Vibrational Spectral Simulation for Peptides of Mixed Secondary Structure: Method Comparisons with the Trpzip Model Hairpin

Petr Bouř*,† and Timothy A. Keiderling*,‡

Institute for Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo náměstí 2, 16610 Praha 6, Czech Republic, and Department of Chemistry, University of Illinois at Chicago (m/c 111), 845 W. Taylor St., Chicago, Illinois 60607-7061

Infrared absorption and vibrational circular dichroism (IR and VCD) spectra of model fragments of TrpZip-style β-hairpin structures are simulated using density functional theory (DFT) methods to estimate the influence of fragment size, end effects, conformational irregularities, peptide side chains, and solvent. Different fragmentation schemes, computing the strands and turn segments separately, were tested by varying the sizes of each and their respective overlaps. For suitably overlapping fragments, atomic property tensors were found to be reliably transferable, as tested by their ability to generate simulated spectra in good agreement with results from ab initio DFT computations for the entire peptide. This fragment approach significantly reduces computational times and opens up a wider range of systems that can be studied with a DFT-based approach as compared to previous methods based on uniform repeating sequences. However, vacuum calculations do not adequately represent the frequency dispersion of solvated molecules, and thus, some alternate strategies for solvation correction are explored for improving the simulation accuracy. Unlike for regular periodic secondary structure, the solvent significantly impacts the spectral shapes of hairpins, due to the different degrees of hydration of individual amide groups, which can be exposed to or shielded from water due to external vs internal hydrogen bonding. This is amplified by the shielding of selected amides from the solvent due to bulky side chains. The peptide plus solvent was structurally modeled with molecular dynamics methods, and then an electrostatic field-based parametrization correction was added to the force field and intensity tensors to compensate for the solvent dipolar field. The effect of the shielding and subsequent reordering of modes has a larger impact on VCD than IR band shapes.

Introduction

Infrared (IR) absorption spectra of amide modes have long provided a tool for determining average secondary structure of peptides and proteins, because of the dependence of amide vibrational frequencies and intensities (particularly that of the amide I, mostly C=O stretch) on secondary structure.[1–3] Raman spectra are also interpreted on the basis of various amide mode frequencies but with greater emphasis on the amide III vibration (N–H deformation plus C–N stretch coupled to the C–H wag).[4–7] More recently, vibrational circular dichroism (VCD) and Raman optical activity (ROA) have developed as complementary secondary structure analysis tools, with the interpretation of the former, at least, being more dependent on band shape than frequency,[8,9] paralleling the standard approaches to interpreting protein electronic CD in the ultraviolet region.[10,11] Consequently, many different efforts have appeared to theoretically model these dependences and to predict the spectra from the structure or, in a more applied vein, the structure from the spectra. Early approaches used empirical force fields (FF) formulated obtained by fitting parameters to spectra with various algorithms.[12,13] These FFs proved useful for gaining an understanding of the overall spectral patterns (frequencies), but since the parameters were empirically derived, they were limited in predicting structure and did not provide absolute spectral intensities.

Various theoretical methods based on quantum mechanical (QM) determinations of force constants then developed into hybrid approaches.[14–24] Many of these employed QM to determine FF parameters for just single amides which were transferred to longer peptides, after which the modes were empirically coupled, most often by adding an imposed electrostatic interaction, often termed transition dipole coupling (TDC), and coupled wave functions were obtained by diagonalization of the simplified Hamiltonian.[13,14,24,25] In the past decade, we and others have computed vibrational spectra and simulated intensities at an ab initio QM level for a series of peptides of progressively larger sizes as the abilities of computers and theoretical methods have improved.[21,22,23,26–43] Particularly with the advent of density functional theory (DFT) methods for determining FF and IR intensities (atomic polar tensors, APT) and the development by Stephens and co-workers of magnetic field perturbation (MFP) based methods for determining atomic axial tensors (AAT) with gauge invariant orbitals to simulate VCD at the DFT level, spectra for larger peptides could be directly simulated.[44–48]

Nonetheless, ab initio techniques are still limited to relatively small systems, especially if reasonable basis sets and accurate approximation levels are used. Moreover, computational times for ab initio calculations often scale much more steeply with

---

* To whom correspondence should be addressed. E-mail: bour@uochb.cas.cz (P.B.); tak@uic.edu (T.A.K.).
† Academy of Sciences of the Czech Republic.
‡ University of Illinois.

23688 *J. Phys. Chem. B, Vol. 109, No. 49, 2005*

Bouř and Keiderling

size of the calculation than might correspond to expectations. The most notorious example is the exponential scaling of the full configuration-interaction method which is needed to reproduce the essentially local and additive (linear) correlation energy.[49] Thus, taking into account the relatively local nature of molecular vibrations, we developed a property tensor transfer method to use atomic FF, APT, and AAT values obtained for modest-sized oligopeptide fragments and transfer them to larger peptides of the same (local) secondary structure.[27,28,33] The original approach was dependent on the larger molecule having a repetitive structure so that, after rotating and translating the tensors to orient them properly with respect to the bonds, the coordinates of the smaller one fit the structure of the larger one, and the atomic tensor parameters could be transferred to yield a good approximation of the larger molecule's FF, APT, and AAT parameters. This allowed the smaller molecule model to be progressively propagated along the framework of the larger one to establish the optimal structural match for transfer. The method works exceptionally well for IR and VCD of regular helices[27,29,30,32,34,50,51] and also well for IR of amide I bands of regular, extended β-sheets.[26,36,37] The constraints of the small molecule approximation, i.e., neglect of long-range coupling, have been tested and shown to be minor in regular structures,[26,29,33,36] although for sheets it is difficult to get relevant experimental data for comparison, since formation of an extended sheet brings together multiple strands whose number and alignment are difficult to control.

By contrast, β-hairpins are formed from two antiparallel β-strands linked by a turn. Since β-hairpins are monomeric in solution, they offer better experimental control, but also their properties can be calculated at a higher level, if the structures are known. Thus hairpins provide very nice models for some fundamental issues in protein folding. We have been able to track thermodynamic processes with IR and CD and to isotopically label specific positions in various hairpins for site-specific conformational probes.[31,37,52] However, simulating spectra for such structures poses several additional problems. First, even in the strand sequences, these structures are not constrained to the kinds of uniform main chain torsion (φ, ψ) angles such as seen in regular helices or extended sheets, since real hairpins are twisted and the ends are frayed (the terminal H-bonds are often not formed). Second, the turn is a decidedly different type of conformation, not uniform or repeating. Thus, our original method developed for periodic structures cannot propagate an oligomer model into the turn, and another model of adequate conformation must be used for those residues. Third, the inner directed, cross-strand C=O groups are H-bonded to amides, while those on the outside, and those in the turn, are bound only to solvent. In vacuum simulations, these outer directed residues will not be treated equivalently, which could develop an anomalous frequency pattern. In an extended multistrand sheet this is a simple edge effect, which can be minimized by adding more stands, but in a hairpin, outer directed amides involve more than half the residues and presumably need explicit consideration.

In this work we address the problem of utilizing ab initio determined spectral parameters from different kinds of fragments to create a force field and intensities for a single hairpin. We have taken a targeted approach, focusing on a specific structure, to serve as a model for the general problem of the use of fragments to simulate spectra of larger biopolymeric systems. As a model, we have chosen the TrpZip peptides developed by Cochran and co-workers[53,54] since these are very stable (denaturing well above room temperature) and have established NMR

TABLE 1: Experimental Main Chain Torsion Angles in the Trpzip1 and Trpzip2 Peptides Used in the Simulations

	φ	ψ	ω	χ
			Trpzip1	
Ser (N-end)	−179.9	151.2	169.7	−166.2
Trp	−70.8	116.2	176.4	−63.9, −68.9
Thr	−117.3	154.8	−176.5	59.9, −60.1
Trp	−94.7	129.3	−169.7	174.6, 83.3
Glu	−113.5	46.1	177.9	−152.1, −164.4
Gly	89.1	−56.9	178.7	
Asn	−132.3	−12.7	179.3	−140.5, −43.7
Lys	−134.0	159.5	171.8	46.8, 173.9, −66.9, −61.1
Trp	34.0	158.6	175.4	−64.2, −84.9
Thr	−151.1	146.2	−179.2	−174.4, 65.8
Trp	−91.7	131.1	176.1	−170.4, 78.5
Lys (C-end)	−101.4	113.3	175.3	−70.1, −62.7, −146.2, −72.97
			Trpzip2	
Ser (N-end)	−176.3	129.9	170.4	−172.1
Trp	−77.1	127.1	172.9	−69.2, −75.3
Thr	−128.8	147.5	−176.0	−59.5, −178.2
Trp	−83.0	129.3	175.8	175.5, 86.6
Glu	−141.1	122.7	174.5	−177.1, 179.4, 77.7
Asn	31.2	71.2	178.0	−54.2, 174.4, −177.5
Gly	81.6	−2.3	−176.1	
Lys	−138.8	169.5	168.0	−138.9, 78.5, −174.6, −165.5
Trp	−101.3	135.6	169.8	−67.7, −78.4
Thr	−123.9	135.6	169.8	−67.7, −78.4
Trp	−85.1	100.4	−179.7	−176.9, 75.9
Lys (C-end)	−99.8	134.8	180.0	−61.9, −63.2, −72.6, 165.1

structures. Furthermore, we and others have obtained extensive IR and even some VCD data for them (Setnička et al., unpublished results).[55,56] In addition, the stabilization brought to these hairpins by the hydrophobic interaction of the tryptophan side chains poses another variation on the spectral simulation, that of the impact of side chain shielding on solvation of the amides. Thus, we have tested various models of solvation to account for the Trp influence, indirectly by desolvation, on the backbone amide spectrum.

