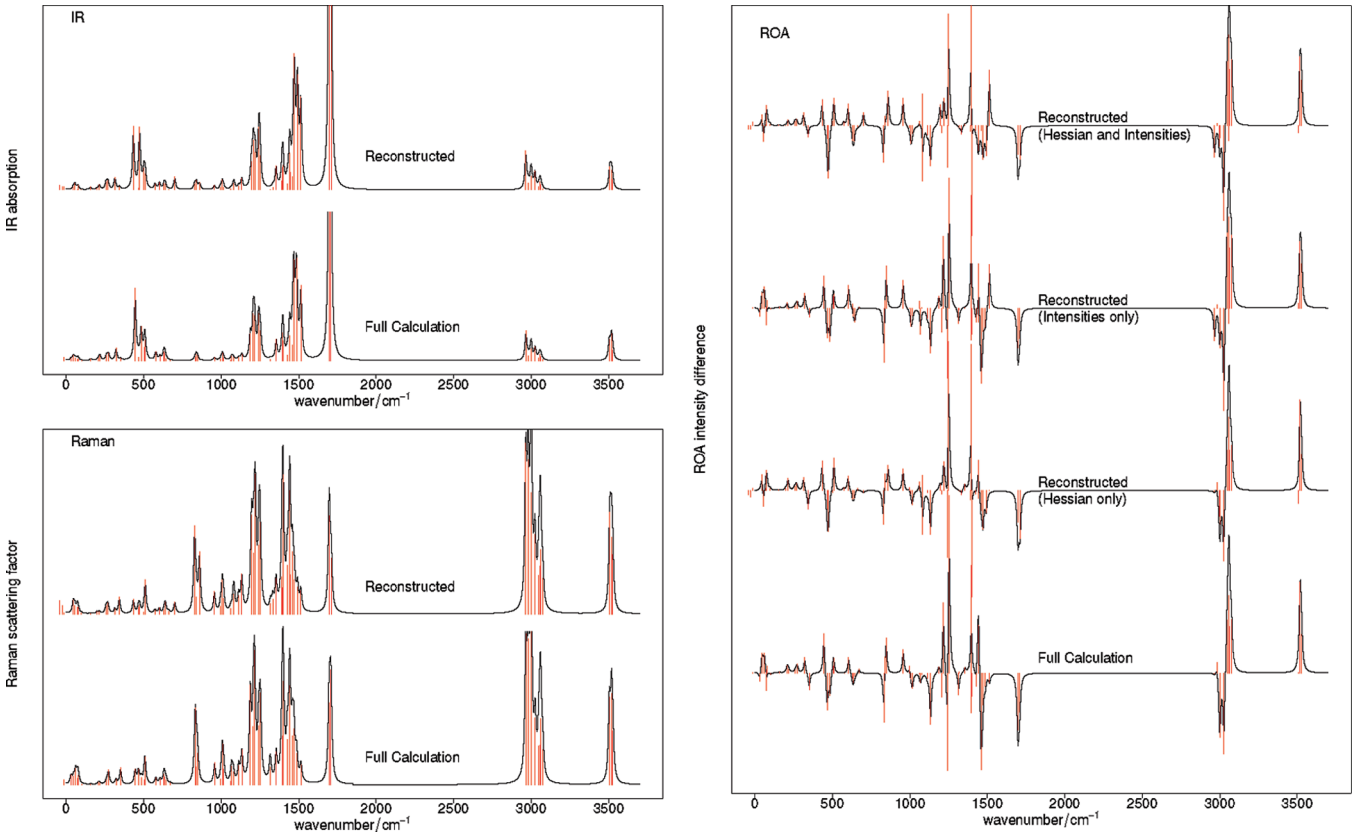

Figure 5. Vibrational spectra of an N-methyl-acetamide trimer reconstructed from two dimer calculations with the CTTM (IR, top left; Raman, bottom left; ROA, right). In the Raman spectra, the region below 2000 cm $^{-1}$ has been magnified by a factor of 8.

another one by considering the two C-terminal NMA units. For these dimers, the atomic coordinates are fixed to those in the trimer, and only the positions of the additional terminal hydrogen atoms are optimized. The IR, Raman, and ROA spectra of the trimer are then reconstructed by transferring the Hessians and property tensor derivatives from the two dimers to the trimer. Note that since the atomic coordinates of the dimers match those of the trimer, a rotation or translation is actually not required in this case, and the Hessian matrix elements and property tensor derivatives calculated for the dimers could be used directly.

The comparison of the IR and Raman spectra from the full calculation and from the CTTM are compared in the plots in the left column of Figure 5. Even though there are some slight differences (e.g., the peak at about 850 cm $^{-1}$ in the Raman spectrum from the full calculation is split into two peaks in the reconstructed Raman spectrum), the agreement is very good. Also for the ROA spectra, shown in the right column of Figure 5, the full calculation (bottom spectrum) and the reconstructed spectrum from the CTTM (top spectrum) match closely in most regions. However, there are also some clear deviations: For instance, around 1500 cm $^{-1}$ (where combinations of amide II and side chain CH_3 bending vibrations appear), the spectrum from the full calculation shows a couplet, while in the reconstructed spectrum, there are only negative peaks. Similarly, there are deviations in the amide III region between 1200 and 1300 cm $^{-1}$.

To shed light on the origin of these deviations, Figure 5 also includes the spectra obtained if only the Hessian matrix is transferred and the property tensor derivatives from the full calculation are used (labeled "Hessian only") and if only the property tensor derivatives are transferred but the Hessian from the full calculation is used (labeled "Intensities only"). In both cases, there are still differences in the full calculation, but if the Hessian from the full calculation is used, the spectrum agrees more closely with the reference. Therefore, it appears that the ROA spectrum is, in contrast to the IR and Raman spectra, rather sensitive to (small) changes in the normal modes.

5. ANALYSIS OF THE CTTM FOR AN $(ALA)_{20}$ POLYPEPTIDE

Since the CTTM has extensively been applied to calculate vibrational spectra of biomolecules, in particular polypeptides, it is crucial to assess its accuracy for such systems. As a typical test case, we choose a polypeptide of 20 alanine residues $(Ala)_{20}$ in an α-helical conformation. Full DFT calculations of the IR and Raman spectra as well as the ROA spectrum of this model system have previously been performed in our group and were analyzed in detail.[29,31,32] The structure of this $(Ala)_{20}$ α-helix is shown in Figure 6a. Even though these calculations might only partly agree with experimental spectra because of limitations of the computational methodology (i.e., the approximations applied for the exchange-correlation functional as well as the neglect of anharmonic effects) and the neglect of solvent effects, these full calculations can serve as a reference for the CTTM. Since the calculations for the small fragments are based on the very same approximations, any deviations of the reconstructed vibrational spectra from the full calculations are caused solely by the CTTM.

Of course, the accuracy of the CTTM will strongly depend on the size of the small fragments employed. The structure of the

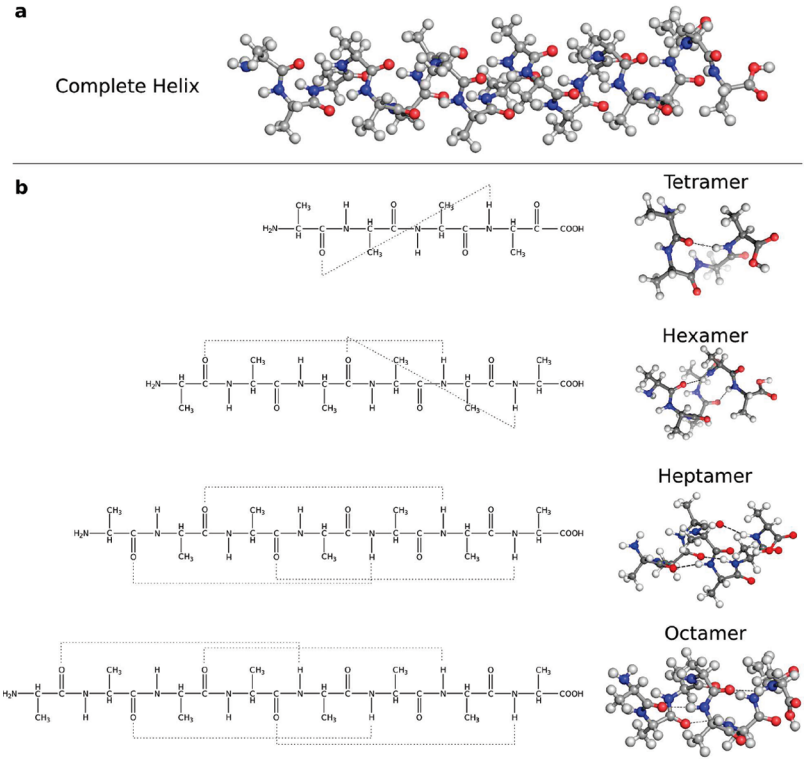

Figure 6. (a) Molecular structure of the $(Ala)_{20}$ α-helix. (b) Molecular structures of the different fragments used to reconstruct the vibrational spectra with the CTTM. In the octamer, there is one peptide bond, which has one H bond in both directions.

