Structure and Vibrational Motion of Insulin from Raman Optical Activity Spectra

3 **AUTHORS**, INCLUDING:

Jakub Kaminský

Ústav Organické Chemie a Biochemie AV ČR, v…

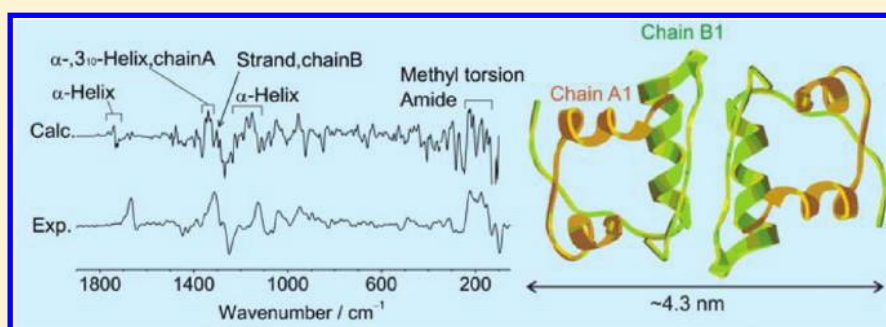**36** PUBLICATIONS **430** CITATIONS

SEE PROFILE

Structure and Vibrational Motion of Insulin from Raman Optical Activity Spectra

Shigeki Yamamoto,* Jakub Kaminský, and Petr Bouř*

Institute of Organic Chemistry and Biochemistry, Academy of Sciences, Flemingovo nám. 2, 166 10, Prague 6, Czech Republic

Ⓢ Supporting Information

ABSTRACT: The Raman optical activity (ROA) spectroscopic technique has been applied in the past to many biologically relevant systems including peptides, proteins, sugars, and even viruses. However, theoretical interpretation of the spectra relies on lengthy quantum-chemical computations, which are difficult to extend to larger molecules. In the present study, ROA and Raman spectra of insulin under a range of various conditions were measured and interpreted with the aid of the Cartesian-coordinate tensor transfer (CCT) method. The CCT methodology yielded spectra of insulin monomer and dimer of nearly ab initio quality, while at the same time reproducing the experimental data very well. The link between the spectra and the protein structure could thus be studied in detail. Spectral contributions from the peptide backbone and the amino acid side chains were calculated. Likewise, specific intensity features originating from the α-helical, coil, β-sheet, and 3_{10}-helical parts of the protein could be deciphered. The assignment of the Raman and ROA bands to intrinsic molecular coordinates as based on the harmonic force field calculation revealed their origin and degree of locality. Alternatively, the relation of the structural flexibility of insulin to the inhomogeneous broadening of spectral bands was studied by a combination of CCT and molecular dynamics (MD). The present study confirms the sensitivity of the ROA technique to some subtle static and dynamic changes in molecular geometry, and many previous ad hoc or semiempirical spectral-structure assignments could be verified. On the other hand, a limitation in longer-range tertiary structure sensitivity was revealed. Unlike for smaller molecules with approximately equal contributions of the electric dipole (α), quadrupole (A), and magnetic dipole (G′) polarizabilities, only the electric dipolar polarization (α) interactions seem to dominate in the protein ROA signal. The simulations concern the largest molecule for which such spectra were interpreted by *a priori* procedures and significantly enhance protein folding studies undertaken by this technique.

A direct structural characterization of a protein in its natural aqueous environment is important for elucidation of its biological function and determination of its stability under environmental changes. Likewise, the interactions with other proteins or substrates, conformational flexibility, balance, and changes in secondary and higher-order structures, and the solvation pattern are of utmost interest in biochemistry. For insulin, for example, understanding the structure and dynamics is critical for its use in medicine.[1] This is still a challenging task, complicated by flexible parts of proteins without canonical structure.[2] The Raman optical activity (ROA) spectroscopy[3] has been increasingly used in the protein studies in the past decade, as it can directly provide, at least in principle, detailed information about the structure in solutions. Proteins do not have to be crystallized as in X-ray diffraction studies. Owing to the fast optical response the ROA spectra are always algebraic sums of conformer signals, unlike in NMR spectroscopy.

ROA detects a small difference in Raman scattered intensities corresponding to right or left circularly polarized light.[3,4] It can thus provide both the absolute configuration of a chiral molecule[5] and finer structural properties.[6] For example, in model peptides in solution, we demonstrated that conformational sensitivity of the ROA is fully comparable to NMR spectroscopy.[7] In some cases, ROA can capture conformers that cannot be resolved by other techniques.[6a,8]

In protein studies, ROA spectroscopy has been applied both to natural and modified amino acid sequences.[6a,9] ROA bands corresponding to protein parts of specific secondary structure have been identified and assigned in an empirical way, usually based on model peptides and a comparison to their crystal

structure. The unknown relation between the solution and crystal geometries makes this procedure problematic. Many assignments are thus only tentative. For instance, a negative ROA band at extended amide III region is exhibited by many secondary structures; a positive ROA band around 1310 cm^{-1} was assigned to the α-helix but without any knowledge about its dependence on the geometry or position in the peptide. Low-frequency ROA bands below 800 cm^{-1} were never assigned. Yet they often originate in large-scale delocalized vibrational modes and can be sensitive to the higher-order folding of protein chains.

It was thus soon recognized that quantum mechanical simulation of the spectra provide a more universal way to interpret the experiment.[10] Owing to the fast analytical implementations, it is nowadays possible to reliably calculate ROA spectra of small and medium molecules routinely within a reasonable time.[5a,11] However, for bigger, hydrated, or flexible molecules the calculations quickly become very costly in terms of required time and computer resources. The biggest systems subjected to quantum mechanical calculations of ROA so far included helical decaalanine,[6f] valinomycin (12 amino acids),[6a,8b] and a β domain of metallothionein (31 amino acids, vacuum calculation only).[12] The insulin molecule with 51 amino acids in one monomer unit thus represents a new challenge because of the steep dependence of computational time on the number of atoms.

To enhance the computations we use the Cartesian coordinate tensor transfer (CCT)[13] as a very flexible technique enabling ROA and other vibrational property calculations for very big molecules. It utilizes locality of vibrational properties, significantly reduces the computational time, and provides nearly the same accuracy as an exact calculation. The original big molecule is divided into smaller fragments, and their property tensors (force field, Raman and ROA polarizabilities, etc.) are calculated at a higher level of approximations and then transferred back to the original molecule, atom by atom. In our implementation[13,14] the corresponding atoms are chosen manually. Lately, an automatic procedure has been presented in a separate program.[15] In both cases, the assignment is based on the covalent bonding pattern and chemical similarity of atoms.

The CCT method provided, for example, vibrational circular dichroism (VCD)[16] or ROA[6a,b,8,17] spectra of large molecules. ROA modeling of flexible molecules can also be enhanced by CCT when a large number of conformers needs to be averaged.[6a,8] In some cases a larger error of transferred ROA tensors was reported than for Raman.[8b,15,18] However, some problems were caused by an error in computer implementation, and for a suitably chosen fragmentation scheme the residual error can be reduced to negligible levels.[18]

Bovine insulin was chosen as an example of hormone protein playing a crucial role in the regulation of mammalian metabolism. It is the most effective and durable drug in the treatment of diabetes[19] and is associated with some classes of obesity.[20] It contains 51 amino acids and about 800 atoms in its monomer state, depending on protonation. The dimer was also simulated in this work. As it is apparent from an X-ray structure of a dimer of bovine insulin shown in Figure 1 (PDB code 2A3G[21]) each monomer unit contains two helices and one strand in chain A and one helix and two strands in chain B. In total α-helix and 3$_{10}$-helix account for about 47% and 13% of the secondary structure, respectively.

The structure of insulin in solution has been studied as well, but a consensus on the detailed geometry was achieved neither for the native[22] nor the non-native states.[23] Previously reported ROA spectra of native insulin[24] exhibited a typical amide I (C=O stretching) ROA band, believed to be specific to α-helical proteins. A renewed interest in ROA of insulin is also due to the discovery that the technique can monitor structural changes under an amyloidal fibril formation, including intermediate denaturation states.[24c] Amyloid deposits are found, for example, in the pancreatic islets of patients with type 2 diabetes mellitus.[25] In the present study we chose the native insulin form with the most distinct spectral features, in order to rationalize the interpretation of the ROA spectra, to investigate their relation to the structure, and to obtain an insight into the molecular flexibility and vibrational properties. To the best of our knowledge the precise vibrational optical activity simulation of such a system has not been reported yet.
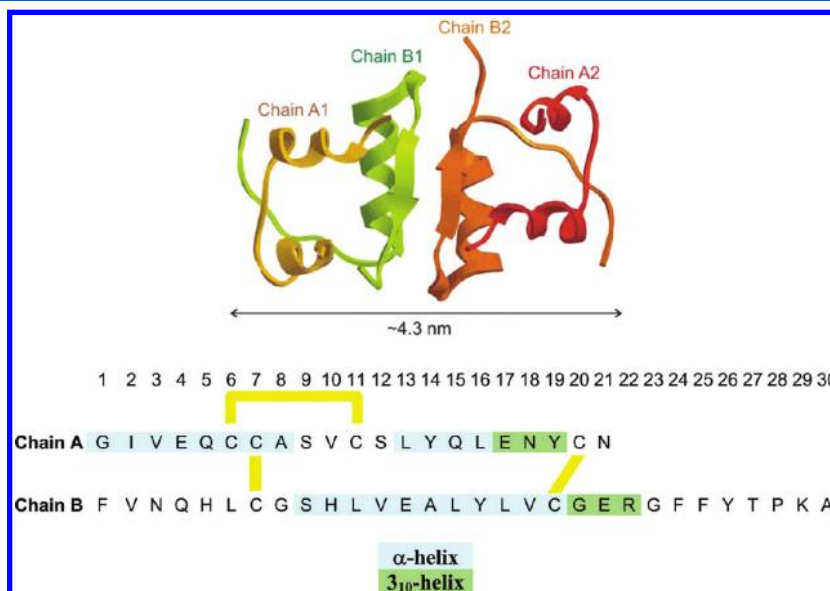