The initial models used for the antiparallel β-hairpin structures were based on the solution-phase NMR determined structures of the Tryptophan zippers[53]

Ser-Trp-Thr-Trp-Glu-Gly-Asn-Lys-Trp-Thr-Trp-Lys (SWT-WEGNKWTWK; Trpzip1) and

Ser-Trp-Thr-Trp-Glu-Asn-Gly-Lys-Trp-Thr-Trp-Lys (SWT-WENGKWTWK; Trpzip2),

which differ only in the sequence of the two residues forming the tight β-turn (underlined). Both peptides form secondary structures having two antiparallel hydrogen-bonded strands, while the packing of the Trp residues somewhat differ between them, which may affect the solvent structure. Trpzip1 adopts thus a well-defined β-hairpin conformation with a type II′ β-turn, and the cross-strand tryptophan rings pack against one another, with less contact between adjacent tryptophan pairs. Trpzip2 has a very similar structure, primarily deviating from Trpzip1 by having a type I′ β-turn at residues 6 and 7. An overview of the experimentally determined torsion angles can be found in Table 1. Because results obtained with the two sequences were quite similar, we here analyze in detail only the results we have obtained for Trpzip1.

Methods

Ab Initio Models. To study fragmentation effects, we simulated IR and VCD spectra for the target hairpin and for various component fragments starting from the published NMR geometries[53] (Protein Databank Nos. 1LE0 and 1LE1 for Trpzip1 and Trpzip2, respectively) as described in Table 1. The geometry was recreated in the Tinker program[57] (to assign an
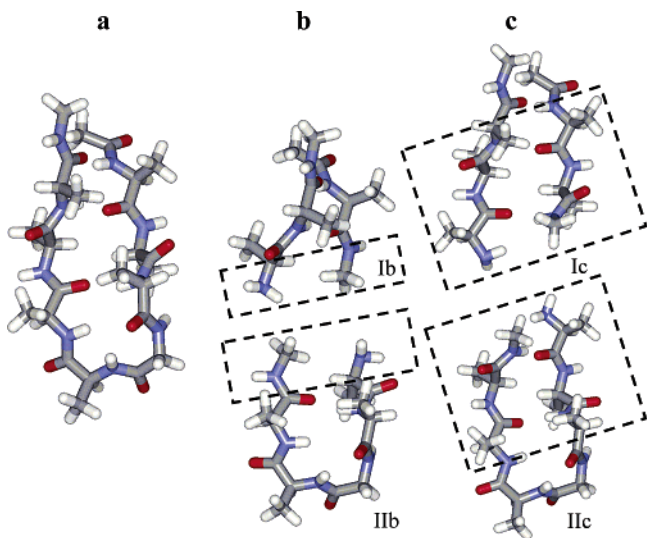
Figure 1. Structures used for the fragmentation test: the 9-amide geometry (**a**) and its decomposition into two smaller (**b**) and bigger fragments (**c**) for the strand (**I**) and turn (**II**) parts of the hairpin. Overlapping parts of the fragments are approximately indicated by the dashed lines.

atomic force field). To remove the worst inaccuracies of that builder, the structure was allowed to relax slightly (50 minimization steps with no constraints, using the Amber force field).[58] This corrects bond distances, angles, and atomic overlaps; the torsions changed little so that the strand conformations did not detectably shift. Then, using the Gaussian programs[59] and our own routines, the structure was optimized at the DFT level using our method of partial optimization in normal coordinates,[60,61] which held modes with $\omega \in (-300, 300 \text{ cm}^{-1})$ fixed. This results in an effective constraint of the (φ, ψ) angles to their starting values but typically provides better convergence behavior than a conventional constrained geometry minimization. Finally, the DFT vibrational spectra (IR and VCD frequencies and intensities) were calculated at the BPW91/6-31G** level. The BPW91 functional[62,63] was used as the primary method for FF determination because it provides good vibrational frequencies at a reasonable computational cost.[64] Hybrid functionals, which are often chosen for vibrational spectral studies, do not improve the amide I and II frequencies at this level yet lead to much longer computational times, limiting the size of peptide we can consider at the ab initio DFT level. Due to the large molecular structures considered for this comparative fragmentation test, basis sets with diffuse functions, which better represent hydrogen bonding[26,65] (Kim and Keiderling, unpublished results), were not employed.

For fragmentation tests, spectral simulations were made using FF, APT, and AAT parameters obtained for the fragments at the DFT level and then directly transferred to the full β-hairpins (onto equivalent geometries in the larger peptide via tensor rotations) by a modification of our Cartesian tensor transfer method,[28] which might be termed a "gluing" or "splicing" together of overlapping fragments. Transferred parameters from the ab initio results for the smaller component fragments described above were combined to yield the FF for the longer peptides of interest (see examples in Figure 1). Note that since all the atoms in their target geometrical environments are explicitly calculated in this approach, there is no propagation (translation and rotation) of a small fragment along the sequence of the large oligopeptide, as would have been characteristic of our original approach. It might be noted that Cho and co-workers have independently developed a somewhat hybrid approach

whereby ab initio calculated spectral parameters for single residues are transferred onto peptides of various structures.[23,24] This is a distinctly different approximation from our transfer method using model oligopeptides, in that we compute all the interactions and couplings at an ab initio (DFT) level for the correct local structure of interest in a peptidic environment before doing the transfer to the extended (large) structure.

Spectra were simulated for both the N—H and N-deuterated peptides, the latter of which correspond to the way they are normally measured in aqueous solution, due to their being dissolved in and exchanged with D_2O.[31,37] Experimentally, this removes the coupling and spectral overlap of the amide I mode (1660–1630 cm^{-1}, depending on structure) with water vibrations (H—O—H bending mode) at ∼1645 cm^{-1}, which moves to ∼1195 cm^{-1} (for D_2O).[23,66] Simultaneous simulations of N—H containing peptides allow us to investigate potential effects on the amide II, accessible by measurement in (ordinary) H_2O. Deuteration also has a huge effect on the N—H, N—D stretching (amide A) modes; however, these are normally only measured in nonaqueous solvents.[3]

Explicit hydration of the Trpzip1 hairpin structure was modeled using the Tinker molecular dynamics (MD) software.[57] Starting from the experimental Trpzip1 structure[53] we built a model peptide (Ac-SWTWEGNKWTWK-NH$_2$) and its all-alanine analogue where all residues except the Asn and Gly were replaced by Ala. Each of the peptides was placed in a cubic water box (31.03 Å on a side) constructed using a standard routine in Tinker. The Amber[58] peptide force field was combined with a separate water potential.[67] Using periodic boundary conditions and an NpT ensemble ($p = 1$ atm, $T = 298$ K, time step of 1 fs), we performed a full MD run on the peptide plus water as well as a restricted MD run where the peptide atoms were constrained to the starting structure and only the solvent molecules were allowed to move.

For these model peptides, the electrostatic potentials for each peptide atom were computed by summing the contributions of all the water partial charges (O = −0.84 e, H = 0.42 e) in instantaneous configurations, utilizing programs written in house. These were then averaged over an ensemble of 1000 MD configurations, each separated by 100 fs steps. Applying the semiempirical correction described in detail elsewhere,[68–70] we could use these potentials to estimate the influence of the explicit solvent on spectral frequencies and intensities and could additionally obtain an estimate of band broadening due to fluctuating solvent interactions. The correction is based directly on the solvent electrostatic potentials φ_j at each amide group atom $j = 1-N$. For example, diagonal force constants for a peptide in the solvent, Λ_I, were calculated from the vacuum peptide FF values, Λ_{0I}, as $\Lambda_I = \Lambda_{0I} + \Sigma_{j=1,N} b_{Ij} \varphi_j$, where the empirical coefficients $\{b_{Ij}, \Sigma_{j=1,N} b_{Ij} = 0\}$ were obtained from a fit of ab initio computations performed on N-methylacetamide–water clusters.[69] These MD-based solvent-corrected results are presented for a D_2O solution (and N—D-exchanged amides) to eliminate the vibrational interference of H_2O with the amide I.