α-helix is stabilized by hydrogen bonds, which could strongly affect the vibrational frequencies as well as the property tensor derivatives. In too small fragments, these hydrogen bonds are not present. Furthermore, the size of the fragments determines the number of off-diagonal elements of the Hessian that are included in the CTTM.

To test the sensitivity of the CTTM to the size of the small fragments, four different sizes are compared. As the smallest fragment, an alanine tetramer $(Ala)_4$ is used. This fragment is cut from the C-terminal end of the α-helix and contains only one internal hydrogen bond. Next, a hexamer $(Ala)_6$, also taken from the C-terminal end of the α-helix, has been employed. In this hexamer, there are two hydrogen bonds. Additionally, a heptamer $(Ala)_7$ and an octamer $(Ala)_8$ have been considered. The structures of these larger fragments have been cut from the central part of the $(Ala)_{20}$ (starting from the eighth and from the second amino acid counting from the C-terminus for the heptamer and the octamer, respectively) and thus have an mostly undistorted α-helical structure. The octamer is the smallest possible fragment in which the central peptide group is involved in hydrogen bonds both at the N−H and at the C═O group. The structures of these four different small fragments are shown in Figure 6b. For all of these small fragments, all molecular coordinates, except for the ones of the terminal NH_2 group, have been fixed in order to keep their structures as close to the full helix as possible. Therefore, the small fragments are no minimum structures with respect to the energy anymore, and the Hessians calculated for these fragments have negative eigenvalues.

The IR and Raman spectra obtained for the $(Ala)_{20}$ α-helix with the CTTM are compared to the full calculation in Figure 7. In general, a good agreement between the full calculation and the CTTM is found already for the spectra reconstructed from the alanine tetramer. When going to a larger fragment size, the spectra only change slightly. There are only a few exceptions where the agreement is worse. First, in the region between 3300 and 3600 cm^{-1}, where the N−H stretching (amide A) vibrations appear, the CTTM cannot reproduce the full calculations. Since the hydrogen bonds in the α-helix are formed between the N−H and the C═O groups of the backbone, these N−H stretching vibrations are especially sensitive to the (partial) neglect of these hydrogen bonds in the CTTM. Only for the octamer, the shape of the amide A band resembles the full calculation, but its intensity is still underestimated, both in the IR and in the Raman spectrum. The peak at 3309 cm^{-1} in the full calculation (slightly below the amide A band), which stems from the N−H stretch vibration of the terminal NH_2 group, is missing in all CTTM spectra. Since we took the structure of our fragments from the central parts or from the C-terminal end of the helix, the structure and chemical environment of the terminal NH_2 are not well described.

Second, the IR and Raman spectra from the CTTM do not agree with the full calculation in the extended amide III region between 1100 and 1350 cm^{-1} (i.e., for the amide III and the C_α−H bending vibrations). While the positions of most of the bands in this region agree in all cases, their relative intensities change significantly between the different fragment sizes. It is known that the extended amide III region is particularly sensitive to structural changes because of the coupling between the classical amide III and the C_α−H bending vibrations.[98,99,34] Finally, there are also deviations in the lower wavenumber region, for instance, around 500 cm^{-1} in the IR spectra or between 650 and 700 cm^{-1} in the Raman spectra. For these features, one can observe that the agreement of the CTTM with the full calculation improves if larger fragments are used.

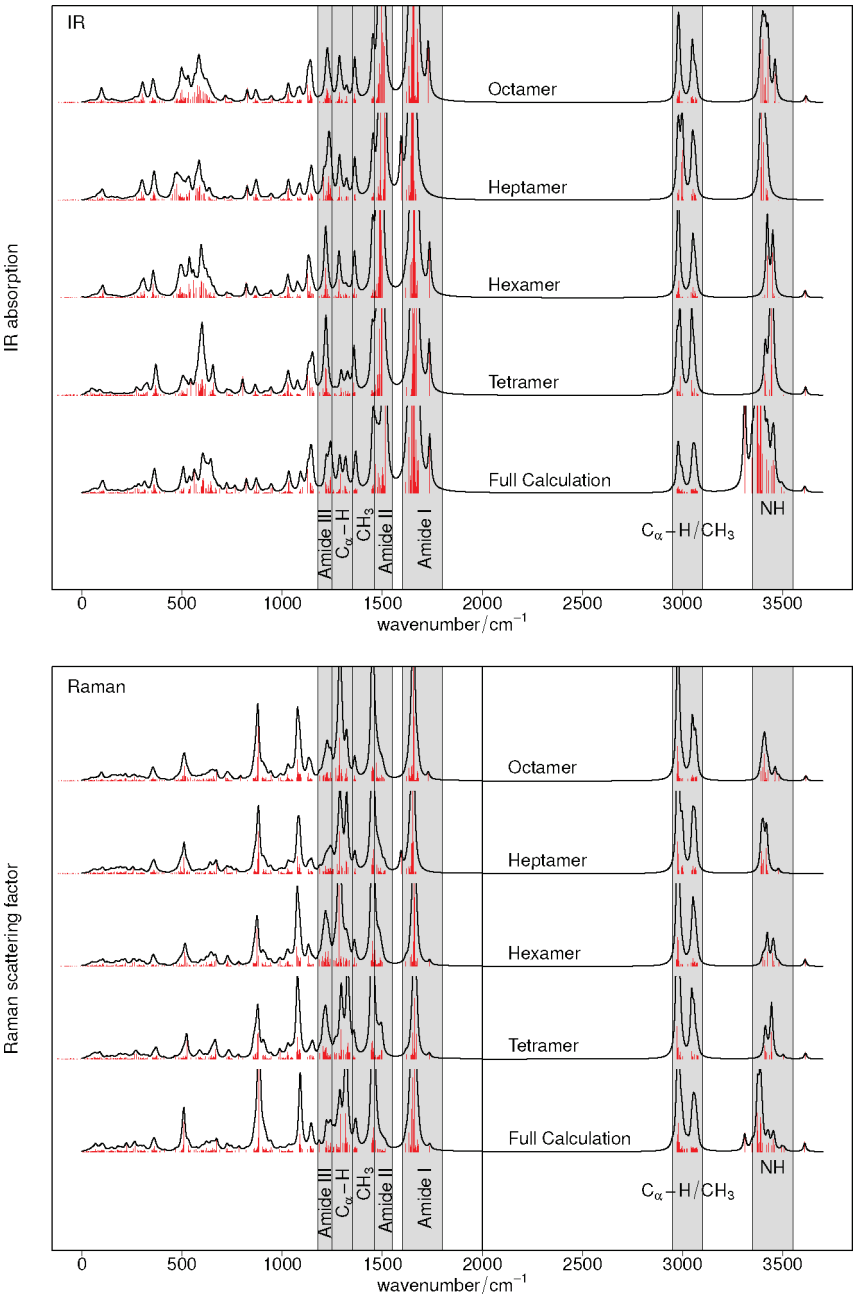

Figure 7. Comparison of the IR and Raman spectra from the full calculation on an $(Ala)_{20}$ α-helix and the spectra obtained with the CTTM by using an alanine tetramer, hexamer, heptamer, and octamer as small fragments. In the Raman spectra, the region up to 2000 cm^{-1} is enlarged by a factor of 20.