Figure 1. X-ray crystal structure of insulin dimer (PDB 2A3G[21]) and the monomer amino acid sequence with three disulfidic bridges (yellow).[26]

■ METHODS

Spectral Measurement. Bovine pancreas insulin was obtained from Sigma-Aldrich and dissolved in HCl aqueous (Milli-Q water) solution to the final concentration of 86 mg/mL. Backscattered Raman and ROA spectra were collected with ChiralRAMAN-2X (Biotools) operating with laser excitation at 532 nm and spectral resolution of 7 cm^{-1}. An additional edge-filter (Semrock) made it possible to lower the frequency cutoff to ∼90 cm^{-1} with respect to the excitation frequency. Laser power at the laser head was 800 mW; sample cell volume was 35 μL. Typically, the spectra were collected for 12 h. Third-order nine-point Savitzky-Golay smoothing was applied to raw ROA spectra, and the solvent background was subtracted from the Raman spectra (Figure S1). The experimental intensities are given in the number of electrons detected by the CCD camera.

ROA and Raman Spectra Calculations. The starting geometry of insulin monomer was extracted from the X-ray geometry[21] of bovine insulin (Figure 1, structural analysis and assignment of secondary structure in the X-ray geometry was performed with the aid of the STRIDE program[27]). The zinc ion was removed, and both terminal residues (NH_3^+, COOH) and His, Glu, Arg, and Lys side chains were protonated to mimic the experimental acidic pH. From this structure, smaller fragments containing four amino acid residues were generated along the entire peptide backbone. Two residues in each fragment were overlapped with the neighboring fragments to comprise the most important vibrational interactions within the main chain. Additional fragments were also created to consider hydrogen bonds and the strongest van der Waals interactions of side chains. In total, 49 fragments were produced, as specified in detail in Figures S2 and S3.

For analogous spectral simulations of an insulin dimer, an additional fragment was created for the contact region. To compensate for slight differences in the crystal geometries of the monomeric parts, the dimer geometry was obtained by rotating the monomer A1 and B1 chains around the C_2 symmetry axis.

The fragments were partially optimized in the normal mode coordinates by fixing the modes between i300 (imaginary) and 300 cm^{-1}.[28] This procedure enabled to relax high-frequency vibrational modes visible in the spectra and conserve the X-ray geometry, presumably close to the solvent structure. The B3PW91[29]/6-31++G** level was used in vacuum or with the CPCM[30] model of the aqueous environment. The optimization was performed with the Qgrad program,[31] and all ab initio computations were performed with Gaussian 09.[32] Harmonic force field (FF) and derivatives of the electric dipole-electric dipole polarizability (α) were calculated for the fragments at the same level as the geometry. The B3PW91 provided in the past excellent ROA spectral profiles for valinomycin.[6a] Local parts of the **G′** and **A** tensor derivatives needed for ROA[3] were calculated at the HF/rDP level. The rDP basis set was specifically optimized to provide good ROA spectra,[11c] and the combination of computational levels reduced the computational time. From the fragments, all tensors (FF, α, **G′**, and **A**) were transferred back to the whole insulin by CCT.[13,14] Backscattered Raman and ROA spectra were generated with Lorentzian peaks with full width of half-maximum (fwhm) of 10 cm^{-1} and the Boltzmann temperature correction for 298 K.

Molecular Dynamics (MD) Simulations. To investigate small-scale structural deviations caused by the temperature motion and their effects on the spectra, a short MD run within the TINKER[33] program environment was performed. According to the PDB X-ray structure specified above, insulin monomer geometry was rebuilt by the PROTEIN module of TINKER, by setting the principal torsion angles to the X-ray values, so that the Amber99 force field could be assigned to all atoms.[34] An intermediate energy minimization was performed, and the geometry was then placed in a cubic box (35.5 Å a side) of water molecules, yielding an insulin monomer surrounded by 1172 water molecules. After another minimization, production MD was run for 10 ns, within the periodic boundary conditions, NpT ensemble at 300 K and 1 atm, Beeman propagation, and 1 fs integration step. 10000 geometry snapshots were saved in 1 ps intervals and analyzed. For each geometry, molecular property tensors were obtained by CCT from the whole X-ray structure described in the previous section. As for the X-ray geometry, the Raman and ROA spectra were generated and averaged over the snapshot geometries.

Analysis of the Insulin Vibrational Motion. Potential energy distribution (PED) within the harmonic approximation enables a convenient assignment of molecular vibrational motions to local coordinates.[35] It was not feasible to manually define a nonredundant set of coordinates for the whole molecule (2343 or 4692 coordinates for monomer or dimer, respectively). Instead, a computer program was used for the monomer, in order to automatically define a redundant set of 4295 internal coordinates comprising all distances (stretchings), bond (bendings), and torsion angles based on the covalent bond pattern. The internal coordinate displacements $\Delta \mathbf{I}$ are proportional to the Cartesian displacements $\Delta \mathbf{X}$, which can be expressed as $\Delta \mathbf{I} = \mathbf{B} \, \Delta \mathbf{X}$. By bold letter we denote vectors and matrices. The normal mode vibrational coordinates $\mathbf{Q}$ were obtained by the usual procedure.[35,36] They are related to the displacements as $\Delta \mathbf{X} = \mathbf{S} \, \mathbf{Q}$. Then contribution of each coordinate i to the vibrational potential energy of mode m can be approximately estimated as $v_{i,m}' = (\sum_{\lambda=1}^{3Nat} S_{i,\lambda} B_{m,\lambda})^2$ or as a dimensionless quantity $v_{i,m} = v_{i,m}'/\sum_{i=1}^{M} v_{i,m}'$, where Nat is the number of atoms and $M = 4295$ is the number of internal coordinates. The X-ray geometry with transferred FF was used for this analysis.

The MD results were also analyzed, to provide information on larger-scale lower-frequency motions. Based on $N = 10000$ snapshots, we calculated for each residue a root-mean-square deviation $\delta_i = (\sum_{n=1}^{N} v_{n,i}/(NL_i))^{1/2}$ from an average geometry $\mathbf{r}_{ave}$, where $v_i = \sum_{\lambda=1}^{Li} (r_{n,\lambda} - r_{ave,\lambda})^2$ where λ runs over Li atomic coordinates in the residue i, and n is the snapshot index. To gain an insight into molecular interactions, two-residue correlation coefficients were also calculated as $cc_{i,j} = (\sum_{n=1}^{N} v_{n,i} v_{n,j})/((\sum_{n=1}^{N} v_{n,i}^2)(\sum_{n=1}^{N} v_{n,j}^2))^{1/2}$.

■ RESULTS AND DISCUSSION

Experimental Spectra. Experimental ROA and Raman spectra of insulin in HCl aqueous solution (pH < 1) are shown in Figure 2, upper traces in the two panels. They are very similar to previous measurements conducted for pH 2.5−3.1[24c] and pH 1.6.[24a,b] However, a better signal-to-noise ratio and a wider spectral range up to ∼80 cm^{-1} were achieved in this study. The range widening enabled to capture the positive ROA band at ∼180 cm^{-1} already indicated in a recent study,[24c] and a new ROA signal, the negative 131 and 91 cm^{-1} bands.