Results

Hairpin Fragmentation. A number of different, somewhat independent computational tests were undertaken to determine optimal ways of addressing the problem of spectral simulation for a larger heterostructured peptide. The fragmentation test, perhaps the most important outcome, is shown first since it is basic to the other tests. We wanted to know how the method of transferring atomic force field and spectral intensity tensors[28] performs for the heterogeneous hairpin geometry. For this test
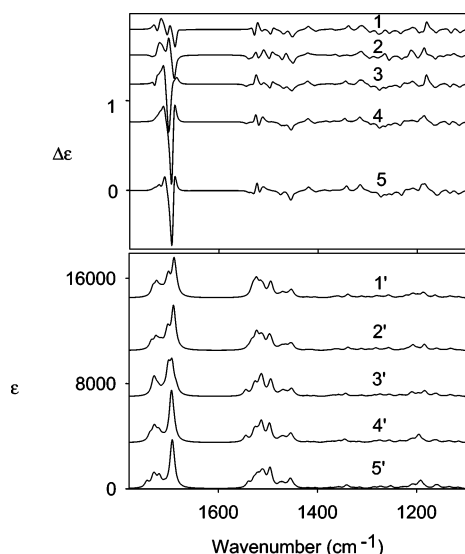
Figure 2. VCD (top) and absorption (bottom) spectra of the (protonated, N−H) model 9-amide hairpin (**a** in Figure 1) calculated by transfer from (1, 1′) the smallest fragments (**Ib** and **IIb** in Figure 1), (2, 2′) the same fragments but with the PM3 force field correction, (3, 3′) those calculated by the transfer from intermediate, 1-amide overlap fragments (**IIb, Ic**), and (4, 4′) transfer from the largest (2-amide overlap) fragments (**Ic, IIc** in Figure 1) all compared with (5, 5′) the full ab initio (BPW91/6-31G**) hairpin **a** computation.

we constructed a target hairpin that consisted of an Ala-based sequence forming a 9-amide loop segment (all residues were replaced by Ala except for the Gly residue in the turn, which is structure **a** in Figure 1) whose conformation was based on the Trpzip1 NMR structural parameters. [This truncated sequence eliminates the frayed terminal residues and leaves a resultant peptide whose 89 atoms (336 electrons, typically 885 basis functions, 6-31G**) are compatible with reasonable computation times using standard basis sets and our current capabilities. By way of example, the analytical second derivative calculation at the BPW91/6-31G** level took about 200 h on one Intel Xeon 2GHz processor and required 27 GB of disk space.] After relaxation of the higher frequency modes with a restrained normal mode optimization,[60,61] fully DFT-based IR and VCD spectra were simulated for this structure, and the results are illustrated in Figure 2, traces 5 and 5′, respectively.

These full hairpin results, obtained solely at the DFT level, were used as a standard for judging reliability of several simulations using transferred parameters obtained from different-sized fragments that develop various overlap patterns. Initially, smaller, minimal-overlap fragments of 49 and 54 atoms (structures **Ib** and **IIb** in Figure 1) were constructed to correspond to the strand termini and 5-amide loop parts, respectively, of the structure in Figure 1a. These provide minimal representations and thus the most efficient DFT computational means for two fragments to contain all the residues in **a**.

For the second transfer calculation, the hairpin **a** was divided into larger fragments of 69 and 74 atoms, as illustrated by **Ic** and **IIc** in Figure 1, which have greater overlap but consequently require longer computation times. Yet these fragment calculations took notably less resources than the full peptide in **a**; for example, the BPW91/6-31G** second-derivative computation for the 74 atom fragment required 73 h and 2.8 GB of hard disk space. Note that the two central hydrogen-bonded amide pairs in **a** were included in both fragments in **c**. Such overlap permits one to eliminate parameters arising from atoms forced to lie at nonrealistic termini in the fragments and retain only

those corresponding to a sequence of coupled amides. Of course, even for these substantially larger calculations, long-range couplings are not computed, but these are likely to be of little importance as long as cross-strand interactions are included.[26,36]

The fragment structures were optimized with constrained normal coordinates (as noted above and in detail separately[60,61]) at the DFT level, and then the new harmonic force field and intensity tensors were calculated at the same DFT level as used for the other fragments. For comparison to the fully ab initio results, the fragment tensors were finally transferred onto structure **a** and the spectra of **a** was calculated by diagonalizing its resultant FF using our routine methods.[28] In the transfer from fragments **c**, all residues in a continuing sequence (internal) are finally computed with realistic parameters (diagonal and near neighbor coupling), since the fragments used have each residue represented with amides before and after them in the separate sequences, and only terminal residues are computed with amides next to open end groups. However, with the minimal fragmentation from part b, due to missing sequential interactions, end effects are likely to influence amides that should be associated with "internal" residues which are also part of the junction between fragments.

In Figure 2, the resulting simulated spectra obtained from different schemes for the transfer of parameters from the two sets of fragments above are compared to the fully DFT-generated VCD and IR results for part a in the bottom traces 5 and 5′, respectively. The top traces 1 and 1′ show the spectra obtained for **a** with transfer from the minimal sized fragments of **b**, which have sensible agreement for IR frequencies but a distorted intensity pattern as compared to the full DFT computation (structure **a**, trace 5, 5′). The amide I VCD in 1′ is significantly different in intensity and sign pattern from 5′ suggesting that such a minimal transfer leads to poor representation of the interactions between amides that are needed for the VCD. The shape of the amide II IR computed at ~1490−1540 cm^{-1} is different as well, and only an approximate VCD shape agreement is obtained using the minimal fragmentation. As an intermediate test, the transfer with the smaller fragments **b** was repeated, but PM3 force constants were computed and used to correct the otherwise neglected long-range interactions (those components of the FF not included in the transfer). The PM3 correction (Figure 2, traces 2, 2′) somewhat improves the spectral shapes over those obtained with the smaller fragments (traces 1, 1′), particularly with respect to the intensity of the amide I IR and VCD (~1650 cm^{-1}), but does not correct for the fundamental end-effect errors introduced by transfer of parameters for a terminal amide onto residues in a sequence.

By contrast, the transfer from the fragments having intermediate overlap (e.g. Figure 1, **IIb** with **Ic**, whose spectra are in Figure 2, traces 3, 3′) provides IR and VCD much closer to the ab initio result in terms of intensity and sign patterns. Reasonable IR dispersion and VCD sign pattern are obtained for the amide I, but details of the mode splitting differ from the full-DFT results in traces 5, 5′. Similarly the amide II bands in the IR are more resolved in 3′ than in 5′ and some of the weak components in the 1450−1500 cm^{-1} region differ in frequency between traces 3′ and 5′.

Going further, when both the larger fragments, **c**, allowing maximum overlap (i.e. **Ic, IIc**) are used, both the VCD and absorption profiles (traces 4, 4′) improve, giving almost perfect agreement with the full DFT result (traces 5, 5′). Indeed, as can be seen by overlaying these spectra with a heightened intensity scale in Figure 3, the transferred (black) and fully DFT spectra (red) are almost identical (with two very minor band
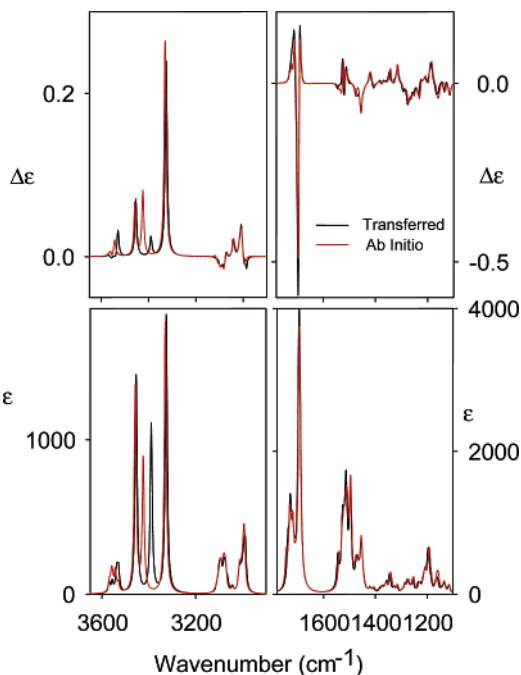
Vibrational Spectral Simulation for Peptides

J. Phys. Chem. B, Vol. 109, No. 49, 2005 **23691**

Figure 3. Overlap of the absorption (ϵ) and VCD ($\Delta\epsilon$) spectra of the fragment **a** (in Figure 1, protonated) obtained by the transfer method (**Ic, IIc**, red) with DFT derived intensities (black).

shape fluctuations in the amide I and II VCD) through the mid-IR region (1100−1800 cm^{-1}). However, the approximations of the transfer methods do result in some residual deviations that cannot be excluded, as documented in the left-hand part of Figure 3 where the C−H and N−H stretching regions of the spectra are plotted. Here, the N−H vibration at 3424 cm^{-1} in the fully DFT calculation was obtained at a significantly lower frequency, i.e., 3390 cm^{-1}. This heightened sensitivity of the NH group to fine geometry and force field changes may be due to alteration of the H-bond geometry close to a terminal group that could occur in the normal mode minimization. [Note: these vacuum simulated modes are not measurable in real systems in solution.] Since, however, the character of the vibration (normal mode displacement pattern) was conserved and because of the success in replicating other regions of the spectra, for the mid-IR region, we find the overall accuracy of the transfer method to be very good when fragments have significant overlap.