A more quantitative comparison can be obtained by looking at the differences in the vibrational frequencies of each individual transition. For such a comparison, we identify the normal modes corresponding to the same transition in the full and CTTM calculations by identifying those pairs of normal modes that have the largest overlap (as defined, for instance, in ref 30). For the CTTM calculations starting from the different fragment sizes, Figure 8 highlights the error in the wavenumber for each of the normal modes. The comparison shows quite significant errors in the vibrational frequencies. For each of the bands, the errors in the individual vibrational frequencies are not systematic but scatter (usually around zero) by about 20 to 50 cm^{-1}. When ignoring the low-frequency vibrations (below about 300 cm^{-1}), the largest errors of up to 80 cm^{-1} are found for the N−H stretching (amide A) vibrations. Also, for the extended amide III region, larger deviations are observed. Nevertheless, it appears that for most of the other regions the errors in the vibrational frequencies—even though they are considerable—do not affect the overall IR and Raman spectra significantly.

For the ROA spectrum, the comparison of the CTTM with the full calculation is shown in the upper part of Figure 9. Some features found in the ROA spectrum from the full calculation can (at least qualitatively) also be found in all the spectra, e.g., the negative peak for the skeletal stretch vibrations at about 1150 cm^{-1}, the strong positive peak at the lower-wavenumer end of the C_α−H stretching region, or the small positive peak for the symmetric CH_3 bending vibrations. Other features, such as the negative peak at the lower-wavenumber end of the amide III

dx.doi.org/10.1021/ct2001478 |J. Chem. Theory Comput. 2011, 7, 1867–1881

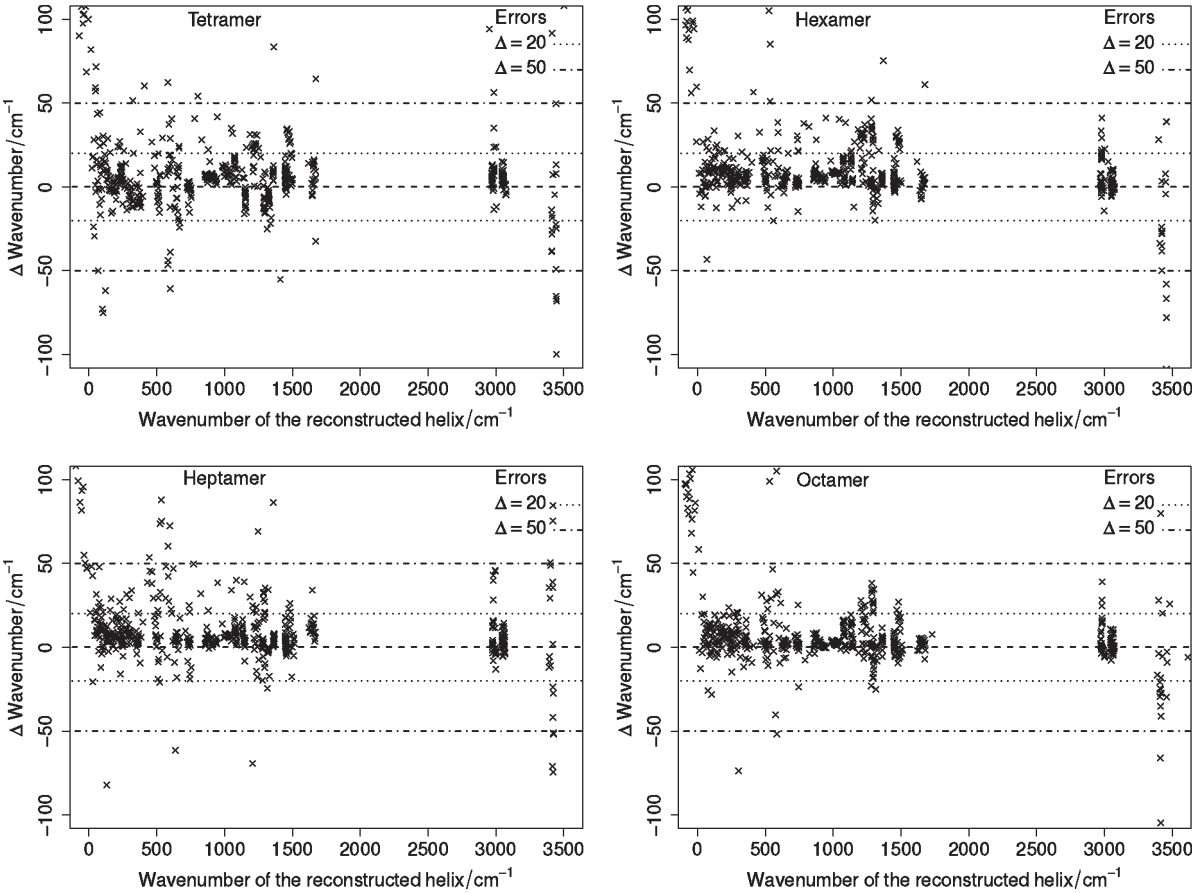

Figure 8. Comparison of the vibrational frequencies obtained with the CTTM starting from a tetramer, hexamer, heptamer, and octamer to those obtained in a full calculation. For each normal mode in the reconstructed spectrum, the corresponding normal mode in the full calculation is identified by considering the overlap between the modes.

region or the couplet for the asymmetric CH_3 stretching vibrations at 1450 cm^{-1} appear with the CTTM only for the heptamer and the octamer. In general, one notices that there is a much stronger dependence on the size of the fragments used in the CTTM, and in some regions, the spectrum changes significantly with fragment size. An example is the amide I region: While there is a negative−positive couplet in the full calculation,[31] a positive peak is found with the CTTM using the tetramer; a couplet is found for the hexamer, and negative peaks are found for the octamer. It is important to point out that, while in some regions the CTTM is able to approach the full calculation if large enough fragments are used, this is not true in general. Even with the octamer, the CTTM produces—in contrast to the full calculation—a strong negative amide II band, an additional negative peak in the amide III region, and an additional positive peak at about 1000 cm^{-1}. Moreover, there are significant deviations from the full calculation in the extended amide III (i.e., amide III and C_α−H bending) region.

One reason for the dissatisfying performance of the CTTM for ROA spectra could be the influence of the terminal residues in the small fragments. These have a different electronic structure than the central residues, which will affect both the Hessian and the property tensor derivatives but are still considered in the reconstruction of the Hessian and property tensor derivatives of the central parts of the full helix. To test their influence on the reconstructed ROA spectra, we repeated the CTTM calculations, but this time the terminal residues are excluded when mapping the atoms of the small fragment to the full α-helix (except for the "ends" of the helix). For the tetramer, the two central peptide groups; for the hexamer, the three central peptide units; and for the heptamer and octamer, the four and three central peptide groups, respectively, are included. The spectra obtained with these smaller overlaps in the CTTM are shown in the lower part of Figure 9. However, while the ROA spectra change in some parts, the disagreement between the CTTM and the full calculations remains.

To better understand why the CTTM is in disagreement with the full calculation of the ROA spectra, we also performed calculations where the CTTM is applied only to the property tensor derivatives, but where the Hessian from the full calculation is used ("Intensities only"), as well as calculations where the CTTM is applied only for the Hessian, but where the property tensor derivatives from the full calculation are used ("Hessian only"). The results are shown in Figure 10. For the "Intensities only" spectra (shown in the upper part), there are still large deviations from the full calculation, and in some parts, the ROA spectrum changes completely when the fragment size is increased. On the other hand, for the "Hessian only" CTTM calculations, the ROA spectra agree rather well with the full calculation. Nevertheless, there are still some smaller differences, for instance, in the extended amide III region or for the amide II band. Therefore, the dissatisfying performance of the CTTM for ROA calculations is mainly caused by the transfer of the property tensors, not by approximations introduced for the Hessian.