The stability of insulin in the wide range of pH is spectacular; its spectra in 20% acetic acid (Figure S4) and in 0.1 M HCl at pH 3 (not shown) are quite similar to those in Figure 2. According to literature insulin forms a monomer in 20% acetic
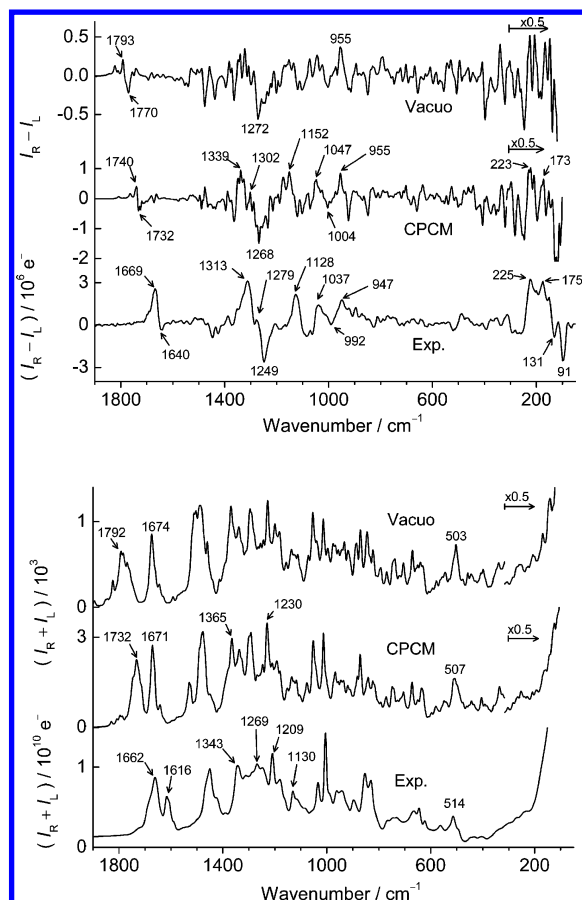
Figure 2. ROA (top) and Raman (bottom) spectra of monomer insulin calculated (B3PW91/6-31++G**//HF/rDP) in vacuo and with the CPCM water model, and experimental data (86 mg/mL in 0.5 M HCl). Note the intensity scale change in the calculated spectra below 350 cm^{-1}.

acid, dimer at pH 1.6 (HCl), and a tetramer at pH 3.0 (HCl) at concentration up to 2 mg/mL.[23a,37] Even at concentrations needed for ROA experiments (100 mg/mL, pH 1.6 in HCl), the dimeric form was determined as prevalent.[24b] Although we may thus suppose that we measured it as dimer in HCl (pH < 1) and monomer in the acetic acid, the difference in the association state was not reflected in the ROA or Raman spectral shapes within the experimental error.

Previously, some ROA bands of insulin were assigned to certain structural features on empirical grounds. These include the ± amide I couplet around 1650 cm^{-1}, typical for α-helical proteins;[24b] in our experiment, the magnitude of the positive lobe (1669 cm^{-1}) is significantly larger than that of the negative one (1640 cm^{-1}). The ROA peaks in the extended amide III region between 1340 and 1200 cm^{-1} were also recognized to contain the $^\alpha$C−H and N−H deformations and the $^\alpha$C−H stretching and to be sensitive to the secondary structure of proteins. The positive peak at 1313 cm^{-1} has been assigned to α-helices in hydrophobic environment,[24b,38] and the negative peak at 1250 cm^{-1} was suggested to characterize the antiparallel β-sheet of the C-terminus of the chain B of insulin dimer.[24b]

Vacuum and CPCM Simulations of the Spectra. The main experimental features in the Raman and ROA spectra are reproduced by both vacuum and PCM computations, including the ROA sign pattern (Figure 2, upper two traces in each panel). However, the CPCM correction clearly provides a better

agreement in spectral shapes, relative intensities, and frequencies within the entire spectral region. For example, the amide I Raman band calculated to be placed at 1792 cm^{-1} in vacuum shifts to 1732 cm^{-1} upon the CPCM correction, much closer to the experimental value of 1662 cm^{-1}. The remaining error can be explained by explicit hydrogen bonding and anharmonic effects.[39] The calculated negatively biased amide I ROA couplet becomes more positive in CPCM, although the positive bias in experiment is larger. This can perhaps be explained by hydrogen bonding of the carboxyl groups, incompletely represented by the CPCM solvent model.[39c,40] The results are in agreement with previous findings for peptides and proteins where the Raman and ROA signals are dominated by nonpolar molecular parts, reasonably well simulated in vacuum. For example, for metallothionein protein, a vacuum computation also yielded the main ROA experimental features.[12] Similarly, in the case of valinomycin potassium complex and its ROA mostly stemming from the isopropyl residues, a negligible effect of the solvent was found.[8b]

As also observed in a model hexapeptide with aromatic residues,[6b] the ring vibrations of tyrosine at ~1616, ~1050, ~750 cm^{-1}, and histidine at ~1550 cm^{-1} yield very strong Raman signals, perhaps overestimated in the computations when compared to an average intensity; this may be attributed to insufficient accuracy of the DFT method and the harmonic approximation for the aromatic systems.[41] In spite of internal achirality of the aromatic chromophores, they can also significantly contribute to ROA via through-space polarization and chiral covalent links.[6e]

At lower wavenumbers, the vacuum and CPCM frequencies are closer to each other. However, smaller shifts and a high density of vibrational transitions still bring about significant changes in the spectral shapes under theoretical hydration, e.g. the intensity loss of the 1405 cm^{-1} Raman band, or the more distinct ROA pattern within the 1400−900 cm^{-1} range. Around 1300 cm^{-1} (predominantly CH bending) the calculated CPCM frequencies are still higher by ~20 cm^{-1} than in the experiment, most likely due to some anharmonicity of the force field.[42] Below 900 cm^{-1}, the solvent typically weakens the relative ROA signal.

Below 350 cm^{-1}, the Raman intensity is almost featureless, and its magnitude in the experiment may be affected by the incomplete subtraction of water or cell-window signal. Nevertheless, the computations reproduce the general trend fairly well, including the sharp intensity increase below 190 cm^{-1}. Rather surprisingly, this region contains a rich ROA spectrum, e.g. the −/+/+ broad bands at 280/225/175 cm^{-1} and a sharp negative band at 91 cm^{-1}. These can be interpreted by the CPCM computation although the arbitrary bandwidth of 10 cm^{-1} seems too narrow for this region. The vacuum computation is less accurate, but it still provides the basic spectral pattern. Typically, methyl torsion, amide out-of-plane motions, and delocalized vibrations contribute to these spectral bands.

The potential energy distribution of insulin's internal coordinates as calculated for the CPCM model is exemplified in Figure 3. Most vibrations contain many coordinates/molecular groups, but some characteristic vibrations can be localized within the entire range of measured frequencies, such as single- and double-bond CO stretching (~1130 and 1750 cm^{-1}) and OH bending (1150 cm^{-1}). In particular, the disulfidic cystine link has many distinct frequencies (CS stretch at ~700 cm^{-1}, CSS bending at 300 cm^{-1}, and SS torsion below 200 cm^{-1}) although their spectral intensity may be overlapped by other
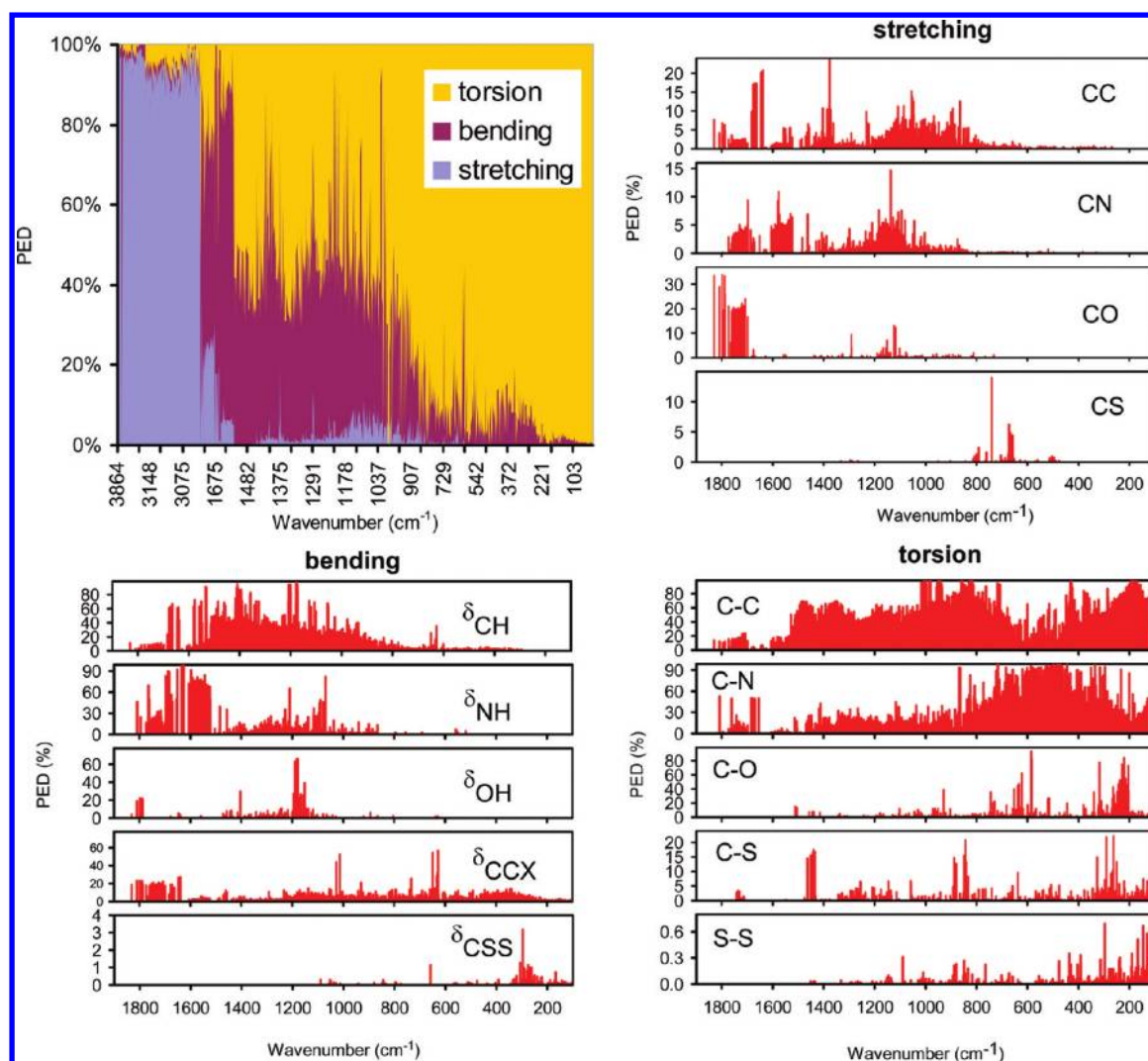
Figure 3. Calculated (B3PW91/CCT/6-31++G**/CPCM) potential energy distribution (PED, top left) of internal insulin coordinates and relative PEDs for the selected stretching, bending, and torsion coordinates for vibrations within 1900−100 cm⁻¹.