Amide Group Hydration and Solvent Electrostatic Field. The importance for amide group frequencies of external electrostatic fields due to dipolar character of the solvent molecules has been recognized in previous works.[22,23,70,71] In fact, most of the solvent shift of the important amide modes can be reproduced by such a field correction.[69] Similar studies on water vibrational spectra suggest that other polar chromophores or possibly those subject to solvent hydrogen bonding behave like this.[72,73] Thus, for the Trpzip1 NMR structure (Ac-SWTWEGNKWTWK-NH$_2$) and its simplified alanine analogue (Ac-A$_5$GNA$_5$-NH$_2$) in a cube of water (e.g. Figure 4), we computed an average field at the solute due to the solvent water molecules for an ensemble of configurations generated in an MD simulation of the system. The field correction for the amide I mode was computed on the basis of the difference in electrostatic potential at the amide carbonyl oxygen and carbon atoms. Although the alanine-based peptide was not stable in the Trpzip1 starting conformation and started to unfold for a longer MD simulation (>1 ns), its conformation did not deviate significantly from the initial Trpzip1 backbone within the averaging time used here (0.1 ns), even for the unconstrained
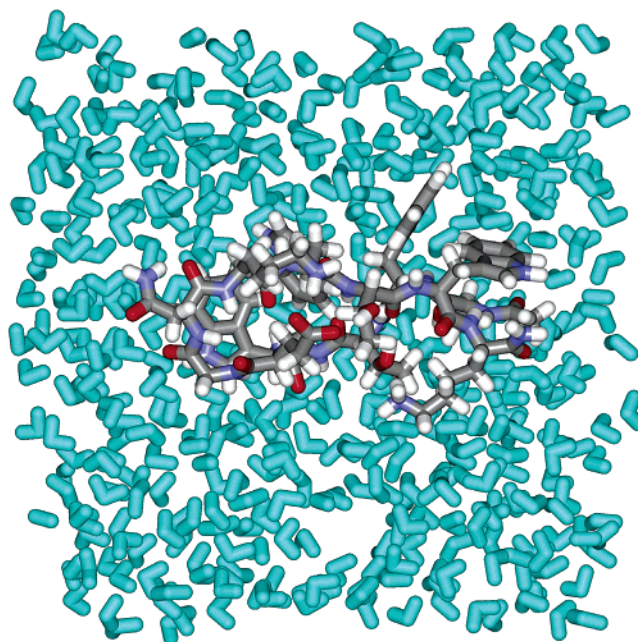

Figure 4. Trpzip1 peptide model in the box of water molecules. (Waters in front of the peptide are not visible.)

MD. The peptide with the full complement of realistic Trpzip1 side chains did not unfold even for longer times, which is consistent with the experimentally observed stability of the Trpzip1 molecule.[55,56]

The average differences between the potentials at the oxygen and carbon atoms of the two peptides are plotted for each carbonyl group in Figure 5. The restricted and unrestricted simulations provide qualitatively similar results, which can be explained by the relative stability of the peptide structure and the fast motion of water molecules as compared to the peptide chain movements. However, the full peptide (Figure 5a, upper part) has distinct contrasts in the pattern of potential differences from the alanine-based model (Figure 5b, lower part). In the Ala peptide, each residue has access to solvent, modulated primarily by cross-strand hydrogen bond formation. Those H-bonded amide C=O groups, on alternate positions (nos. 2, 4, 9, 11), have less solvent-induced electrostatic field ($\Delta\phi$). The turn C=Os (the so-called $i + 1$, $i + 2$ residues, nos. 7, 8) are exposed to solvent, consistent with the large $\Delta\phi$ values indicated for those positions (Figure 5). Somewhat surprisingly, the preturn C=O (ith residue, no. 6) is also affected, even though it should be cross-strand H-bonded to form the turn but is clearly exposed meaning the turn is partially distorted. In the full peptide with side chains, individual amide (C=O) groups are hydrated differently, and the regularity of the all-Ala pattern breaks down. The N-terminal part still shows the alteration noted above, but the turn residues are less regular, though they still seem to be exposed to solvent. However, residues 9−12 are shielded from solvent. This is due to the large Trp groups being clustered on one side of the hairpin and excluding solvent. The shielding causes the amides following residues 10 and 12 to have no apparent solvent field, $\Delta\phi$. The inversion of $\Delta\phi$ for the amide between residues number 2 and 3 is another example of the side chain impact.

An alternative representation of the different degrees by which the various amide groups can be accessed by the solvent can be seen in Figure 6, where radial distribution functions between water hydrogen and carbonyl oxygen atoms are plotted for the constrained simulations. Clearly, for the full peptide, the C=O residues 2, 4, 9, and 11 which make the cross-strand H-bonds
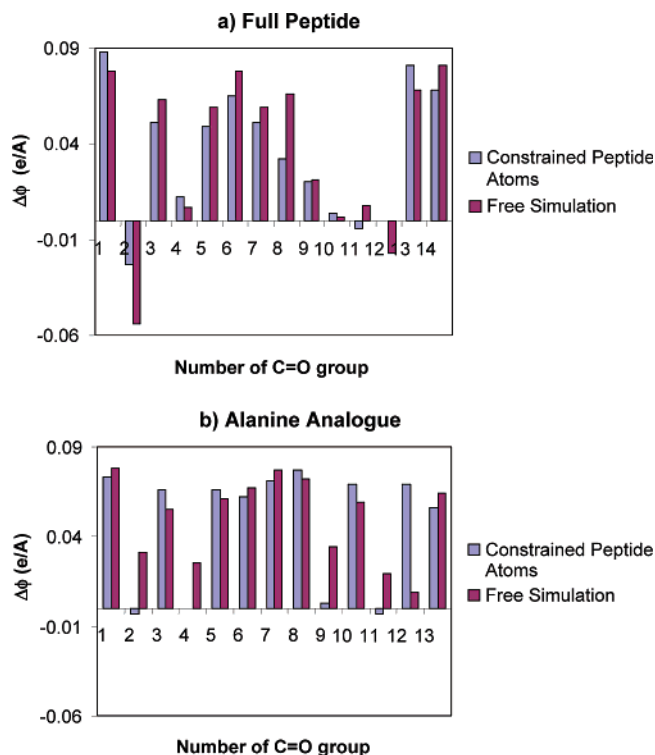
Figure 5. Solvent electrostatic field differences at the carbonyl carbon and oxygen atoms, obtained as the molecular dynamics average for constrained and unconstrained simulations, for the full system (Ac-SWTWEGNKWTWK-NH$_2$, at the top) and for peptide where all side chains were replaced by CH$_3$ groups (at the bottom). The amide groups are numbered consecutively from the C-terminus; the Asn residue on the turn is given no. 14.
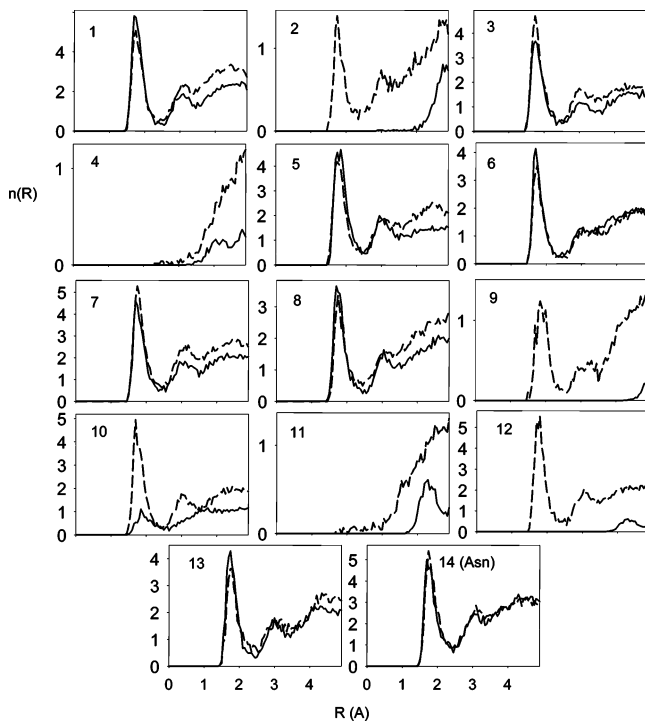

Figure 6. Radial distribution functions of water hydrogen atoms, as obtained by the MD simulation for the 14 carbonyl oxygens of the Trpzip1 peptide (solid line) and the alanine-like analogue (dashed line). The numbering of the carbonyls corresponds to that in Figure 5.

in the hairpin do not have a high probability for solvation by water. Also residues 10 and 12 have much reduced hydration because of the bulky tryptophan residues. For the latter two
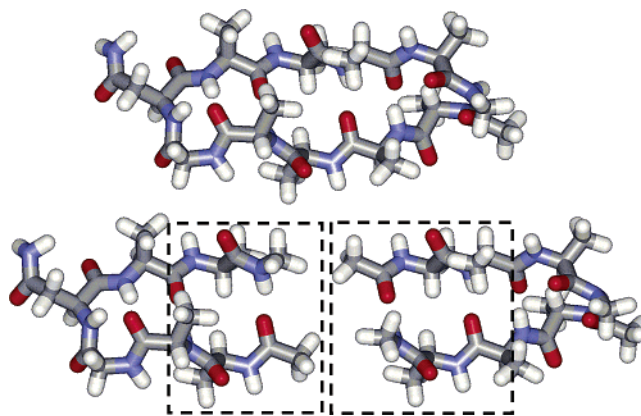

Figure 7. Ala-like fragment (Ac-A$_4$GNA$_4$NHMe) mimicking Trpzip1 (above, 113 atoms) and the two fragments from which ab initio parameters were transferred (below: turn segment, 73 atoms; β-strands, 84 atoms). The overlapped parts are emphasized by the dashed rectangles.

residues, more extensive hydration is seen for the alanine-based peptide (dashed lines) where the side-chain shielding is reduced. However, for residue 2, hydration is only partially increased (compare the vertical scales). Within the same peptide the electrostatic patterns given in Figure 5 appear more sensitive to variations in hydration than do the radial densities. This observation is important for the spectral modeling, since the fields, $\Delta\phi$, directly correlate with the frequency shifts, particularly for the highly sensitive amide I mode.