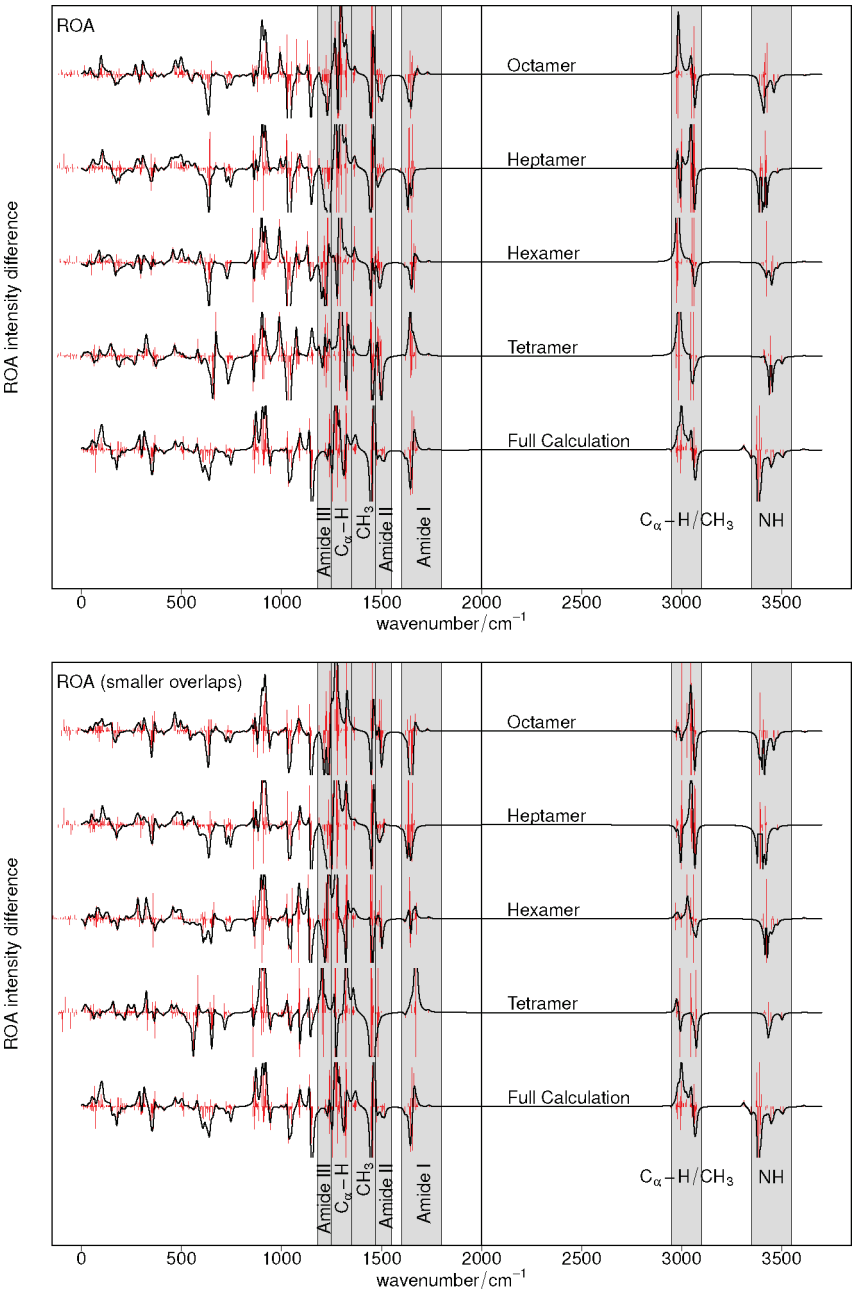

Figure 9. Comparison of the ROA spectra from the full calculation on an $(Ala)_{20}$ α-helix and the spectra obtained with the CTTM by using an alanine tetramer, hexamer, heptamer, and octamer as small fragments. In the top part, all atoms of the small fragments are included for the CTTM calculation, while in the bottom part, the terminal residues of the small fragments are not considered. In all spectra, the region up to 2000 cm^{-1} is enlarged by a factor of 20.

As the ROA intensities depend on the derivatives of three different property tensors (the electric-dipole−electric-dipole polarizability tensor α, the electric-dipole−magnetic-dipole tensor $\mathbf{G}'$, and the electric-dipole−magnetic-quadrupole polarizability tensor $\mathbf{A}$),[9] it is instructive to investigate the degree to which the individual tensors are affected by the CTTM. The α tensor cannot be the reason for the disagreement with the full calculation because good results were obtained for the Raman intensities. To determine which one of the other two tensors is responsible for the errors, pseudospectra treating the two ROA invariants $\beta(\mathbf{G}')^2$ and $\beta(\mathbf{A})^2$ separately are plotted in Figure 11. In these pseudospectra, only the property tensor derivatives have been reconstructed with the CTTM, whereas the Hessian from the full calculation has been used. From the plots, it is evident that the derivatives of the $\mathbf{A}$ tensor are less affected by the CTTM than the $\mathbf{G}'$ tensor derivatives. In addition, the invariants $\beta(\mathbf{G}')^2$ and $\beta(\mathbf{A})^2$ enter the ROA intensity expression in a 3:1 ratio (the pseudospectra in Figure 11 have been scaled accordingly). Keeping in mind that the $\mathbf{A}$ tensor hardly contributes to the full ROA spectrum,[100] it is obvious that the error in the CTTM-constructed ROA spectrum stems from the $\mathbf{G}'$ tensor derivatives when compared to the full ROA calculation.

To further understand the origin of the errors in the ROA intensities introduced by the CTTM, an analysis in terms of local