protein parts. In spite of the delocalized character of many normal mode vibrations the localized coordinates thus provide a good basis to analyze the relation of the ROA spectrum to the secondary structural features. Note that other analyses based on the ROA normal mode intensity,[43] localized vibrations,[44] or group ROA contributions[45] have also been proposed.

Based on the spectral agreement we can conclude that the insulin crystal structure is by and large maintained in solution. The occasional smoothing and/or broadening of the experimental ROA bands if compared to the one-conformer computations is explicable by a higher flexibility of insulin in solution.[6d]

The Magnitude and Basis Set Dependence of Intrinsic ROA Component. The ROA signal can be approximately thought of as a sum of dipolar interactions between chirally oriented groups dependent on the electric polarizability α and irreducible components requiring derivatives of the optical activity tensors $\mathbf{G}'$ and $\mathbf{A}$.[3] Note that in a common origin $\mathbf{G}'$ and $\mathbf{A}$ also contain an α-contribution. For small molecules either component (α and $\mathbf{G}'/\mathbf{A}$) needs to be included for the spectra simulation to be accurate. For larger systems, several studies indicated that the irreducible component can sometimes be neglected.[10b,46] This is also true for insulin. As shown in Figure 4 the polar (α) component is quite dominant and within
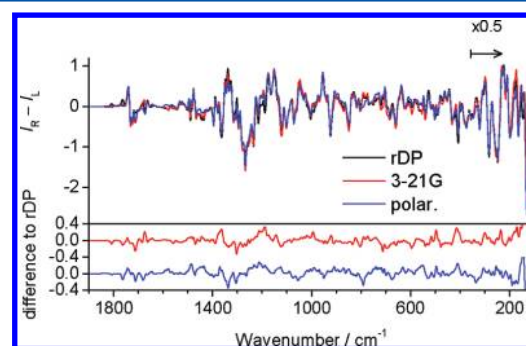

Figure 4. ROA spectra of monomer insulin with the local parts of the $\mathbf{G}'$ and $\mathbf{A}$ intensity tensors calculated at the HF/rDP (black line) and HF/3-21G (red) levels. In the polarization model based on the polarizability α (blue line), these tensors are omitted. The differences against the HF/rDP result are magnified at the bottom, FF and the α tensor were calculated using B3PW91/6-31++G**/CPCM(water), intensity below 350 cm⁻¹ was reduced by a factor of 2.

the polarization model[47] provides a spectral shape which is very close to that obtained while including the irreducible $\mathbf{G}'$ and $\mathbf{A}$ parts. We explain this by the higher sensitivity of the $\mathbf{G}'/\mathbf{A}$ part to the geometry variations than to the longer-distance dipolar

interactions and subsequent averaging of the G'/A contribution over the many amino acid residues.

The relative independence of the irreducible ROA signal (cf. lower part of Figure 4) on the basis set represents yet another surprising aspect. In general, ROA intensities are much more sensitive to the basis set variations than are the Raman ones.[10b,11a,48] However, such sensitivity involves a variation of the force field and the α-polarizability component. The variation under the rDP and HF/3-21G basis set change is rather modest for most ROA bands although very significant differences can be observed under 200 cm^{-1}.

The results indicate that the polarization model is suitable for the calculation of ROA in big molecules, as proposed earlier in other studies of peptidic systems.[17,47,49] This model reduces the computational cost of evaluating the G' and A tensor derivatives. Nevertheless, in practice this convenience is relatively minor (although still notable for insulin) given the much higher computational cost of calculating the force field. Likewise, the fast and analytical techniques implemented at the DFT level allow for cost-efficient computations of all the intensity tensors.[10b,11a−d,f]

Effect of Dimerization on the Spectra. In principle, ROA and Raman spectra reflect the dimerization in many ways, as local changes of polarizability tensors, mutual polarization of molecular parts, and coupling of vibrational motions.[8b,18,45b] In the dimer, an antiparallel β-sheet structure between the B-chains connects the two units. In the simulated spectra of insulin in Figure 5, however, the effects are quite minor, especially above 200 cm^{-1}. This is, nevertheless, consistent with the observation that the two solvent environments (pH < 1 HCl, and 20% acetic acid) favoring dimer and monomer provide nearly the same spectra. Clearly, the signal from the contact atoms is too small and overlapped with other transitions, which was verified by simulation of the ROA spectrum of the linking part only (not shown).

A more detailed inspection of Figure 5 showing the calculated spectra does reveal some small but distinct changes caused by the dimerization. In the dimer, amide I ROA couplet (~1735 cm^{-1}) becomes more negative, and the +/− couplet at ~950 cm^{-1} gets weaker compared to the monomer. Previously, the negative peak at 1250 cm^{-1} in experimental ROA spectrum was assigned to the antiparallel β-sheet of the C-terminus of the chain B in insulin dimer.[24b] Our simulation does not support this theory, as the band is developed already in the monomer. ROA peaks below 200 cm^{-1} seem to be more sensitive to the dimerization, but these are difficult to analyze. Similar relative insensitivity to the mutual peptide chain interactions can also be observed for the summary signal of chains A and B in monomer (Figure S5). The changes in the Raman spectra (Figure 5, bottom) are tiny as well, e.g. slight intensity redistribution in the CH bending 1200−1400 cm^{-1} region.

Contribution of the Main and Side Chains to the ROA and Raman Spectra. Within the monomer model, approximate contributions of the amide, $^{\alpha}CH$, $^{\alpha}CH^{\beta}CH$, and more distant side chain groups were obtained by deleting the relevant atomic derivatives of tensors α, G', and A. The difference spectra, I_{total} - $I_{deleted}$, with the B3PW91/6-311++G**/CPCM(water) FF provided spectra shown in Figure 6. They reveal how the amino acid parts contribute to the intensities. For example, ROA and Raman amide I bands at ~1700 cm^{-1} indeed originate predominantly from the amide groups, with a small contribution from side chain carbonyls. Note that the local chiral centers $^{\alpha}CH$ do not contribute to this ROA band. Thus, similar to the vibrational circular dichroism,[50] amide I ROA signal reflects mostly the secondary structure of the protein.

These data suggest that the amide I ROA signal was generated by a chiral arrangement of amide groups in the secondary or higher peptide structures, rather than by a local chirality of the amino acid, in turn explaining the sensitivity of this ROA band to the secondary structures of proteins in general.

The ROA pattern in the frequency region of 1500−800 cm^{-1} mostly originates from both the $^{\alpha}CH^{\beta}CH_n$ moiety (with the dominance of $^{\alpha}CH$) and the amide groups. This is consistent with some previous works where these bands were found to be sensitive to the rotation of the $^{\alpha}C−^{\beta}C$ bond, that is to the conformation of peptide side chains.[6b,8b] The coupling between the side and main chain groups, and perhaps also the contribution of the more distant side chain parts (Figure 6), make the ROA bands quite complex but at the same time sensitive to the higher-order structures as encountered in some experiments.[24b] Unlike for ROA, distant side chains contribute more than the
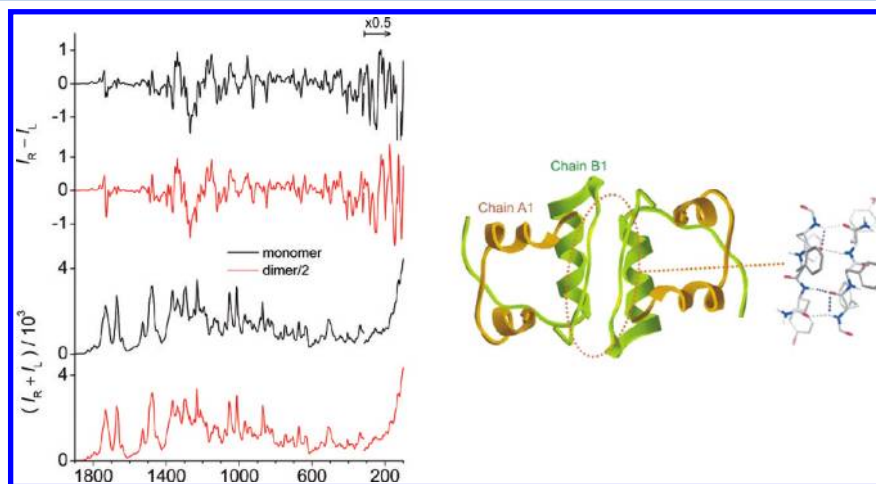