Influence of Solvation on Vibrational Spectra. To put this dynamical shielding and structural insight together with spectral computation and thereby simulate the influence of the solvent on the spectra, a simplified peptide (Ac-A$_4$GNA$_4$NHMe, 12-amides total, counting the Asn side chain) was derived from the experimental Trpzip1 geometry by use of the Tinker editor just as done for the analogous hairpin fragment **a** in Figure 1. However, since this molecule was too big for direct DFT computation, we constructed its FF and intensity parameters by property transfer from the smaller, yet strongly overlapping fragments displayed in Figure 7. Here the strand fragment I, (Ac-Ala$_3$-NHMe)$_2$, has 84 atoms (820 basis functions) and the turn segment, Ac-AAGNAA-NHMe, has 73 atoms (735 basis functions), both posing substantial computational requirements. The fragment FF and the intensity tensors were obtained at the BPW91/6-31G** level. On the basis of the tests discussed above and results in Figures 2 and 3, we assume that this calculation, which has a two amide overlap between fragments, is virtually equivalent to a direct DFT result.

The amide I′ absorption and VCD spectra of the N-deuterated 12-mer obtained from the vacuum calculation are shown in Figure 8, traces a, a′. Though shifted to high wavenumber, the absorption profile resembles that expected for a β-sheet conformation,[26,36] namely having a sharp, intense peak computed at 1700 cm^{-1} and lower intensity higher frequency bands at ∼1720−1760 cm^{-1}. Obviously, comparison to regular sheets is not entirely relevant because only few amide groups in this peptide actually have a β-sheet conformation, but those that are there influence the intensity distribution, presumably by coupling to the other amides in a manner to bring out the characteristic β-sheet amide I splitting pattern. The conceptually simplest method of solvent correction is to redo the DFT calculations with addition of the COSMO continuum solvent model. This approach yields a more compact and intense amide I′ band profile (Figure 8b,b′) which is shifted down by ∼30−50 cm^{-1} from the vacuum results and better resembles the observed

Vibrational Spectral Simulation for Peptides

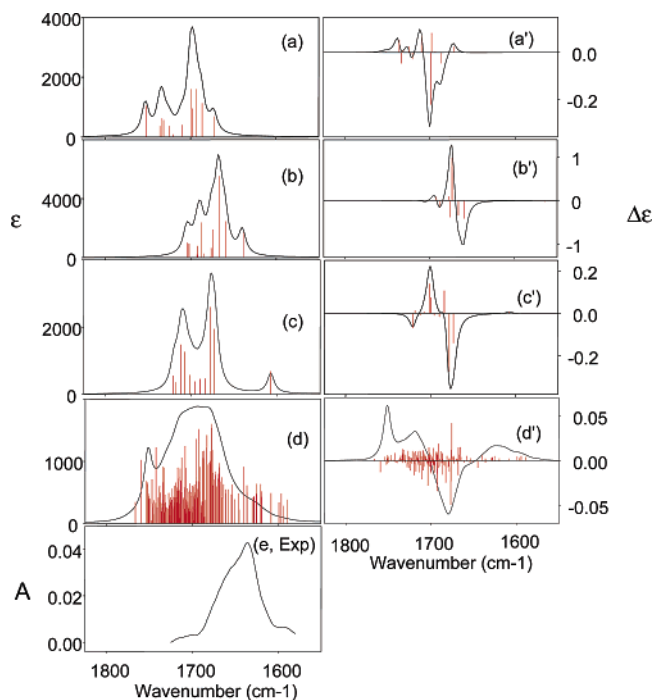*J. Phys. Chem. B, Vol. 109, No. 49, 2005* **23693**

Figure 8. Trpzip1 absorption (left) and VCD (right) spectra of simulated all-Ala-like peptide from Figure 7 at the BPW91/6-31G** level (a, a′) in a vacuum (b, b′) with the COSMO solvent model, (c, c′) with the correction to the average MD electrostatic field, and (d, d′) by direct MD averaging using the empirical correction obtained for *N*-methylacetamide at the BPW91/6-31G** level. The experimental spectrum (e) could be obtained only for absorption.

amide I′ in TrpZip1 (Figure 8e). The normal mode ordering is also changed, which has obviously a larger impact on the predicted VCD band shape than on the absorbance. Indeed, the changes in normal mode ordering could be verified by the visualization of the dynamic displacements (not shown) and can be clearly attributed to the interaction with the solvent, which is different for each amide group in the peptide. Because of the delocalization of the vibrations and the mode coupling, there is not a simple explanation of the calculated changes in the VCD pattern. The observed change is in contrast to results for the helical peptides studied previously, where the solvent had only minor effect on the overall band shape but obviously larger effects on the underlying mode distribution.[29,33,34,50]

In a different approach to modeling the solvent effect, the water electrostatic potential, averaged from our MD calculations of the peptide in the solvent cube (Figure 4), was used to add an empirical correction to the ab initio force field of the simplified alanine-based 12-amide peptide (Figure 7).[23,68–70] The fitting coefficients previously obtained for the linear dependence of the force field and dipolar derivatives on the electrostatic field obtained for clusters of *N*-methylacetamide (NMA) and water[69] obtained at the BPW91/6-31G** level were used to create corrections to the FF, APT, and AAT tensors for 100 MD configurations. The configurations were separated by 500 fs steps so that they could be considered independent, and tensors (FF and APT only) for each of them were corrected for the electrostatic potential on the corresponding atoms of the full peptide. Thus shielding, determined from the ensemble of MD structures of the full peptide plus solvent, was taken into account even though the side chains are not explicitly represented in the DFT calculations.

In Figure 8c,c′ the effect of the average electric field on the IR and spectra can be seen. Clearly the IR intensity has a large shift coalescing into two major features. The low-energy peak

at 1607 cm⁻¹ arises from the C=O on the last amide in the strand next to the turn (*i*th residue) which has a bifurcated H-bond with contributions from the turn amide N−H (*i* + 2, Asn N−H) and the cross-strand amide N−H (*i* + 3 residue). It is also the lowest energy mode in the vacuum- and COSMO-based calculations, but in those calculations it is not so drastically split from the other amide I modes. Such a large shift is consistent with the solvent exposure computed in terms of $\Delta\phi$ in Figure 5. The VCD on the other hand has strong similarity to the COSMO result in Figure 8b, meaning the coupling between amides did not change drastically by adding the field correction. The average amide I frequencies obtained with this parametrized method are not improved as compared to the COSMO correction, in contrast to our experience form computing peptides with explicit H-bound waters using DFT methods.[26,34] Frequency is not really the issue if coupling is represented well, since this method could develop better overall frequencies by merely reparametrizing. However, this method does provide a means of estimating the bandwidth arising from fluctuation of the water around the peptide. By summation over many MD configurations in the ensemble, inhomogeneous line widths arising from water structure fluctuations could be simulated as shown in Figure 8d,d′ (note: only a fraction of the transitions calculated is plotted for Figure 8d,d′ because of plotting software limits). For each configuration a spectrum was obtained by applying the correction to the vacuum force field and APTs of the same simplified peptide, but again the solvent electrostatic field for each amide group was obtained for the fully hydrated trpzip model, in this case including side chains. As can be seen (Figure 8d,d′), after such a correction the amide I is now greatly broadened, even more than found experimentally. The VCD has the same intrinsic sign pattern, but a new positive band is seen to low wavenumber and the high wavenumber intensity distribution is shifted significantly, reflecting the increased IR dispersion, both of which pose problems for this model.

Finally, we can estimate the influence of the peptide side chains on the absorption and VCD spectra by comparing the MD-based correction for the full, side-chain shielded Trpzip1 MD simulation and those for an analogous peptide where all the side chains (except Asn and Gly) were replaced by methyls, i.e., converted to alanine residues. It should be clear that the basic FF and intensity parameters are the same for both Trpzip1 models and the same as for the results above (Figure 8), having been transferred from parameters determined from single DFT computations on the all-Ala hairpin fragments (Figure 7). The absorption and VCD spectra simulated without (traces a, a′) and with (traces b, b′) the shielding of the trpzip side chains are plotted in Figure 9. For this comparison, the results are based on the fit to NMA computed at the Becke3LYP/6-311++G** level, which provides somewhat more realistic bandwidth for the amide I (compare trace b Figure 9 and trace d in Figure 8). It is important to remember that the fitting used to correct FF and intensity parameters is based on DFT computations of solvated NMA, neither on actual experimental results nor on results for an oligopeptide of relevance to this molecule. Thus, even though corrected, the amide I frequencies will still be higher than desired for agreement with experiment.[69] As always with such simulations, the relative changes of various modes in the band are more important than are absolute frequencies for structure correlation. Apparently, while the IR spectra simulated with and without the shielding are quite similar except for the slight amide I average frequency shift and intensity variations, the VCD pattern does change quite significantly,
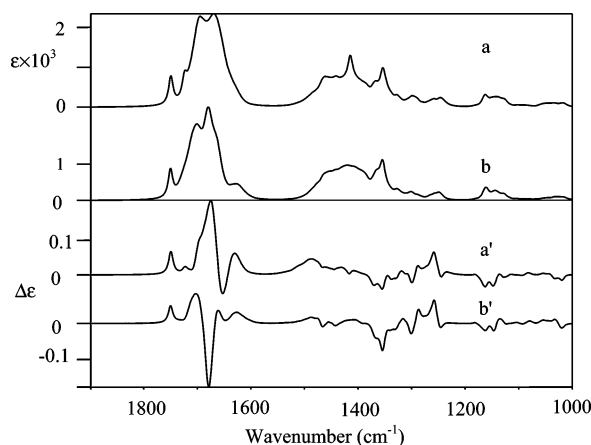
Figure 9. Simulated absorption (top) and VCD (bottom) spectra of the model 12-amide hairpin using the MD solvent correction. The correction was either applied to the Ala peptide directly (a, a′) or by use of the shielding corresponding to the Trpzip1 side chains (b, b′).

developing a qualitative difference, with the amide I signal being predominantly positive (trace a′) as opposed to dominated by the negative lobe (trace b′). Clearly, this must be a result of reordering of local modes in the shielded environment.