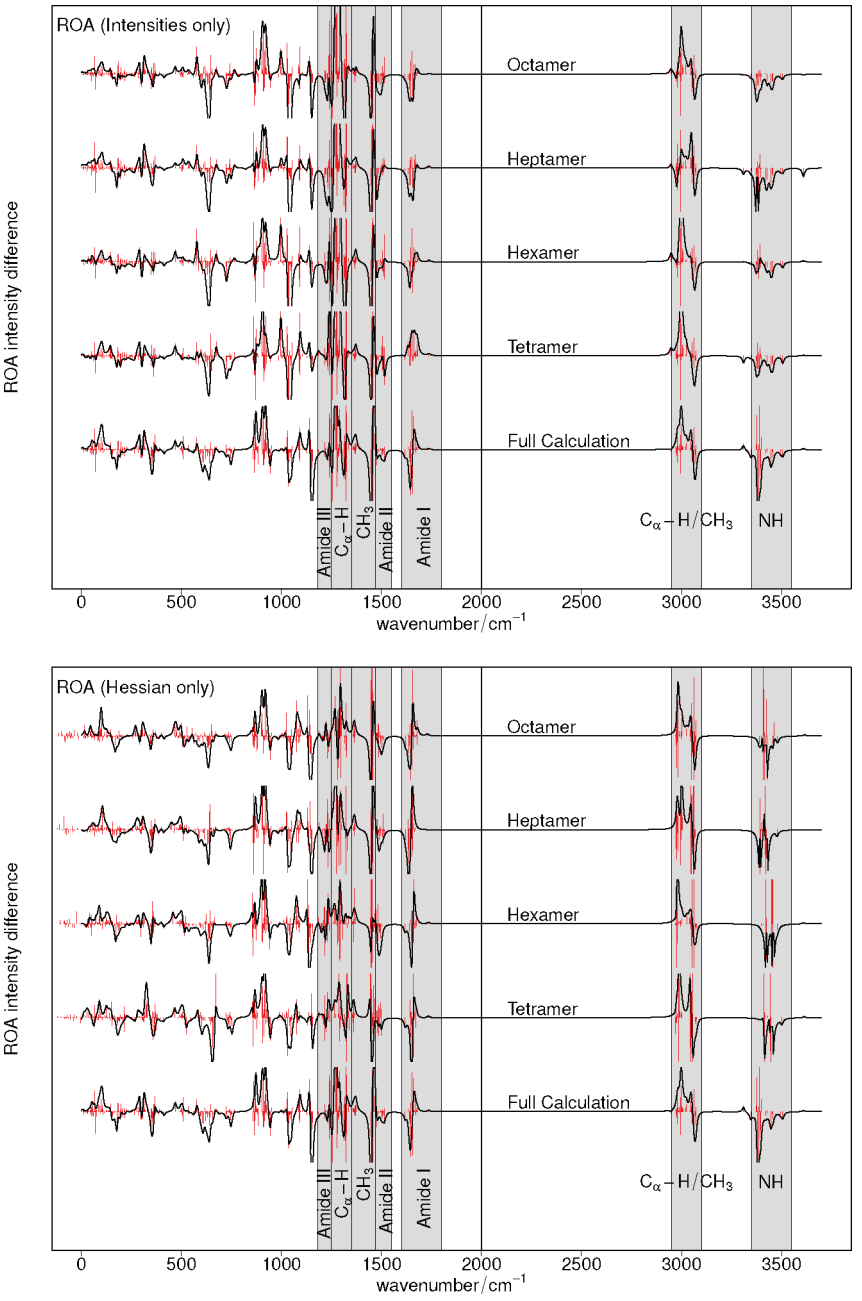

Figure 10. Comparison of the ROA spectra from the full calculation on an $(Ala)_{20}$ α-helix and the spectra obtained with the CTTM applied only for the intensities (top part) or only for the Hessian (bottom part). In all spectra, the region up to 2000 cm^{-1} is enlarged by a factor of 20.

modes[29] can provide additional insight. In such an analysis, the delocalized normal modes contributing to one band are unitarily transformed to a set of localized modes. In general, these localized modes are vibrations of a single amino acid residue or peptide group. Even though they do not correspond to the transitions observed in the experiment, a wavenumber and a ROA intensity can be assigned to each localized mode, and the total intensity of a specific band is invariant under the transformation from normal modes to localized modes. For more details, we refer to refs 29 and 32. For the ROA spectrum of $(Ala)_{20}$, such an analysis has previously been performed in our group for the full ROA calculation,[31] to which we may compare here.

As an example, we consider the amide III band between ca. 1200 and 1260 cm^{-1}. For this band, the full ROA calculation yields a rather small negative peak, while a much stronger negative peak is obtained in the CTTM calculation, even when using the large octamer fragment. The ROA intensities of the amide III localized modes are listed in Table 1 for both the full calculation and the intensities-only CTTM calculation employing the octamer fragment. Note that because the CTTM is applied for the property tensor derivatives only, the normal modes and also the localized modes are identical in both cases. For the terminal residues, the localized-mode intensities from the CTTM qualitatively agree with the full calculation. However, for the central residues, there are larger differences. While in the full calculations, the ROA intensities of the localized modes are approximately -10×10^{-3} $Å^4$/amu, those obtained when applying the CTTM are about twice as large. This is then reflected by the sum of the intensities for the

dx.doi.org/10.1021/ct2001478 |J. Chem. Theory Comput. 2011, 7, 1867–1881

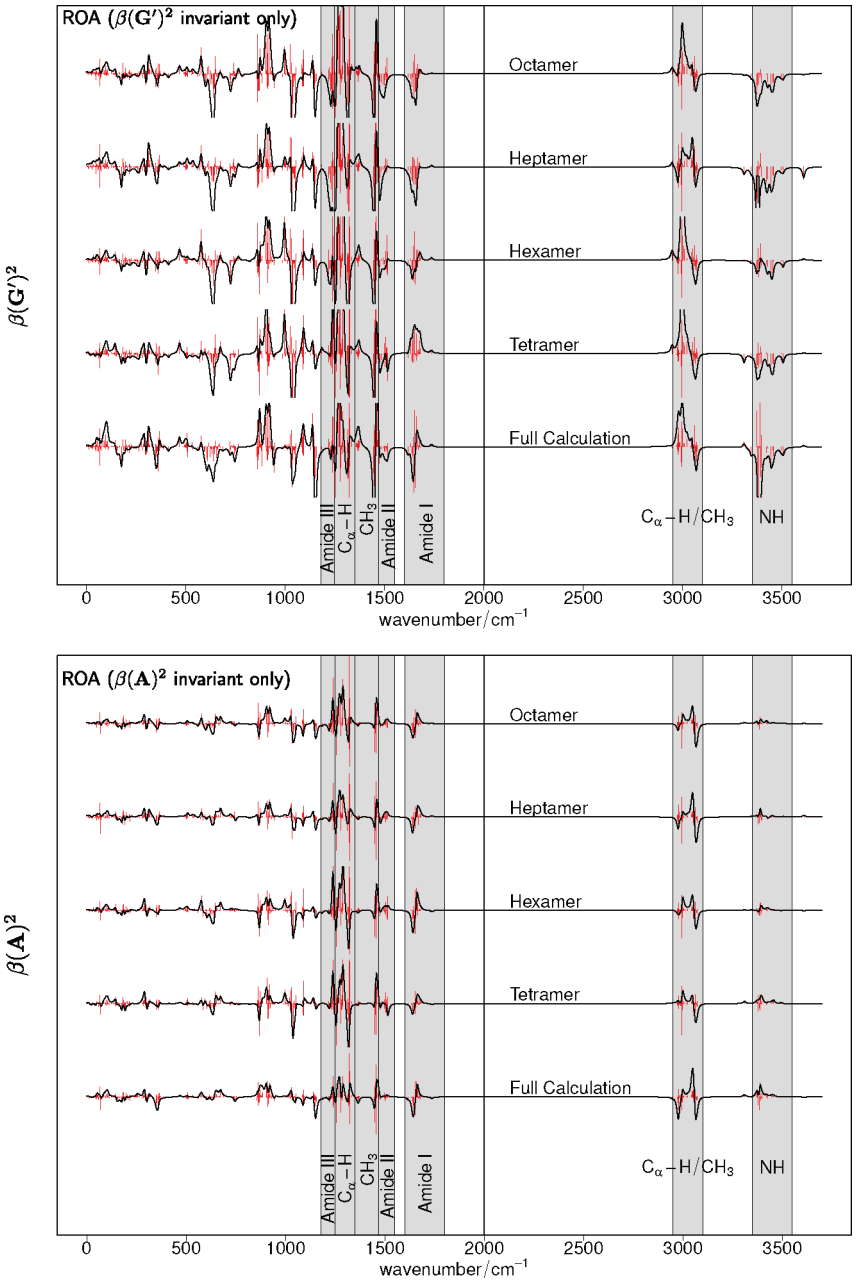

Figure 11. Comparison of the pseudospectra showing only the $\beta(\mathbf{G}')^2$ and $\beta(\mathbf{A})^2$ invariants from the full calculation on an $(Ala)_{20}$ α-helix and from the CTTM. The pseudospectra are scaled such that their sum corresponds to the total ROA intensity, as shown in the upper part of Figure 10. In all spectra, the region up to 2000 cm^{-1} is enlarged by a factor of 20.

amide III band, resulting in a much stronger negative peak in the CTTM calculation. In addition, the ROA intensities of the octamer fragment can also be analyzed in terms of local modes. The resulting wavenumbers and ROA intensities of the localized modes are included in Table 1. A comparison shows that the localized mode intensities of the $(Ala)_{20}$ CTTM calculation are similar to the ones of the octamer fragment calculation. This is not surprising, since the property tensors of the fragment were used to reconstruct the corresponding property tensor derivatives of the complete helix. In conclusion, one would obtain correct ROA intensities if the localized mode intensities of the fragment and the large molecule were similar. However, our analysis shows that, even with the relatively large octamer fragment, this is not the case, in particular for the central residues. Finally, we note that for the other bands in the ROA spectrum, similar observations could be made.