Figure 5. Comparison of the calculated (B3PW91/6-31++G**/CPCM(water)/HF/rDP) normalized insulin monomer and dimer ROA (I_R - I_L) and Raman (I_R + I_L) spectra. The intensity scale changes at 350 cm^{-1}. The additional contact region fragment (between the monomer units) used in the CCT calculation is displayed to the right.
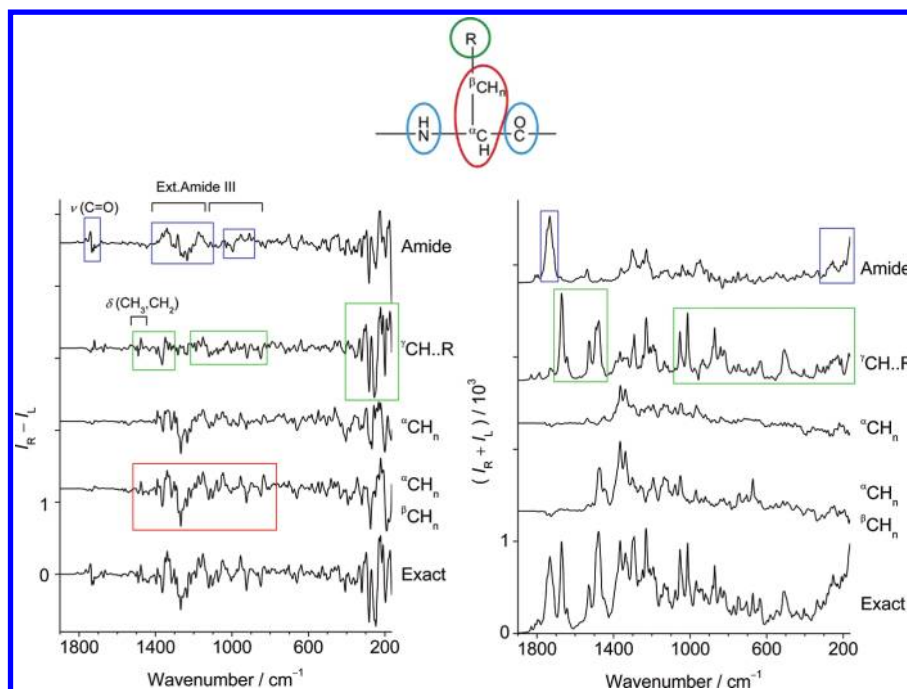
Figure 6. Approximate decomposition of the Raman and ROA insulin spectra into contributions of the amide group (blue), distant parts of the side chains (green), and the $^{\alpha}$CH-$^{\beta}$CH$_2$ moiety (red). The $^{\alpha}$CH contribution is also plotted.

backbone to the Raman intensities, because they often contain bulky and polarizable but not chiral residues.

ROA bands between ~800 and ~400 cm^{-1} are relatively weak and originate in all protein parts, whereas the Raman spectrum in this region is still strong, dominated by the side chains. Below 400 cm^{-1}, the contributions of the side chains to ROA are very large, via participations of the CC and CN torsion (Figure 3) on the vibrations. These bands should thus be more sensitive to the tertiary structure of the protein, which is consistent with the dimerization example shown in Figure 5.

Relation between the Spectra and Protein Secondary Structure. Similarly as for the backbone and side chain parts (Figure 6) we decomposed the spectra of monomer insulin into selected amino acid sequences in chains A and B according to their secondary structure in the main structural units[27] (Figure 7). Corresponding Raman spectra can be seen in Figure S6, alternate decomposition to the helical and coil/strand parts only in Figure S7. A comparison of the ROA spectra of chains A and B revealed that the strongest positive ROA band at 1339 cm^{-1} (1313 cm^{-1} in experiment) chiefly originates from chain A. This is somewhat surprising because both chains contain a high portion of α- or 3$_{10}$-helical structure. However, in chain B, only a weak signal is predicted in this region, as the positive peak of the 3$_{10}$-helical B20−22 residues at ~1330 cm^{-1} is compensated by a negative peak of the α-helical B9−19 residues at ~1318 cm^{-1}. Currently we cannot indentify any particular differences in the α-helical geometries of the A and B strands that would be directly linked to the different spectral behavior. In addition to the empirical assignment of this band to α-helix[24b] our results show that this ROA band mainly comes from the chain A and that the 3$_{10}$-helical structure also significantly contributes to it. The decomposition (Figure 7) shows that helical parts significantly contribute to the amide I ROA band, which is consistent with the previous empirical assignment.[24b] However,

the coil and strand parts in total contribute to about 20% of the intensity in this region (Figure S7).

The sharp positive peak at 1302 cm^{-1} in the calculated ROA spectrum of chain B (Figure 7) corresponds to the experimentally observed shoulder peak at 1279 cm^{-1}. Only the coil and strand residues (B1−8, 23−30) exhibit this positive sharp feature which can be assigned to δ (C$_{\alpha}$H), δ (C$_{\beta}$H), and in-plane amide vibrations. The 1305 cm^{-1} ROA peak of B1−8 originates from the head of chain B (Phe1, Val2, Asn3), while the B23−30 peak at 1302 cm^{-1} originates from its tail (Phe24, Phe25, Tyr26). In a previous study of insulin amyloid fibrils[24c] a similar sharp positive ROA peak was detected at 1271 cm^{-1} and assigned to a β-turn structure. It is close to the shoulder peak at 1279 cm^{-1} of the native insulin and possibly marks the strand formation in the fibrils.

The negative ROA band at 1268 cm^{-1} in the calculated ROA spectrum is present in all insulin parts (Figure 7), although its central frequency in the coils and strands (1235 cm^{-1}) is lower than in the helices (1269 cm^{-1}). Another strong positive ROA peak (at ~1150 cm^{-1} in calculation, 1128 cm^{-1} experimentally) clearly comes from the α-helical parts (A1−8, A13−16, B9−19) and thus can be used as an additional indicator of this conformation.

In Figure 8, the ROA spectra of the three α-helical parts are compared; the main ROA patterns are relatively indifferent to the structural variations, such as the broad −/+/− ROA couplet at ~1270/1152/1100 cm^{-1}, and are indeed experimentally observed in most spectra of α-helical proteins.[24b] We should note, however, that the disordered poly-L-glutamic acid also exhibits such a pattern.[24b] The feature is absent for the 3$_{10}$-helices (A17−19, B20−22), which thus provides a convenient opportunity to discriminate between these two similar structure forms. Also the strong positive/weak negative couplet of the amide I band has been assigned to a characteristic of α-helical protein.[24b] The calculations (Figures 7 and 8) nevertheless
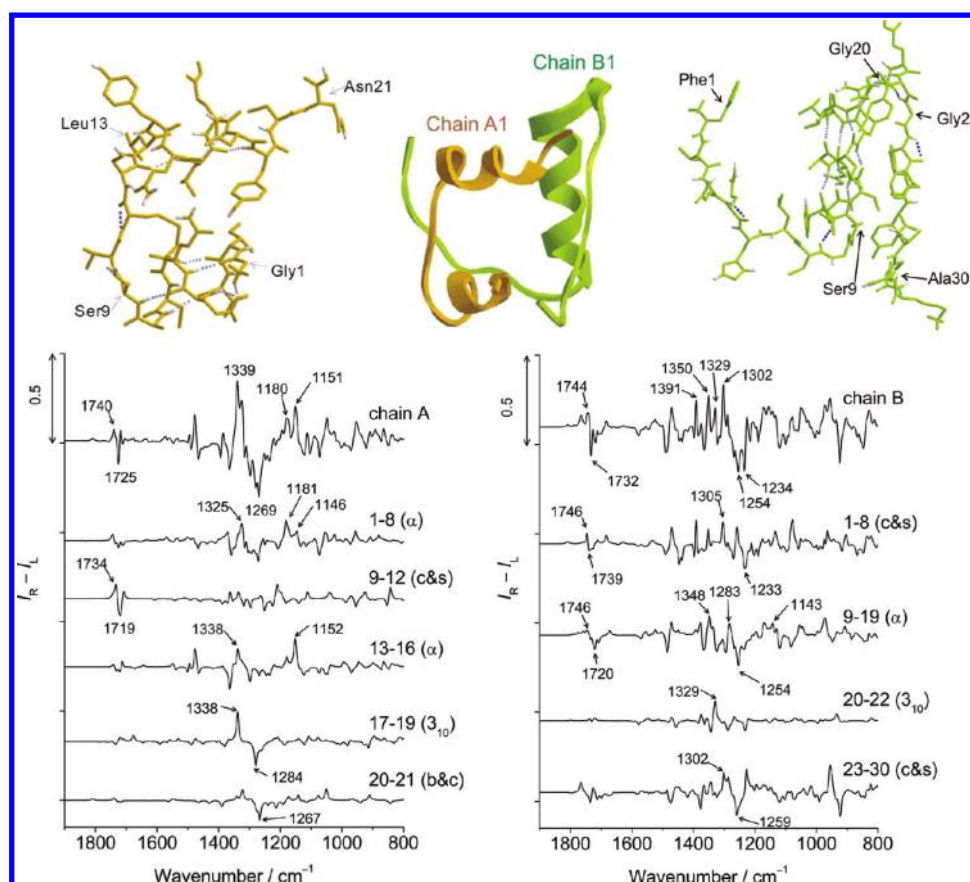
Figure 7. Calculated ROA spectra of specific amino acid sequences in the A chain (left) and B chain (right) of insulin monomer. Chain A: Gly1-Ala8 (α-helix), Ser9-Ser12 (coil and strand), Leu13-Leu16 (α-helix), Glu17-Asn21 (3_{10}-helix), and Cys20-Asn21 (bridge and coil). Chain B: Phe1-Gly8 (coil and strand), Ser9-Cys19 (α-helix), Gly20-Arg22 (3_{10}-helix), and Gly23-Ala30 (coil and strand).