Discussion

Fragmentation Approach. The resource demands of carrying out ab initio computations on large biomolecules are often seen as the limiting factor for modeling realistic systems, but the situation is more complex. We have shown that computational times can be reduced by taking advantage of the local nature of the vibrational excitations, for example, by the fragmentation approach tested here for the trpzip model hairpin. The key to developing computationally accurate FF and intensity parameters for larger peptides is to compute parameters for each residue in a realistic local environment. That means the residues in an extended sequence must have FF parameters computed for an environment including near neighbors and any cross-strand residues to which they are strongly hydrogen-bonded.[26,36] Thus to represent the hairpin, we included overlap between the fragments so that each residue in the continuing sequence was not dependent on parameters computed for a terminal residue in the fragment. The best computations were developed with double amide overlap (i.e. **Ic** and **IIc** in Figure 1), but satisfactory representations of the major spectral features were obtained with single amide overlap (**Ic** and **IIb** or **IIc** and **Ib**). This means that the transferred parameters on sequential residues include diagonal and at least one amide near-neighbor off-diagonal interaction terms. The results with minimal, merely terminal blocking group overlap (**Ib** and **IIb**) were decidedly inferior, which directly suggests a weakness in attempts to use ab initio computed parameters from very small amides, for example, NMA or dipeptides, to develop FF and intensity parameters for large molecules. Such an approach, while in principle quite attractive, is bound to both miss significant interaction terms and insert unrealistic end effects into the FF. On the other hand, by extension of the transfer results above, use of a triamide model for each structural type (not tested here) could develop a correction for at least the nearest-neighbor coupling.[27,29,74] Consequently the accuracy of computed band shapes and particularly the IR and VCD intensity distributions will suffer to the extent that those interactions impact the spectra. In general, interaction terms are vital for providing the amide I modes (and others by extension) their conformational sensitivity.

Thus, if one seeks to relate computed spectra to structure, then a minimal representation will not yield accurate results.

Nonetheless, several reports exist of parametrized models that yield spectral representations in good agreement with experimental values.[22,24,75−80] Such models have value for providing rapid, simple modeling of the changes in a system under structural or other perturbations. By contrast, our results are not parametrized; thus, for us to improve agreement with experimental spectra in terms of the average frequency (diagonal FF) and dispersion of the modes (overall interactions), the basic properties of our model must be enhanced. When tested, those enhancements directly point to the types of physical interactions that are important in getting accurate simulations and yield insight to the method's sensitivity to perturbation and structure. Parametrized models can normally be adjusted to yield the quality of agreement desired. However, the more fitting and parameters used, the less rigorous and physically based will be the interpretation. Thus, both approaches have a role in spectra−structure analyses but yield different insights.

Static Solvent Corrections. As emphasized in the Introduction, the success of previous models of vibrational spectra for regular secondary structures was enhanced by the peptide periodicity. In the final result, the systematic error in amide frequencies that is normal for vacuum DFT computations could be represented as a uniform shift of the simulated amide I band. As we have shown, this error is due primarily to ignoring the solvent, but in many cases (especially helices) the coupling is not affected significantly by adding solvent molecules to the computation.[26,29,34] Spectral shapes, which are most relevant to sensing secondary structure, remained conserved in those cases, since they are mostly determined by interresidue coupling. For irregular structures, such as the β-hairpins, the situation is clearly different at the outset, because individual amide groups are differentially exposed to the solvent. Since these are the structurally diagnostic chromophores, one should ideally understand the effects of the solvent field on the peptide before using vibrational spectral results for structural analyses. This would be critical in extending the methods to proteins where shielding of various amides by other, sequentially remote factors in the folded structure is common. Such understanding will determine how important is correction for solvent effects and the best method of achieving it.

Previously we have shown that DFT-based simulations of IR and VCD spectra had a distinct frequency error that could only be substantially corrected by inclusion of diffuse basis functions in the DFT computation or by explicit consideration of the water solvent molecules, at least representing the inner, directly hydrogen-bonded shell.[29,34,36,70] No correction was perfect, but these two enhancements of the FF made a large difference to the amide I frequencies, whereas added basis set size and added hydration shells beyond this resulted in only smaller shifts and did not seem to change the character of the normal modes. However, even these "first-level" corrections did not fundamentally alter the coupling or affect the relative mode ordering in helical peptides where all the residues were in very similar environments.[29,34] Because of that, aside from frequency, the qualitative IR and VCD band shapes (intensity distributions over the exciton band) of helices are quite well represented even with vacuum (no solvent) calculations.

By contrast, for more general structures, as exemplified by the hairpin models studied here, this is less true. In our previous work we have computed amide I spectra for tight turns and for antiparallel strand fragments, both with and without water.[26,31] In the 5-amide ^{D}Pro-Gly-based turn model, addition of an inner

Vibrational Spectral Simulation for Peptides

J. Phys. Chem. B, Vol. 109, No. 49, 2005 **23695**

hydration shell resulted in a substantial shift to lower wavenumber, as expected, and a contraction of the computed frequency dispersion of the amide modes as compared to the vacuum result.[31] This would normally have led to a reordering of modes in the spectral sequence for the turn, if it were not for the X-^{D}Pro tertiary amide linkage yielding an artificially low amide I frequency. By contrast, for the 2 × 3 (2 strands, 3 amides each) β-sheet model, the average frequency shift on hydration was less. While the dispersion did not actually increase, the intensity distribution changed to make it appear so, since very weak intensity high-frequency modes in the vacuum simulation have more intensity in the solvated case.[26] Combining these results to simulate a solvated hairpin structure would suggest that solvation could cause a relative reordering of modes between those originating in the strand and turn parts of the molecule. This same sort of change is seen in Figure 8, where solvent correction seems to contract the dispersion for the entire hairpin, but the VCD shows shape changes that would result from a reordering of local modes centered on turn and strand.

Empirical MD-Based Solvent Correction. Despite these efforts to model the solvent in some sort of average or static representation, we must recognize they are potentially flawed. Since the solvent molecules are weakly bound to the peptide and to each other, at 300 K they are in rapid motion creating a diverse ensemble of local structures. Ideally one would like to account for this dynamical fluctuation. Clearly a full quantum mechanical representation of the effects of such an ensemble is impossible. Thus we, and others, have employed empirical corrections to the spectral parameters to build a representation of the spectra influenced by the dynamic solvent.[22,23,69−73] In this report we used results of MD simulations of the peptide in a water environment to create an effective electric field at the atoms of interest and then, using a parametrization scaled against previous DFT calculations for water−NMA complexes, scaled the FF and intensity parameters and rediagonalized them to generate corrected spectral responses. With these one could then average over the time course of the MD and generate a distribution of spectral responses, which when summed gave a representation of the line shape seen experimentally in solution.

Our approach is based on a treatment which obtains spectral parameters computed at a high ab initio level for fixed conformation peptides and then perturbs them with an electric field that fluctuates due to the dynamic motion of the solvent. Two aspects are left out in this approach. First scaling to DFT results for hydrated NMA, even for bigger clusters, does not obviate any errors in those calculated spectra, nor does it account for the differences in NMA from sequential peptides. In particular the initial models used for scaling were minimal hydration clusters (just three water molecules in various geometries) to develop the method rather than converge of excellent frequency agreement with experiment.[69] Extended hydration shells certainly will impact frequencies[68,70] but will have little impact on coupling. Second, our results considered only one conformation of the hairpin that was maintained over a 1 ns trajectory. Perhaps for more flexible systems (or longer trajectories), models other than the very stable β-hairpins we have chosen for this example test, one will instead have to adapt empirical and classical dynamics means of developing FF parameters for peptides in solution.[21,22] These issues pose a challenge for future studies, that is, properly accounting for the dynamics of the peptide chain itself, a factor neglected in our instantaneous mode approximation. They additionally point to opportunities for improvement of the accuracy of the method.

Despite these limitations, we can see that the current model provides new information about specific peptide−solvent interactions and provides a broader base for interpreting the experimental data than do vacuum-based or static models.