6. DISCUSSION AND CONCLUSION

By comparing the vibrational spectra reconstructed from small fragments to those from a full calculation, the accuracy of the CTTM has been investigated. For IR and Raman spectra, we find that the CTTM yields spectra which are generally in good agreement with the full calculation. However, we also find that for the vibrational frequencies of the individual normal modes, there are rather large deviations of 20 to 50 cm^{-1}, and in some

Table 1. Wavenumbers Ω_{ii} (in cm $^{-1}$) and ROA Intensities (in 10^{-3} Å^4/amu) for the Amide III Localized Modes of the Full Calculation and the Intensity-Only CTTM Calculation on the $(Ala)_{20}$ Helix and for the Amide III Localized Modes Obtained for the Octamer Fragment Used in the CTTM

	$(Ala)_{20}$ helix			octamer	
residue	Ω_{ii}	ROA int. (full)	ROA int. (CTTM)	Ω_{ii}	ROA int. (full)
1	1225.7	1.87	9.84	1215.0	14.32
2	1198.6	−26.20	−26.58	1217.2	−23.43
3	1219.6	−18.22	−20.24	1218.9	−32.55
4	1250.0	−8.07	−19.30	1228.4	−23.94
5	1235.7	−5.41	−15.42	1239.5	−30.21
6	1228.5	10.71	3.16	1222.0	15.08
7	1237.0	−9.64	−9.27	1237.5	−27.05
8	1242.6	−10.42	−20.61	1225.0	44.62
9	1242.0	−6.85	−22.02		
10	1243.2	−12.93	−18.01		
11	1243.4	−12.62	−20.89		
12	1244.1	−13.95	−19.73		
13	1243.8	−18.18	−20.07		
14	1238.7	−19.76	−41.18		
15	1246.0	−20.44	−43.80		
16	1240.3	−20.43	−32.27		
17	1224.4	16.59	20.09		
18	1239.6	−14.72	−17.43		
19	1223.2	58.46	51.84		
20	1254.3	−14.38	−5.73		
20	1221.7	33.75	32.51		
sum		−110.86	−235.10		−63.16

problematic cases (such as the N−H stretching vibrations), even larger errors are found. The observation that these errors hardly affect the overall IR and Raman spectra can be understood through an analysis in terms of local modes: As was discussed in ref 32, the positions and total intensities of the bands observed in the IR and Raman spectra of polypeptides are determined by local properties (i.e., vibrational frequencies and intensities of localized modes). These local properties can be described adequately by small fragments. On the other hand, the coupling between local modes determines the observed band shapes. Some of these couplings (i.e, off-diagonal elements of the Hessian) are neglected in the CTTM, which affects the individual normal modes but leaves the overall band shapes unchanged, particularly if the nearest-neighbor couplings are included. However, for bands where the local mode properties are very dependent on the environment, such as the N−H stretching vibrations, or in regions that are very sensitive to changes in the coupling between different modes, such as the extended amide III region,[34] the approximations introduced by the CTTM show up as errors in the IR and Raman spectra.

For ROA spectroscopy, a different picture emerges. The CTTM spectra are strongly dependent on the size of the fragments. Even with the largest fragment used here, an octamer, the agreement of the CTTM ROA spectra with the full calculation is not satisfying. However, if the CTTM is applied only to the Hessian and not to the property tensor derivatives, a much better agreement with the full calculation is found. Thus, the poor performance of the CTTM for ROA is due to the approximations introduced when calculating property tensor derivatives for smaller fragments. On the other hand, applying the CTTM to the Hessian and neglecting the vibrational coupling between some residues only has a minor effect. Revisiting the analysis of the ROA spectrum of $(Ala)_{20}$ in terms of local modes[31] helps to understand these errors. The comparison of the localized mode intensities from the CTTM to those from the full calculation clearly shows the reason for these errors. The localized ROA intensities of the octamer fragment are quite different compared to those of the larger $(Ala)_{20}$ α-helix. Already in ref 31, it was found that the ROA intensities of the localized modes are very different for the terminal residues than for the central ones (see Tables in the Supporting Information of ref 31). Thus, because of the neglect of hydrogen bonding in the smaller fragments, the "local" ROA intensities can change significantly. Because this fragment is then used to reconstruct the complete $(Ala)_{20}$ α-helix, wrong ROA intensities are obtained, and since the ROA intensities can even change sign, the effect on the overall spectra is more dramatic than in the case of CTTM-constructed IR and Raman spectra.

Of course, there are several details of our implementation of the CTTM and of its application that could be changed in order to attempt to improve on the present results. Most important, other choices would be possible for the small fragments. Instead of extracting their geometry from the full calculation, the small fragments could be constructed to resemble an idealized α-helix. Moreover, instead of fixing their geometry, the small fragments could be (at least partly) optimized. In their applications of the CTTM method, Bouř, Keiderling and co-workers mostly employed a geometry optimization in normal coordinates for the small fragments, in which the low frequency modes are kept fixed.[101] Furthermore, our calculations used a single small fragment for reconstructing the whole α-helix. Instead, it would be possible to employ different small fragments for different parts of the helix. However, because the small fragments have to overlap, the additional computational cost for the calculations on the small fragments might then render the CTTM more expensive then a full calculation.

Another issue is the definition of the mapping between the atoms of the small fragment and the large molecule. For this, we tested two different options, either to include all atoms of the small fragment or to consider only those of the central residues (i.e., those where hydrogen bonding is accounted for and the electronic structure should be most similar to the one of the large molecule). Finally, if different small fragments (or different parts of the same fragment) can be employed in the reconstruction, there are different ways to decide which one is used. Here, we always choose the one that has the structure most similar to the large molecule. An average of all matching small fragments or a weighted average that prefers the central parts of the small fragment could be employed instead.

Carefully testing these different options might improve the agreement of the CTTM with the full calculations for ROA spectra. Nevertheless, this would also imply that the CTTM applied to ROA spectroscopy is very sensitive to the choice of these parameters. In conclusion, our analysis indicates that applying the CTTM for ROA spectra is strongly dependent on the details of how the CTTM is applied, most importantly on the size and structure of the small fragments. Therefore, many ROA results obtained previously with the CTTM will have to be re-evaluated carefully. It is particularly puzzling that the CTTM was

found to work well in direct comparison to the experiment (see references provided in the Introduction), which might point to a fortunate error compensation. If this is not the case, it might even be necessary to revise the conclusions previously drawn from such calculations. Moreover, it appears to be necessary to also investigate the reliability of the CTTM for the calculation of VCD spectra.

One solution to these problems could be to perform the calculations on the small fragments not for the isolated molecules but for embedded structures which feature an appropriate environment. A first step is to apply a continuum solvation model that could partly account for the effect of hydrogen bonding, as has already been done in applications of the CTTM (whether this is the reason for the high reliability previously reported remains to be shown). More accurate environment models, such as combined quantum mechanics/molecular mechanics (QM/MM) methods,[102,103] have already been adopted for use in computational vibrational spectroscopy, for instance, in the mobile block Hessian method.[42,43] Additionally, more advanced embedding approaches based on a subsystem formulation of DFT have been shown to provide an accurate description of the effect of hydrogen bonds.[104,105] Combining the generalization of such methods to polypeptides and proteins[106] and to the subsystem calculation of polarizabilities[107,108] could lead to accurate and efficient methods for the calculation of ROA spectra that go beyond the CTTM.

■ ASSOCIATED CONTENT

Supporting Information. Detailed results of the validation of our CTTM implementation for L-alanine, coordinates of all considered molecules as well as the corresponding fragments, and further details of the CTTM calculations for $(Ala)_{20}$. This information is available free of charge via the Internet at http://pubs.acs.org/.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: christoph.jacob@kit.edu, markus.reiher@phys.chem.ethz.ch.

■ ACKNOWLEDGMENT

Financial support by the Swiss National Science Foundation (project no. 200020-132542/1) for the Zurich group and from the DFG-Center for Functional Nanostructures at KIT for C.R.J. is gratefully acknowledged. We also would like to acknowledge correspondence with Petr Bouř on the implementation of the CTTM and on the accuracy of CTTM ROA spectra when we discovered the discrepancies reported in this paper in 2010.