reveal that this couplet may also come from the coils and strands (A9−12, B1−8).

Insulin Dynamics and Motional Broadening of Spectral Bands. It has been recognized in the past that the Raman and ROA bands reflect not only the equilibrium structure but also the conformer equilibria[5a,51] and structural fluctuations[6d,41] caused by thermal motion. This can be documented on the ROA and Raman spectra based on averaging of 10000 MD snapshots as compared to both the equilibrium X-ray calculation and experiment in Figure 9.

In the Raman spectrum the averaging does not change the principal features of the crystal-based simulation but many peaks become smoother, and below 1000 cm⁻¹ some weaker bands disappear altogether. Even within the 1200−1400 cm⁻¹ interval the relative intensity pattern changes by the averaging. Overall, the MD curve is much closer to experiment than for the one-conformer model. Further improvement of the simulated spectral shape would primarily require a model explicitly accounting for solvent and carbon hydrogen bonding,[39a,52] which is currently not feasible for insulin, and it would probably also broaden the amide I band more realistically.[39b,53]

The MD-averaged ROA spectrum (Figure 9, left) also conserves the main intensity features obtained by the crystal geometry simulation above 800 cm⁻¹. However, a profound drop in intensity caused by the averaging is apparent below 800 cm⁻¹. This is caused by low-frequency torsional vibrations (Figure 3) of the protein chain and a high sensitivity of the ROA intensity to these geometrical changes. The MD ROA
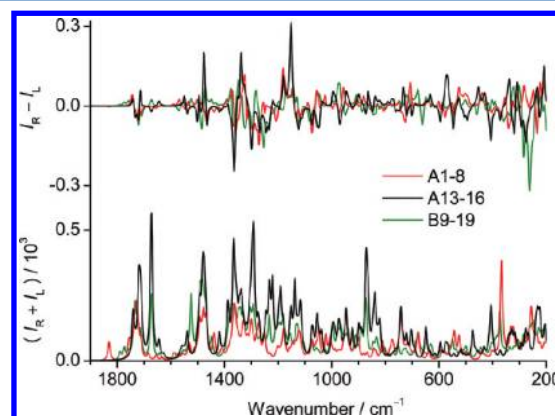

Figure 8. Calculated ROA (top) and Raman (bottom) spectra of the three α-helical segments of insulin, comprising residues 1−8 and 13−16 in chain A and 9−19 in chain B. The intensity is normalized to one amino acid residue.

pattern is thus much closer to the experiment in the low-frequency region. In addition, the ratio of intensities within the 100−800 cm⁻¹ and 800−1800 cm⁻¹ region is more realistic. The too weak signal within the 100−300 cm⁻¹ interval suggests that the averaging based on the classical trajectory may be too extensive, as also observed for other simulations of similar kind.[8a] The ROA signal above 800 cm⁻¹ is not significantly improved by the averaging. In particular, the already weak amide I signal (experimentally at ∼1660 cm⁻¹) further disappears, unlike in
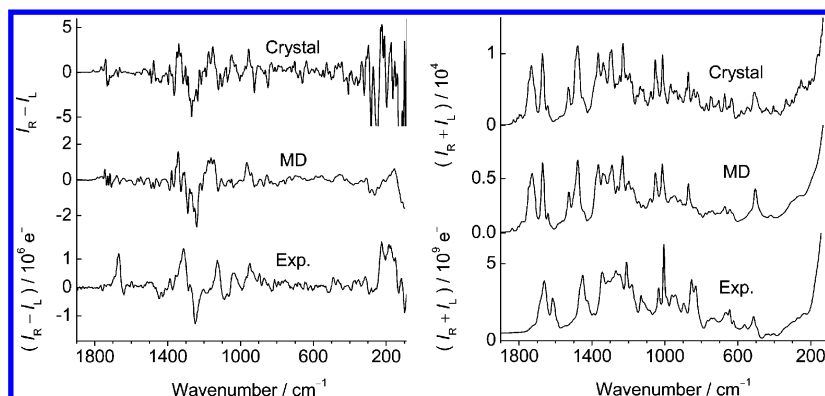
Figure 9. Calculated ROA (left) and Raman (right) of the insulin monomer based on the crystal structure (top), averaged MD snapshots (middle), and the experiment (bottom).

experiment, probably because of inaccuracies of the CPCM solvent model and MD force field.

The time of 10 ns comprised by MD is obviously too short to provide a conformational change of the entire protein.[54] Nevertheless, neither the experimental results nor the time dependence of geometry parameters obtained during MD indicate a propensity of insulin to denature or change folded state. We thus suppose that the simulation does realistically describe the low-amplitude temperature fluctuations of the protein. Previous simulations also suggest that the monomer model is indeed adequate as the two insulin units' motion is almost independent.[55]

The rigidity can also be estimated from the root-mean-square deviations from the equilibrium structure obtained within the MD run (Figure S8). In most cases the deviations are fairly low, below 1 Å, except for the flexible C-terminus of chain A (Figure 1). Rather surprisingly, with respect to the deviations the α-helical sections do not exhibit a different flexibility behavior than the other insulin parts. This is nonetheless consistent with the observed high stability of insulin in a range of pH, indicating a stable main and side chain network. The cysteine residues exhibit highest rigidity (low δ of 0.2−0.3 Å) because of the cystine disulfide bridge.

The correlation coefficients between motions of the amino acids residues displayed as a 2D graph in Figure 10
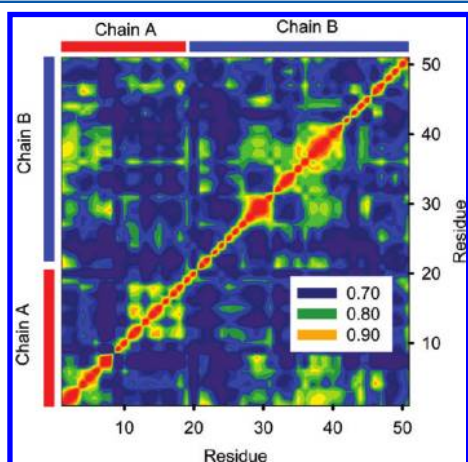

Figure 10. Correlation coefficients between atomic deviations of insulin residues calculated from the molecular dynamics trajectory.

provide another interesting insight into the protein dynamics. We can see that the motion residues that are

closed to each other in the helical segments (1−8, 13−19, 30−40) is more correlated than the rest, which is due to the intramolecular hydrogen bonding. Moreover, even the β-strand part of chain B (25−30) in contact with the helix (Figure 1) is correlated with the helical residues. Finally, the motion of the initial part of chain A (∼1−4) exhibits strong a correlation with amino acids 35−40 in chain B, which clearly reflects the anchoring of chain A in the α-helical part of chain B. Even though rather complex, these correlations can explain some differences in spectral profiles of various α-helical peptide segments (Figures 7 and 8).

Relatively large oscillations of the peptide backbone occurred during the dynamics (Figure S9). Side chain residues exposed to the aqueous environment were in general more flexible than residues buried in the protein. In spite of the fluctuations, the dynamics provided a converged distribution of insulin coordinates, mostly oscillating around equilibrium positions. Computed probability distributions of selected coordinates are plotted in Figure S10 and compared to the X-ray data. The average distance of chain A termini obtained by MD, for example, is almost identical to the crystalline value but with a relatively broad (± ∼ 2 Å) distribution. According to MD, chain B relaxes more in solution, its ends move further apart by about 2 Å, and the α-helix becomes longer by 1.5 Å if compared to the crystal. Some side chains rotate nearly freely in solutions, such as the serine and histidine residues shown in lower part of Figure S10.