The MD modeling of the extent of hydration presented here is in some contrast to assumptions made in developing our earlier all-DFT simulations of explicit water on α-helices or β-strands where we placed water molecules on all available C= O and N−H groups on the basis of energetic considerations only and did a normal-mode restricted DFT optimization.[26,31,34] The results presented here suggest that a geometry with completely H-bonded residues is an overestimate of the available water density if one accounts for motion at 300 K and entropic effects. For helices, both methods provide useful corrections to the FF resulting in improved IR and VCD band shapes in the amide I and II bands, but the explicit water corrections improve the amide I−II splitting and even the amide I coupling (VCD band shapes), while the empirical MD corrections offer solvent fluctuation based FF variations.[34,69] For sheets, these are the first MD corrected attempts we have tried, but explicit water corrections seem to yield primarily better frequencies (near experimental values for amide I) and an apparent, but small, increase in dispersion.[26,68] Here we see that even the reduced density, sequence-dependent radial distribution functions seen in Figure 6 result in changes to the hairpin spectra as compared to the vacuum results. Since these effects are modified by the side chains, the role of the peptide side chains should gain consideration in future studies, particularly for β-sheet models. Probably the role of the sheet twist on solvent accessibility will also need reinvestigation.[81]

Additionally, the MD correction provides one means of simulating bandwidths and comparing to experimental results. If the MD broadened spectra are compared to the experimental spectra (Figure 8, traces d and e, respectively), we see that significantly broader widths are obtained by the simulation with the MD-based correction, but we must note that these are partially products of the approximation and basis sets used.[69] Indeed, if Figure 8d is compared with Figure 9b, it is clear that the overall amide I bandwidth with the BPW91/6-31G**-based correction is larger than with the Becke3LYP/6-311++G**-based correction and that the shape of the latter is in closer agreement with the experimental intensity distribution, notably having a peak intensity at the lower frequency range of the amide I dispersion; however, the anomalous high-frequency component ($\sim$1750−60 cm^{-1}) from the vacuum result (Figure 8a) persists. [Note these two corrections are based on different parameter sets for NMA plus three waters of solvation that were calculated with these different basis sets but here are fit to the same MD-derived potentials for the hairpin peptide.] This narrowing with the 6-31++G** parameter set is consistent with the behavior previously observed for the NMA spectrum.[69] Additionally, bandwidths determined in this MD-based approach may be overestimated due to the inadequacies of the Amber force field, which can favor rare configurations.[72] Furthermore, the average frequency shift obtained with the B3LYP/6-31++G** is somewhat compromised since the B3LYP functional tends to increase the frequency error which the diffuse functions in the 6-31++G** basis set try to correct (Kim and Keiderling, unpublished results).[65] The major limitation on frequency correction with this initial test of the MD model for hairpins remains its referencing to limited solvation NMA theoretical models and not to structurally relevant empirical peptides.[68] However, this correction is included here for sake of comparison with the literature and to point the way to future

applications.[69] The highest frequency, sharp absorption peak calculated for the amide I mode of the hairpin in a vacuum (Figure 8a) is associated with the Asn side chain and is not observed in experiment. Simulations of the amide I for the hairpin in a COSMO field and with the electrostatic correction due to water solvation (Figure 8b,c) show that this local mode moves down in frequency to couple with the Gly C=O and overlap several other modes arising in the strand residues.

VCD Spectra. The simulated VCD shapes (right-hand side of Figure 8) are more sensitive to the mode ordering than are the absorption profiles. Interestingly, the MD solvent correction tends to just spread out the VCD bands seen in a vacuum, reflecting a persistency of the vacuum FF in the MD correction approach. Neither of these agree very well with the COSMO result, but the average field, or static solvent correction, does have significant similarity to the COSMO amide I, being somewhere between it and the vacuum result. The differences are small but involve modes moving past each other in the sequence from high to low frequency which can strongly affect VCD due to its sign character. Unfortunately, the VCD spectra of Trpzip1 could not be measured due to solubility limitations. Simulations carried out as above, including solvent corrections, for the Trpzip2 structure provided qualitatively the same results. Preliminary experimental data suggest that amide I VCD signal of Trpzip2 is surprisingly weak and predominantly negative, which would be consistent with the simulations and the intensity weakening impact of broadening on VCD intensities (compare $\Delta\epsilon$ values for the static, a′ or c′, and dynamic averaging, d′, results in Figure 8, which are ∼5 times weaker; COSMO values cannot be compared) and with the VCD obtained for a series of hairpin peptides.[31,82]

When the shielding of the peptide backbone by the side chains is included in the model, subtle spectral changes are seen. In the absorption spectra of the Ala-based peptide (Figure 9a) there is a small shift of the amide I absorption band toward lower wavenumbers, which is in agreement with expectations, as the carbonyl groups are more exposed to the solvent leading to added hydrogen bonding, at some level, at least to the externally oriented C=O groups. This shift obviously corresponds to the stronger solvent electrostatic field as represented in Figure 5. In addition to this uniform shift, the normal mode ordering changes, as can be estimated from the VCD spectra simulated for the peptide with and without side chains (Figure 9 bottom). The gross, overall character of the VCD shape is conserved implying that the qualitative role of the side chains in modulating the amide signal is small; however, due to shifts in the detailed ordering of the modes in the amide I region, there is a decided VCD pattern difference in the two spectra. When the side chains are removed, the increased intensity of the positive band (at ∼1710 and 1690 cm^{-1} with and without the side chains, respectively) and smaller magnitude of the lower frequency negative signal (at 1690 and 1670 cm^{-1}) unbalances the net negative VCD predicted by all other models. The lower frequency part of the spectra (below 1550 cm^{-1}) is apparently not so sensitive to the side chain solvent shielding. This region involves vibrations of nonpolar groups, such as C−H bending, which should not be very sensitive to the solvent, but also amide II and amide III modes. The latter movements, however, are associated with the N−H deformation coupled with C(O)−N(H) stretching, the second of which is not as exposed to the solvent as the C=O group. Minor changes are evident in the region of 1450−1500 cm^{-1}, which are associated with the amide II vibration.

Conclusions

Vibrational spectra (IR and VCD) for irregular structures can be accurately simulated by fragmenting the molecule into overlapping sequences, each having the target geometry, so that all amides that are in a continuing sequence in the target molecule are represented in at least one fragment complete with amide near neighbors. The restricted normal mode optimization method enables vibrational relaxation of the model molecular fragments with minimal affect on their geometries. DFT computations for such fragments then yield parameters that can be transferred to the larger irregular molecule for spectral simulation. We have additionally shown that for hairpins, unlike for regular peptide structures, the solvent significantly modulates the amide I vibrations, depending on local conformation. Local hydrogen-bonding and side-chain shielding can strongly limit solvent access to the amides critical for spectra−structure analyses. Corrections for water-solvent interactions with the peptide can be modeled on the basis of a parametrized electric field perturbation model that when combined with MD yields estimates of spectral bandwidths that are dependent on the parametrization scheme. This latter approach needs further optimization to account for local interactions perhaps with more complex peptides and solvation models before application to such irregular systems as hairpins.

References and Notes

(1) Mantsch, H. H.; Chapman, D. *Infrared Spectroscopy of Biomolecules*; Wiley-Liss: Chichester, U.K., 1996.

(2) Haris, P. I.; Chapman, D. *Biopolymers* **1995**, *37*, 251.

(3) Keiderling, T. A.; Silva, R. A. G. D. Conformational studies using infrared techniques. In *Synthesis of peptides and peptidomimetics*; Goodman, M., Herrman, G., Eds.; Georg Thieme Verlag: Stuttgart, Germany, 2002; Vol. E22b; p 715.

(4) Williams, R. W. *Methods Enzymol.* **1986**, *130*, 311.

(5) Schweitzer-Stenner, R.; Eker, F.; Huang, Q.; et al. *J. Phys. Chem. B* **2002**, *106*, 4294.

(6) Asher, S. A.; Ianoul, A.; Mix, G.; Boyden, M. N.; Karnoup, A.; Diem, M.; Schweitzer-Stenner, R. *J. Am. Chem. Soc.* **2001**, *123*, 11775.

(7) Asher, S. A.; Mikhonin, A. V.; Bykov, S. *J. Am. Chem. Soc.* **2004**, *126*, 8433.

(8) Barron, L. D.; Hecht, L.; Blanch, E. W.; Bell, A. F. *Prog. Biophys. Mol. Biol.* **2000**, *73*, 1.

(9) Keiderling, T. A. Peptide and protein conformational studies with vibrational circular dichroism and related spectroscopies. In *Circular Dichroism: Principles and Applications*, 2nd ed.; Berova, N., Nakanishi, K., Woody, R. W., Eds.; Wiley-VCH: New York, 2000; p 621.

(10) Sreerama, N.; Woody, R. W. Circular dichroism of peptides and proteins. In *Circular Dichroism Principles and Applications*; Nakanishi, K., Berova, N., Woody, R. W., Eds.; Wiley-VCH: New York, 2000; p 601.

(11) Venyaminov, S. Y.; Yang, J. T. Determination of Protein Secondary Structure. In *Circular Dichroism and the Conformational Analysis of Biomolecules*; Fasman, G. D., Ed.; Plenum Press: New York, 1996; p 69.

(12) Krimm, S. Vibrational spectroscopy of polypeptides. In *Modern Polymer Spectroscopy*; Zerbi, G., Ed.; Wiley-VCH: 1999; p 239.

(13) Krimm, S.; Bandekar, J. *Adv. Protein Chem.* **1986**, *38*, 181.

(14) Lee, S. H.; Krimm, S. *Biopolymers* **1998**, *46*, 283.

(15) Krimm, S. Interpreting Infrared Spectra of Peptides and Proteins. In *Infrared Analysis of Peptides and Proteins: Principles and Applications*; ACS Symposium Series; Singh, B. R., Ed.; American Chemical Society: Washington, DC, 2000; p 38.

(16) Csaszar, A. G.; Perczel, A. *Prog. Biophys. Mol. Biol.* **1999**, *71*, 243.

(17) Torii, H.; Tatsumi, T.; Kanazawa, T.; Tasumi, M. *J. Phys. Chem. B* **1998**, *102*, 309.

(18) Mirkin, N. G.; Krimm, S. *J. Mol. Struct.* **1996**, *377*, 219.