■ REFERENCES

(1) Vass, E.; Hollosi, M.; Besson, F.; Buchet, R. *Chem. Rev.* **2003**, *103*, 1917–1954.
(2) Barth, A. *Biochim. Biophys. Acta, Bioenerg.* **2007**, *1767*, 1073–1101.
(3) Carey, P. R. *J. Biol. Chem.* **1999**, *274*, 26625–26628.
(4) Thomas, G. J., Jr. *Biopolymers* **2002**, *67*, 214–225.
(5) Holzwarth, G.; Hsu, E. C.; Mosher, H. S.; Faulkner, T. R.; Moscowitz, A. *J. Am. Chem. Soc.* **1974**, *96*, 251–252.
(6) Keiderling, T. A. In *Circular Dichroism: Principles and Applications*, 2nd ed.; Berova, N., Nakanishi, K., Woody, R. W., Eds.; Wiley-VCH: New York, 2000; pp 621−666.
(7) Silva, R. A. G. D.; Kubelka, J.; Bouř, P.; Decatur, S. M.; Keiderling, T. A. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 8318–8323.
(8) Kubelka, J.; Silva, R. A. G. D.; Keiderling, T. A. *J. Am. Chem. Soc.* **2002**, *124*, 5325–5332.
(9) Barron, L. D. *Molecular Light Scattering and Optical Activity*, 2nd ed.; Cambridge University Press: Cambridge, U. K., 2004; pp 327−381.
(10) Barron, L. D.; Bogaard, M. P.; Buckingham, A. D. *J. Am. Chem. Soc.* **1973**, *95*, 603–605.
(11) Hug, W.; Kint, S.; Bailey, G. F.; Scherer, J. R. *J. Am. Chem. Soc.* **1975**, *97*, 5589–5590.
(12) Barron, L. D.; Hecht, L.; McColl, I. H.; Blanch, E. W. *Mol. Phys.* **2004**, *102*, 731–744.
(13) Zhu, F.; Isaacs, N. W.; Hecht, L.; Tranter, G. E.; Barron, L. D. *Chirality* **2006**, *18*, 103–115.
(14) Barron, L. D.; Buckingham, A. D. *Chem. Phys. Lett.* **2010**, *492*, 199–213.
(15) Schweitzer-Stenner, R. *J. Raman Spectrosc.* **2001**, *32*, 711–732.
(16) Asher, S. A.; Ianoul, A.; Mix, G.; Boyden, M. N.; Karnoup, A.; Diem, M.; Schweitzer-Stenner, R. *J. Am. Chem. Soc.* **2001**, *123*, 11775–11781.
(17) Schweitzer-Stenner, R. *J. Raman Spectrosc.* **2005**, *36*, 276–278.
(18) Huang, C.-Y.; Balakrishnan, G.; Spiro, T. G. *J. Raman Spectrosc.* **2006**, *37*, 277–282.
(19) Hamm, P.; Lim, M.; Hochstrasser, R. M. *J. Phys. Chem. B* **1998**, *102*, 6123–6138.
(20) Ganim, Z.; Chung, H. S.; Smith, A. W.; DeFlores, L. P.; Jones, K. C.; Tokmakoff, A. *Acc. Chem. Res.* **2008**, *41*, 432–441.
(21) Zhuang, W.; Hayashi, T.; Mukamel, S. *Angew. Chem., Int. Ed.* **2009**, *48*, 3750–3781.
(22) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.* **2007**, *129*, 5887–5897.
(23) Lagutschenkov, A.; Langer, J.; Berden, G.; Oomens, J.; Dopfer, O. *J. Phys. Chem. A* **2010**, *114*, 13268–13276.
(24) Rossi, M.; Blum, V.; Kupser, P.; von Helden, G.; Bierau, F.; Pagel, K.; Meijer, G.; Scheffler, M. *J. Phys. Chem. Lett.* **2010**, *1*, 3465–3470.
(25) Rizzo, T. R.; Stearns, J. A.; Boyarkin, O. V. *Int. Rev. Phys. Chem.* **2009**, *28*, 481.
(26) Nagornova, N. S.; Rizzo, T. R.; Boyarkin, O. V. *J. Am. Chem. Soc.* **2010**, *132*, 4040–4041.
(27) Herrmann, C.; Reiher, M. *Top. Curr. Chem.* **2007**, *268*, 85–132.
(28) Hug, W. *Chem. Phys.* **2001**, *264*, 53–69.
(29) Jacob, Ch. R.; Reiher, M. *J. Chem. Phys.* **2009**, *130*, 084106.
(30) Hug, W.; Fedorovsky, M. *Theor. Chem. Acc.* **2008**, *119*, 113–131.
(31) Jacob, Ch. R.; Luber, S.; Reiher, M. *Chem.—Eur. J.* **2009**, *15*, 13491–13508.
(32) Jacob, Ch. R.; Luber, S.; Reiher, M. *J. Phys. Chem. B* **2009**, *113*, 6558–6573.
(33) Liégeois, V.; Jacob, Ch. R.; Champagne, B.; Reiher, M. *J. Phys. Chem. A* **2010**, *114*, 7198–7212.
(34) Weymuth, T.; Jacob, Ch. R.; Reiher, M. *J. Phys. Chem. B* **2010**, *114*, 10649–10660.
(35) Reiher, M.; Neugebauer, J. *J. Chem. Phys.* **2003**, *118*, 1634–1641.
(36) Reiher, M.; Neugebauer, J. *Phys. Chem. Chem. Phys.* **2004**, *6*, 4621–4629.
(37) Herrmann, C.; Neugebauer, J.; Reiher, M. *New J. Chem.* **2007**, *31*, 818–831.
(38) Kiewisch, K.; Neugebauer, J.; Reiher, M. *J. Chem. Phys.* **2008**, *129*, 204103.
(39) Kiewisch, K.; Luber, S.; Neugebauer, J.; Reiher, M. *Chimia* **2009**, *63*, 270–274.
(40) Luber, S.; Neugebauer, J.; Reiher, M. *J. Chem. Phys.* **2009**, *130*, 064105.
(41) Luber, S.; Reiher, M. *ChemPhysChem* **2009**, *10*, 2049–2057.
(42) Ghysels, A.; Van Neck, D.; Van Speybroeck, V.; Verstraelen, T.; Waroquier, M. *J. Chem. Phys.* **2007**, *126*, 224102.
(43) Ghysels, A.; Van Neck, D.; Brooks, B. R.; Van Speybroeck, V.; Waroquier, M. *J. Chem. Phys.* **2009**, *130*, 084107.