Except for the slight relaxation and oscillation, the structure appeared sufficiently rigid to allow for a plain averaging of the coordinates (after appropriate rotation and overlap of MD snapshots). This average MD geometry is compared to the crystalline one in Figure 11. Clearly, the basic folding pattern remains unchanged. On the other hand, finer differences are apparent not only in the relatively loose strand parts but also in the helical segments. The better agreement of the ROA spectra with experiment if compared to the crystal geometry simulation (Figure 9) confirms that the trends in structural changes of insulin under solvation predicted by MD are realistic.

In summary, we can conclude that the possibility to simulate the protein ROA in full and to analyze fine geometric and dynamical effects significantly broadens the insight into the molecular structure and dynamics that can be obtained by this technique.
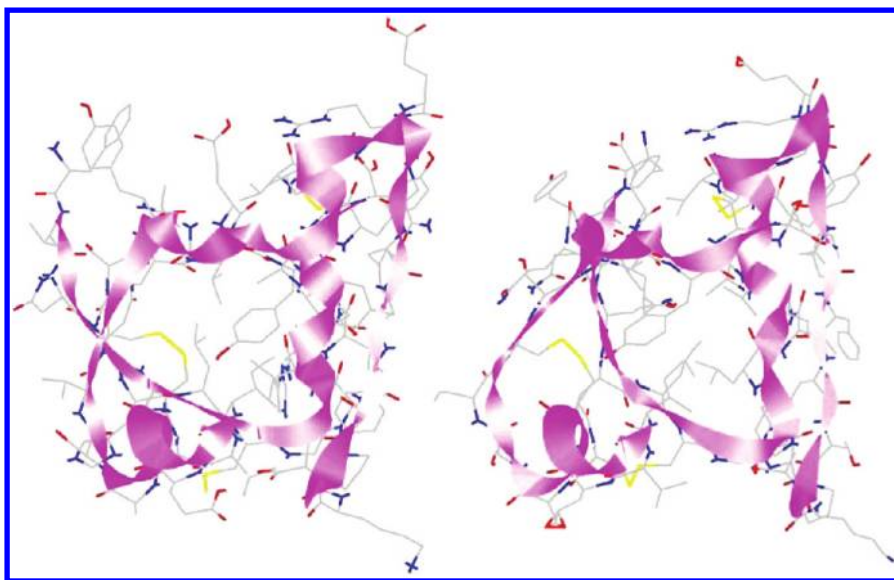
Figure 11. X-ray (PDB 2A3G, left) and average MD (right) insulin geometries.

CONCLUSIONS

Insulin is the largest molecule for which the Raman and ROA spectra were ever interpreted by quantum-chemical procedures. The calculation was made possible by employing the latest analytical derivative techniques and the Cartesian coordinate tensor transfer (CCT) method. The spectra simulated with the X-ray geometry explained virtually all relative intensities in experiment conducted in aqueous solution. The spectral profile could be further improved by accounting for the motional averaging within the combined MD/CCT computation. The results indicate that insulin conserves the crystal structure in solution, except for low-frequency motions of some peptide parts. ROA and Raman spectra of insulin dimer were also calculated; while the dimerization produced small spectral changes in theory, they were too small to be detected experimentally.

The full-protein vibrational spectra simulation significantly enhances the information about the molecular structure that can be obtained from the spectra. It confirmed or corrected previous empirical assignment to a specific peptide secondary structure, and provided new marker bands. Contributions to the resultant spectral profiles of the main and side amino acid chains or specific amino acid sequences could be deciphered. Likewise, the analysis of protein vibrations made it possible to assign specific spectral bands to local molecular groups. Protein ROA exhibited some features not observed for smaller molecules, in particular the dominance of the electric dipole polarizability contribution. We thus believe that the ROA spectroscopy and the vibrational simulation techniques combined can reveal many important properties of macromolecules similar to the insulin hormone.

ASSOCIATED CONTENT

Ⓢ Supporting Information

Experimental and computational details. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: aporoa@gmail.com (S.Y.), bour@uochb.cas.cz (P.B.).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Kim, T.; Rhee, A.; Yip, C. M. *J. Am. Chem. Soc.* **2006**, *128*, 5330−5331.

(2) (a) Uversky, V. N. *Proteins* **2000**, *41*, 415−427. (b) Dunker, A. K.; Romero, P.; Obradovic, Z. *Biophys. J.* **1998**, *74*, A132 Part 2.

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

(4) (a) Barron, L. D.; Buckingham, A. D. *Chem. Phys. Lett.* **2010**, *492*, 199−213. (b) Barron, L. D.; Buckingham, A. D. *Mol. Phys.* **1971**, *20*, 1111−1119. (c) Barron, L. D.; Bogaard, M. P.; Buckingham, A. D. *J. Am. Chem. Soc.* **1973**, *95*, 603−605. (d) Hug, W.; Kint, S.; Bailey, G. F.; Schere, J. R. *J. Am. Chem. Soc.* **1975**, 97.

(5) (a) Haesler, J.; Schindelholz, I.; Riguet, E.; Bochet, C. G.; Hug, W. *Nature* **2007**, *446*, 526−529. (b) Zuber, G.; Hug, W. *Helv. Chim. Acta* **2004**, *87*, 2208−2234.

(6) (a) Yamamoto, S.; Watarai, H.; Bouř, P. *ChemPhysChem* **2011**, *12*, 1509−1518. (b) Hudecová, J.; Kapitán, J.; Baumruk, V.; Hammer, R. P.; Keiderling, T. A.; Bouř, P. *J. Phys. Chem. A* **2010**, *114*, 7642−7651. (c) Luber, S.; Reiher, M. *ChemPhysChem* **2010**, *11*, 1876−1887. (d) Kapitán, J.; Baumruk, V.; Kopecký, V. Jr.; Bouř, P. *J. Phys. Chem. A* **2006**, *110*, 4689−4696. (e) Jacob, C. R.; Luber, S.; Reiher, M. *ChemPhysChem* **2008**, *9*, 2177−2180. (f) Herrmann, C.; Ruud, K.; Reiher, M. *ChemPhysChem* **2006**, *7*, 2189−2196. (g) Mukhopadhyay, P.; Zuber, G.; Beratan, D. N. *Biophys. J.* **2008**, *95*, 5574−5586.

(7) Buděšínský, M.; Daněček, P.; Bednárová, L.; Kapitán, J.; Baumruk, V.; Bouř, P. *J. Phys. Chem. A* **2008**, *112*, 8633−8640.

(8) (a) Kaminský, J.; Kapitán, J.; Baumruk, V.; Bednárová, L.; Bouř, P. *J. Phys. Chem. A* **2009**, *113*, 3594−3601. (b) Yamamoto, S.; Straka, M.; Watarai, H.; Bouř, P. *Phys. Chem. Chem. Phys.* **2010**, *12*, 11021−11032.

(9) (a) Wen, Z. Q.; Hecht, L.; Barron, L. D. *J. Am. Chem. Soc.* **1994**, *116*, 443−445. (b) Barron, L. D.; Hecht, L.; Ford, S. J.; Bell, A. F.; Wilson, G. *J. Mol. Struct.* **1995**, *349*, 397−400. (c) Ford, S. J.; Cooper,

2449

dx.doi.org/10.1021/ac2032436 | *Anal. Chem.* 2012, 84, 2440−2451

A.; Hecht, L.; Wilson, G.; Barron, L. D. *J. Chem. Soc., Faraday Trans.* **1995**, *91*, 2087−2093. (d) Nafie, L. A.; Yu, G. S.; Freedman, T. B. *Vib. Spectrosc.* **1995**, *8*, 231−239. (e) Wilson, G.; Ford, S. J.; Cooper, A.; Hecht, L.; Wen, Z. Q.; Barron, L. D. *J. Mol. Biol.* **1995**, *254*, 747−760. (f) Wilson, G.; Hecht, L.; Barron, L. D. *J. Chem. Soc., Faraday Trans.* **1996**, *92*, 1503−1509. (g) Blanch, E. W.; Bell, A. F.; Hecht, L.; Day, L. A.; Barron, L. D. *J. Mol. Biol.* **1999**, *290*, 1−7. (h) Jalkanen, K. J.; Nieminen, R. M.; Knapp-Mohammady, M.; Suhai, S. *Int. J. Quantum Chem.* **2003**, *92*, 239−259. (i) Blanch, E. W.; McColl, I. H.; Hecht, L.; Nielsen, K.; Barron, L. D. *Vib. Spectrosc.* **2004**, *35*, 87−92.

(10) (a) Polavarapu, P. L.; Hecht, L.; Barron, L. D. *J. Phys. Chem.* **1993**, *97*, 1793−1799. (b) Ruud, K.; Thorvaldsen, J. *Chirality* **2009**, *21*, E54−E67.

(11) (a) Cheeseman, J. R.; Frisch, M. J. *J. Chem. Theory Comput.* **2011**, *7*, 3323−3334. (b) Cheeseman, J. R.; Shaik, M. S.; Popelier, P. L. A.; Blanch, E. W. *J. Am. Chem. Soc.* **2011**, *133*, 4991−4997. (c) Zuber, G.; Hug, W. *J. Phys. Chem. A* **2004**, *108*, 2108−2118. (d) Reiher, M.; Liegeois, V.; Ruud, K. *J. Phys. Chem. A* **2005**, *109*, 7567−7574. (e) Šebestík, J.; Bouř, P. *J. Phys. Chem. Lett.* **2011**, *2*, 498−502. (f) Liegeois, V.; Ruud, K.; Champagne, B. *J. Chem. Phys.* **2007**, *127*, 204105.