Vibrational Spectral Simulation for Peptides

J. Phys. Chem. B, Vol. 109, No. 49, 2005 **23697**

(19) Schweitzer-Stenner, R.; Sieler, G.; Mirkin, N. G.; Krimm, S. *J. Phys. Chem.* **1998**, *102*, 118.

(20) Eker, F.; Cao, X.; Nafie, L.; Schweitzer-Stenner, R. *J. Am. Chem. Soc.* **2002**, *124*, 14330.

(21) Schweitzer-Stenner, R. *J. Phys. Chem. B* **2004**, *108*, 16965.

(22) Lee, H.; Kim, S. S.; Choi, J. H.; Cho, M. *J. Phys. Chem. B* **2005**, *109*, 5331.

(23) Choi, J. H.; Cho, M. *J. Chem. Phys.* **2004**, *120*, 4383.

(24) Choi, J. H.; Kim, J. S.; Cho, M. *J. Chem. Phys.* **2005**, *122*, 174903.

(25) Torii, H. *J. Phys. Chem. A* **2004**, *108*, 7272.

(26) Bouř, P.; Keiderling, T. A. *J. Phys. Chem. B* **2005**, *109*, 5348.

(27) Bouř, P.; Kubelka, J.; Keiderling, T. A. *Biopolymers* **2000**, *53*, 380.

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

(29) Kubelka, J.; Silva, R. A. G. D.; Bouř, P.; Decatur, S. M.; Keiderling, T. A. Chirality in peptide vibrations. Ab Initio computational studies of length, solvation, hydrogen bond, dipole coupling and isotope effects on vibrational CD. In *Chirality: Physical Chemistry*; ACS Symposium Series; Hicks, J. M., Ed.; American Chemical Society: Washington, DC, 2002; Vol. 810; p 50.

(30) Silva, R. A. G. D.; Kubelka, J.; Decatur, S. M.; Bouř, P.; Keiderling, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8318.

(31) Hilario, J.; Kubelka, J.; Keiderling, T. A. *J. Am. Chem. Soc.* **2003**, *125*, 7562.

(32) Huang, R.; Kubelka, J.; Barber-Armstrong, W.; Silva, R. A. G. D.; Decatur, S. M.; Keiderling, T. A. *J. Am. Chem. Soc.* **2004**, *126*, 2346.

(33) Kubelka, J. IR and VCD spectroscopy of model peptides. Theory and Experiment. Ph.D. Thesis, University of Illinois at Chicago, 2002.

(34) Kubelka, J.; Huang, R.; Keiderling, T. A. *J. Phys. Chem. B* **2005**, *109*, 8231.

(35) Kubelka, J.; Keiderling, T. A. *J. Am. Chem. Soc.* **2001**, *123*, 6142.

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

(37) Setnička, V.; Huang, R.; Thomas, C. L.; Etienne, M. A.; Kubelka, J.; Hammer, R. P.; Keiderling, T. A. *J. Am. Chem. Soc.* **2005**, *127*, 4992.

(38) Bohr, H. G.; Jalkanen, K. J.; Elstner, M.; Frimand, K.; Suhai, S. *Chem. Phys.* **1999**, *246*, 13.

(39) Knapp-Mohammady, M.; Jalkanen, K. J.; Nardi, F.; Wade, R. C.; Suhai, S. *Chem. Phys.* **1999**, *240*, 63.

(40) Jalkanen, K. J.; Suhai, S. *Chem. Phys.* **1996**, *208*, 81.

(41) Schweitzer-Stenner, R.; Eker, F.; Huang, Q.; Griebenow, K. *J. Am. Chem. Soc.* **2001**, *123*, 9628.

(42) Abdali, S.; Jalkanen, K. J.; Cao, X.; Nafie, L. A.; Bohr, H. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2434.

(43) Jalkanen, K. J.; Elstner, M.; Suhai, S. *J. Mol. Struct. (THEOCHEM)* **2004**, *675*, 61.

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

(45) Cheeseman, J. R.; Frisch, M. J.; Devlin, F. J.; Stephens, P. J. *Chem. Phys. Lett.* **1996**, *252*, 211.

(46) Devlin, F. J.; Finley, J. W.; Stephens, P. J.; Frisch, M. J. *J. Phys. Chem.* **1995**, *99*, 16883.

(47) Bak, K. L.; Devlin, F. J.; Ashvar, C. S.; Taylor, P. R.; Frisch, M. J.; Stephens, P. J. *J. Phys. Chem.* **1995**, *99*, 14918.

(48) Stephens, P. J.; Devlin, F. J.; Ashvar, C. S.; Chabalowski, C. F.; Frisch, M. J. *Faraday Discuss.* **1994**, *99*, 103.

(49) Olsen, J.; Roos, B. O.; Jorgensen, P.; Jensen, H. J. *J. Chem. Phys.* **1988**, *89*, 2185.

(50) Bouř, P.; Kubelka, J.; Keiderling, T. A. *Biopolymers* **2002**, *65*, 45.

(51) Kubelka, J.; Silva, R. A. G. D.; Keiderling, T. A. *J. Am. Chem. Soc.* **2002**, *124*, 5325.

(52) Kuznetsov, S. V.; Hilario, J.; Keiderling, T. A.; Ansari, A. *Biochemistry* **2003**, *42*, 4321.

(53) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5578.

(54) Russell, S. J.; Blandl, T.; Skelton, N. J.; Cochran, A. G. *J. Am. Chem. Soc.* **2003**, *125*, 388.

(55) Du, C. G.; Zhu, Y. J.; Huang, C. Y.; Gai, F. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 15915.

(56) Yang, W. Y.; Pitera, J. W.; Swope, W. C.; Gruebele, M. *J. Mol. Biol.* **2004**, *336*, 241.

(57) Ponder, J. W. *Tinker*, version 3.8; Washington University School of Medicine: St. Louis, MO, 2000.

(58) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

(59) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.6 ed.; Gaussian, Inc.: Pittsburgh, PA, 1998.

(60) Bouř, P.; Keiderling, T. A. *J. Chem. Phys.* **2002**, *117*, 4126.

(61) Bouř, P. *Collect. Czech. Chem. Commun.* **2005**, *70*, 1315.

(62) Becke, A. D. Exchange-correlation approximations in density-functional theory. In *Modern electronic structure theory*; Yarkony, D. R., Ed.; World Scientific: Singapore, 1995; Vol. 2; p 1022.

(63) Becke, A. D. *J. Chem. Phys.* **1988**, *88*, 2547.

(64) Bouř, P.; McCann, J.; Wieser, H. *J. Chem. Phys.* **1998**, *108*, 8782.

(65) Kubelka, J.; Keiderling, T. A. *J. Phys. Chem. A* **2001**, *105*, 10922.

(66) Chen, X. G.; Schweitzer-Stenner, R.; Krimm, S.; Mirkin, N. G.; Asher, S. A. *J. Am. Chem. Soc.* **1994**, *116*, 11141.

(67) Lybrand, T. P.; Ghosh, I.; McCammon, J. A. *J. Am. Chem. Soc.* **1985**, *107*, 7793.

(68) Bouř, P. *J. Chem. Phys.* **2004**, *121*, 7545.

(69) Bouř, P.; Michalík, D.; Kapitán, J. *J. Chem. Phys.* **2005**, *122*, 144501 1.

(70) Bouř, P.; Keiderling, T. A. *J. Chem. Phys.* **2003**, *119*, 11253.

(71) Ham, S.; Kim, J. H.; Kochan, H.; Cho, M. *J. Chem. Phys.* **2003**, *118*, 3491.

(72) Corcelli, S. A.; Lawrence, C. P.; Skinner, J. L. *J. Chem. Phys.* **2004**, *120*, 8107.

(73) Hayashi, T.; Jansen, T. C.; Zhuang, W.; Mukamel, S. *J. Phys. Chem. A* **2005**, *109*, 64.

(74) Kubelka, J.; Kim, J.; Bouř, P.; Keiderling, T. A. *Vib. Spectrosc.* **2005**, submitted for publication.

(75) Brauner, J. W.; Dugan, C.; Mendelsohn, R. *J. Am. Chem. Soc.* **2000**, *122*, 677.

(76) Fang, C.; Wang, J.; Charnley, A. K.; Barber-Armstrong, W.; Smith, A. B., III; Decatur, S. M.; Hochstrasser, R. M. *Chem. Phys. Lett.* **2003**, *382*, 586.

(77) Kim, Y. S.; Wang, J. P.; Hochstrasser, R. M. *J. Phys. Chem. B* **2005**, *109*, 7511.

(78) Paul, C.; Wang, J.; Wimley, W. C.; Hochstrasser, R. M.; Axelsen, P. H. *J. Am. Chem. Soc.* **2004**, *126*, 5843.

(79) Measey, T.; Schweitzer-Stenner, R. *Chem. Phys. Lett.* **2005**, *408*, 123.

(80) Measey, T.; Hagarman, A.; Eker, F.; Griebenow, K.; Schweitzer-Stenner, R. *J. Phys. Chem. B* **2005**, *109*, 8195.

(81) Bouř, P.; Keiderling, T. A. *J. Mol. Struct. (THEOCHEM)* **2004**, *675*, 95.

(82) Zhao, C.; Polavarapu, P. L.; Das, C.; Balaram, P. *J. Am. Chem. Soc.* **2000**, *122*, 8228.