(44) Bouř, P.; Sopková, J.; Bednárová, L.; Maloň, P.; Keiderling, T. A. *J. Comput. Chem.* **1997**, *18* (5), 646–659.
(45) Bouř, P.; Zaruba, K.; Urbanova, M.; Setnicka, V.; Matejka, P.; Fiedler, Z.; Kral, V.; Volka, K. *Chirality* **2000**, *12*, 191–198.
(46) Kubelka, J.; Keiderling, T. *J. Am. Chem. Soc.* **2001**, *123*, 6142–6150.
(47) Mazaleyrat, J.; Wright, K.; Gaucher, A.; Wakselman, M.; Oancea, S.; Formaggio, F.; Toniolo, C.; Setnicka, V.; Kapitan, J.; Keiderling, T. *Tetrahedron: Asymmetry* **2003**, *14*, 1879–1893.
(48) Huang, R.; Setnicka, V.; Etienne, M. A.; Kim, J.; Kubelka, J.; Hammer, R. P.; Keiderling, T. A. *J. Am. Chem. Soc.* **2007**, *129*, 13592–13603.
(49) Grahnen, J. A.; Amunson, K. E.; Kubelka, J. *J. Phys. Chem. B* **2010**, *114*, 13011–13020.
(50) Bouř, P.; Kapitan, J.; Baumruk, V. *J. Phys. Chem. A* **2001**, *105*, 6362–6368.
(51) Bouř, P.; McCann, J.; Wieser, H. *J. Phys. Chem. A* **1997**, *101*, 9783–9790.
(52) Bouř, P.; McCann, J.; Wieser, H. *J. Phys. Chem. A* **1998**, *102*, 102–110.
(53) Bouř, P.; Kubelka, J.; Keiderling, T. *Biopolymers* **2000**, *53*, 380–395.
(54) Kubelka, J.; Keiderling, T. *J. Am. Chem. Soc.* **2001**, *123*, 12048–12058.
(55) Hilario, J.; Kubelka, J.; Syud, F.; Gellman, S.; Keiderling, T. *Biopolymers* **2002**, *67*, 233–236.
(56) Andrushchenko, V.; Wieser, H.; Bouř, P. *J. Phys. Chem. B* **2002**, *106*, 12623–12634.
(57) Hilario, J.; Kubelka, J.; Keiderling, T. *J. Am. Chem. Soc.* **2003**, *125*, 7562–7574.
(58) Huang, R.; Kubelka, J.; Barber-Armstrong, W.; Silva, R.; Decatur, S.; Keiderling, T. *J. Am. Chem. Soc.* **2004**, *126*, 2346–2354.
(59) Andrushchenko, V.; Wieser, H.; Bouř, P. *J. Phys. Chem. B* **2004**, *108*, 3899–3911.
(60) Bouř, P.; Keiderling, T. *J. Phys. Chem. B* **2005**, *109*, 5348–5357.
(61) Setnicka, V.; Huang, R.; Thomas, C.; Etienne, M.; Kubelka, J.; Hammer, R.; Keiderling, T. *J. Am. Chem. Soc.* **2005**, *127*, 4992–4993.
(62) Kubelka, J.; Huang, R.; Keiderling, T. *J. Phys. Chem. B* **2005**, *109*, 8231–8243.
(63) Bouř, P.; Andrushchenko, V.; Kabelac, M.; Maharaj, V.; Maharaj, V. *J. Phys. Chem. B* **2005**, *109*, 20579–20587.
(64) Bouř, P.; Keiderling, T. *J. Phys. Chem. B* **2005**, *109*, 23687–23697.
(65) Kim, J.; Huang, R.; Kubelka, J.; Bouř, P.; Keiderling, T. A. *J. Phys. Chem. B* **2006**, *110*, 23590–23602.
(66) Turner, D. R.; Kubelka, J. *J. Phys. Chem. B* **2007**, *111*, 1834–1845.
(67) Kapitan, J.; Baumruk, V.; Bouř, P. *J. Am. Chem. Soc.* **2006**, *128* (7), 2438–2443.
(68) Zhu, F.; Kapitan, J.; Tranter, G. E.; Pudney, P. D. A.; Isaacs, N. W.; Hecht, L.; Barron, L. D. *Proteins* **2008**, *70*, 823–833.
(69) Budesinsky, M.; Sebestik, J.; Bednarova, L.; Baumruk, V.; Safarik, M.; Bouř, P. *J. Org. Chem.* **2008**, *73*, 1481–1489.
(70) Kapitan, J.; Zhu, F.; Hecht, L.; Gardiner, J.; Seebach, D.; Barron, L. D. *Angew. Chem., Int. Ed.* **2008**, *47*, 6392–6394.
(71) Budesinsky, M.; Danecek, P.; Bednarova, L.; Kapitan, J.; Baumruk, V.; Bouř, P. *J. Phys. Chem. A* **2008**, *112*, 8633–8640.
(72) Sebek, J.; Kapitan, J.; Sebestik, J.; Baumruk, V.; Bouř, P. *J. Phys. Chem. A* **2009**, *113* (27), 7760–7768.
(73) Hudecova, J.; Kapitan, J.; Baumruk, V.; Hammer, R. P.; Keiderling, T. A.; Bouř, P. *J. Phys. Chem. A* **2010**, *114*, 7642–7651.
(74) Yamamoto, S.; Straka, M.; Watarai, H.; Bouř, P. *Phys. Chem. Chem. Phys.* **2010**, *12*, 11021–11032.
(75) Herrmann, C.; Ruud, K.; Reiher, M. *ChemPhysChem* **2006**, *7*, 2189–2196.
(76) Jacob, Ch. R.; Luber, S.; Reiher, M. *ChemPhysChem* **2008**, *9*, 2177–2180.
(77) Luber, S.; Reiher, M. *J. Phys. Chem. B* **2010**, *114*, 1057–1063.
(78) Luber, S.; Reiher, M. *Chem. Phys.* **2008**, *346*, 212–223.
(79) Quinet, O.; Champagne, B. *J. Chem. Phys.* **2001**, *115*, 6293–6299.
(80) Rappoport, D.; Furche, F. *J. Chem. Phys.* **2005**, *122*, 064105.
(81) Quinet, O.; Liegeois, V.; Champagne, B. *J. Chem. Theory Comput.* **2005**, *1*, 444–452.
(82) Liegeois, V.; Ruud, K.; Champagne, B. *J. Chem. Phys.* **2007**, *127*, 204105.
(83) Ruud, K.; Thorvaldsen, A. J. *Chirality* **2009**, *21*, E54–E67.
(84) Bast, R.; Ekström, U.; Gao, B.; Helgaker, T.; Ruud, K.; Thorvaldsen, A. J. *Phys. Chem. Chem. Phys.* **2011**, *13*, 2627.
(85) Altmann, S. L. *Rotations, Quaternions and Double Groups*; Oxford University Press: New York, 1986; pp 201−223.
(86) Kearsly, S. K. *J. Comput. Chem.* **1990**, *11* (10), 1187–1192.
(87) Kneller, G. R. *Mol. Simul.* **1991**, *7*, 113–119.
(88) Neugebauer, J.; Reiher, M.; Kind, C.; Hess, B. A. *J. Comput. Chem.* **2002**, *23*, 895–910.
(89) SNF 4.0.0 program for the quantum chemical calculation of vibrational spectra. Neugebauer, J.; Herrmann, C.; Luber, S.; Reiher, M. www.reiher.ethz.ch/software/snf (accessed Apr 04, 2011).
(90) Turbomole. Ahlrichs, R. http://www.turbomole.com (accessed Apr 04, 2011).
(91) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822–8824.
(92) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
(93) Schäfer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829–5835.
(94) Turbomole basis set library. ftp://ftp.chemie.uni-karlsruhe.de/pub/basen (accessed Apr 04, 2011).
(95) Turbomole auxiliary basis sets. ftp://ftp.chemie.uni-karlsruhe.de/pub/jbasen (accessed Apr 04, 2011).
(96) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *240*, 283–289.
(97) PyVib2, a program for analyzing vibrational motion and vibrational spectra. Fedorovsky, M. http://pyvib2.sourceforge.net (accessed Apr 04, 2011).
(98) Oboodi, M. R.; Alva, C.; Diem, M. *J. Phys. Chem.* **1984**, *88*, 501–505.
(99) Diem, M.; Lee, O.; Roberts, G. M. *J. Phys. Chem.* **1992**, *96*, 548–554.
(100) Luber, S.; Herrmann, C.; Reiher, M. *J. Phys. Chem. B* **2008**, *112*, 2218–2232.
(101) Bouř, P.; Keiderling, T. A. *J. Chem. Phys.* **2002**, *117*, 4126–4132.
(102) Senn, H.-M.; Thiel, W. *Top. Curr. Chem.* **2007**, *268*, 173–290.
(103) Senn, H. M.; Thiel, W. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.
(104) Jacob, Ch. R.; Neugebauer, J.; Jensen, L.; Visscher, L. *Phys. Chem. Chem. Phys.* **2006**, *8*, 2349–2359.
(105) Kiewisch, K.; Eickerling, G.; Reiher, M.; Neugebauer, J. *J. Chem. Phys.* **2008**, *128*, 044114.
(106) Jacob, Ch. R.; Visscher, L. *J. Chem. Phys.* **2008**, *128*, 155102.
(107) Neugebauer, J. *J. Chem. Phys.* **2009**, *131*, 084104.
(108) Neugebauer, J.; Curutchet, C.; Munoz-Losa, A.; Mennucci, B. *J. Chem. Theory Comput.* **2010**, *6*, 1843–1851.