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

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

(14) Bouř, P. Academy of Sciences: Prague, 1997−2009.

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

(16) (a) Andrushchenko, V.; Bouř, P. *J. Comput. Chem.* **2008**, *29*, 2693−2703. (b) Andrushchenko, V.; Bouř, P. *J. Phys. Chem. A* **2007**, *111*, 9714−9723. (c) Andrushchenko, V.; Wieser, H.; Bouř, P. *J. Phys. Chem. B* **2002**, *106*, 12623−12634. (d) Kubelka, J.; Keiderling, T. A. *J. Am. Chem. Soc.* **2001**, *123*, 12048−12058.

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

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

(19) Arbid, E.; Kidron, M. *J. Diabetes Sci. Technol.* **2009**, *3*, 562−567.

(20) Kahn, S. E.; Hull, R. L.; Utzschneider, K. M. *Nature* **2006**, *444*, 840−846.

(21) Smith, G. D.; Pangborn, W. A.; Blessing, R. H. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2005**, *61*, 1476−1482.

(22) (a) Mark, A. E.; Berendsen, H. J. C.; van Gunsteren, W. F. *Biochemistry* **1991**, *30*, 10866−10872. (b) Uversky, V. N.; Garriques, L. N.; Millett, I. S.; Frokjaer, S.; Brange, J.; Doniach, S.; Fink, A. L. *J. Pharm. Sci.* **2003**, *92*, 847−858. (c) Legge, F. S.; Budi, A.; Treutlein, H.; Yarovsky, I. *Biophys. Chem.* **2006**, *119*, 146−157. (d) Todorova, N.; Legge, F. S.; Treutlein, H.; Yarovsky, I. *J. Phys. Chem. B* **2008**, *112*, 11137−11146.

(23) (a) Nielsen, L.; Frokjaer, S.; Brange, J.; Uversky, V. N.; Fink, A. L. *Biochemistry* **2001**, *40*, 8397−8407. (b) Jiménez, J. L.; Nettleton, E. J.; Bouchard, M.; Robinson, C. V.; Dobson, C. M.; Saibil, H. R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9196−9201.

(24) (a) Wen, Z. Q.; Hecht, L.; Barron, L. D. *J. Am. Chem. Soc.* **1994**, *116*, 443−445. (b) Barron, L. D.; Hecht, L.; Blanch, E. W.; Bell, A. F. *Prog. Biophys. Mol. Biol.* **2000**, *73*, 1−49. (c) Yamamoto, S.; Watarai, H. *Chirality* **2012**, *24*, 97−103.

(25) (a) Dzwolak, W.; Loksztejn, A.; Galinska-Rakoczy, A.; Adachi, R.; Goto, Y.; Rupnicki, L. *J. Am. Chem. Soc.* **2007**, *129*, 7517−7522. (b) Khemtémourian, L.; Doménech, E.; Doux, J. P. F.; Koorengevel, M. C.; Killian, J. A. *J. Am. Chem. Soc.* **2011**, *133*, 15598−15604.

(26) Ryle, A. P.; Sanger, F.; Smith, L. F.; Kitai, R. *Biochem. J.* **1955**, *60*, 541−556.

(27) Heinig, M.; Frishman, D. *Nucleic Acids Res.* **2004**, *32*, W500−W502.

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

(29) Perdew, J. P.; Burke, K.; Wang, Y. *Phys. Rev. B* **1996**, *54*, 16533−16539.

(30) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999−3093.

(31) Bouř, P. Academy of Sciences, Prague: Prague, 2006.

(32) 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. A., Jr.; 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, Inc.: Wallingford, CT, 2009.

(33) Ponder, J. W. 3.8 edn.; Washington University School of Medicine: Saint Louis, 2000.

(34) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049−1074.

(35) (a) Wilson, E. B. *J. Chem. Phys.* **1939**, *7*, 1047−1052. (b) Wilson, E. B.; Decius, J. C.; Cross, P. C. *Molecular vibrations*; Dover: New York, 1980.

(36) Papoušek, D.; Aliev, M. R. *Molecular Vibrational/Rotational Spectra*; Academia: Prague, 1982.

(37) Attri, A. K.; Fernandez, C.; Minton, A. P. *Biophys. Chem.* **2010**, *148*, 28−33.

(38) (a) McColl, I. H.; Blanch, E. W.; Hecht, L.; Barron, L. D. *J. Am. Chem. Soc.* **2004**, *126*, 8181−8188. (b) Blanch, E. W.; Morozova-Roche, L. A.; Cochran, D. A. E.; Doig, A. J.; Hecht, L.; Barron, L. D. *J. Mol. Biol.* **2000**, *301*, 553−563.

(39) (a) Hopmann, K. H.; Ruud, K.; Pecul, M.; Kudelski, A.; Dračínský, M.; Bouř, P. *J. Phys. Chem. B* **2011**, *115*, 4128−4137. (b) Bouř, P.; Keiderling, T. A. *J. Chem. Phys.* **2003**, *119*, 11253−11262. (c) Bouř, P. *J. Chem. Phys.* **2004**, *121*, 7545−7548. (d) Kubelka, J.; Huang, R.; Keiderling, T. A. *J. Phys. Chem. B* **2005**, *109*, 8231−8243.

(40) Bouř, P.; Michalík, D.; Kapitán, J. *J. Chem. Phys.* **2005**, *122*, 144501.

(41) Kapitán, J.; Hecht, L.; Bouř, P. *Phys. Chem. Chem. Phys.* **2008**, *10*, 1003−1008.

(42) Daněček, P.; Kapitán, J.; Baumruk, V.; Bednárová, L.; Kopecký, V. Jr.; Bouř, P. *J. Chem. Phys.* **2007**, *126*, 224513.

(43) Luber, S.; Reiher, M. *ChemPhysChem* **2009**, *10*, 2049−2057.

(44) Jacob, C. R.; Luber, S.; Reiher, M. *J. Phys. Chem. B* **2009**, *113*, 6558−6573.

(45) (a) Hug, W.; Fedorovsky, M. *Theor. Chem. Acc.* **2006**, *119*, 113−131. (b) Jacob, C. R.; Luber, S.; Reiher, M. *Chem.—Eur. J.* **2009**, *15*, 13491−13508.

(46) (a) Luber, S.; Reiher, M. *Chem. Phys.* **2008**, *346*, 212−223. (b) Buděšínský, M.; Šebestík, J.; Bednárová, L.; Baumruk, V.; Šafařík, M.; Bouř, P. *J. Org. Chem.* **2008**, *73*, 1481−1489.

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

(48) Šebek, J.; Kapitán, J.; Šebestík, J.; Baumruk, V.; Bouř, P. *J. Phys. Chem. A* **2009**, *113*, 7760−7768.

(49) (a) Bouř, P. *J. Comput. Chem.* **2001**, *22*, 426−435. (b) Kapitán, J.; Baumruk, V.; Kopecký, V. Jr.; Pohl, R.; Bouř, P. *J. Am. Chem. Soc.* **2006**, *128*, 13451−13462.

(50) Keiderling, T. A.; Silva, R. A. G. D. In *Synthesis of peptides and peptidomimetics*; Goodman, M., Felix, A., Moroder, L., Toniolo, C., Eds.; Georg Thieme Verlag: Stuttgart, 2002; Vol. E22b.

(51) Pecul, M.; Rizzo, A.; Leszczynski, J. *J. Phys. Chem. A* **2002**, *106*, 11008−11016.

(52) (a) Jalkanen, K. J.; Degtyarenko, I. M.; Nieminen, R. M.; Cao, X.; Nafie, L. A. *Theor. Chem. Acc.* **2008**, *119*, 191−210. (b) Jalkanen,

2450

dx.doi.org/10.1021/ac2032436 | *Anal. Chem.* 2012, 84, 2440−2451

K. J.; Nieminen, R. M.; Frimand, K.; Bohr, J.; Bohr, H.; Wade, R. C.; Tajkhorshid, E.; Suhai, S. *Chem. Phys.* **2001**, *265*, 125−151.

(53) (a) Grahnen, J. A.; Amunson, K. E.; Kubelka, J. *J. Phys. Chem. B* **2010**, *114*, 13011−13020. (b) Choi, J. H.; Cho, M. H. *J. Chem. Phys.* **2009**, *130*, 014503.

(54) (a) Hauser, K.; Krejtschi, C.; Huang, R.; Wu, L.; Keiderling, T. A. *J. Am. Chem. Soc.* **2008**, 2984−2992. (b) Hauser, K.; Ridderbusch, O.; Roy, A.; Hellerbach, A.; Huang, R.; Keiderling, T. A. *J. Phys. Chem. B* **2010**, *114*, 11628−11637.

(55) Falconi, M.; Cambria, M. T.; Cambria, A.; Desideri, A. *J. Biomol. Sruct. Dyn.* **2001**, *18*, 761−772.