Two-Dimensional Infrared Spectroscopy as a Probe of the Solvent Electrostatic Field for a Twelve Residue Peptide†

Jianping Wang,‡,§ Wei Zhuang,|| Shaul Mukamel,|| and Robin Hochstrasser*,‡

Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 9104-6323, and Department of Chemistry, University of California Irvine, Irvine, California 92697-2025

Received: July 19, 2007; In Final Form: September 21, 2007

The linear IR and two-dimensional (2D) IR spectra of the amide-I modes of the 12-residue β-hairpin peptide tryptophan zipper-2 (SWTWENGKWTWK) and its two ^{13}C isotopomers were simulated, with local mode frequencies evaluated by two solution-phase peptide amide-I frequency maps proposed recently: an electrostatic potential map and an electrostatic field map. Both maps predict a set of nondegenerate local amide-I mode transition energies for the hairpin. Spectral simulations using both maps predict the main spectral features of the linear IR and 2D IR experimental results of the ^{13}C-labeled and -unlabeled hairpin. The radial distribution functions obtained using trajectories from classical molecular dynamics simulations demonstrate different water distributions at different sites of the hairpin. Our results suggest that the observed difference of the ^{13}C-shifted band, including its peak position and frequency distributions for different isotopomers, in both linear IR and 2D IR spectra, is likely to be due to the difference in the local environment of the solvated peptide. Ab initio density functional theory calculations show a residue-independent ^{13}C shift of the amide-I mode, further supporting the result. The variations of these shifts are attributed to the residue level heterogeneity of the electrostatic environment of the peptide. Our results show that 2D IR of peptide with single ^{13}C isotopic labeling can be used to probe the electrostatic environment of the peptide local structure.

Introduction

Two-dimensional infrared (2D IR) spectroscopy in conjunction with ab initio quantum computations and molecular dynamics simulations is a promising approach to determining the dynamics of structure changes of peptides.[1,2] However, in order to allow this novel approach to reach its full potential, there is a great need for predictive theories of the vibrational spectra of the peptide backbone in the presence of water. It is now evident that calculations of spectra that do not incorporate specific solvent effects on the frequencies and transition dipoles of the amide unit will not be successful in predicting many essential details of the spectra of peptides. The dynamic shifts caused by hydrogen bonding and other charge effects are substantial and need to be combined with computations of intermode coupling to predict the characteristic infrared spectra of the various secondary structure motifs.

It has been shown through numerous experiments using FTIR[3−5] and 2D IR[6−9]experiments that isotope selective labeling of many types of peptide environments can permit structure and dynamics to be examined on a residue-by-residue basis. One of the main goals of 2D IR methods is to simplify the broadband spectra of peptides and define more clearly the underlying features and their interactions with water. However, the amide vibrations of peptides form groups of modes, each having a near degeneracy of the number of residues, and the modes in these groups are not normally identified separately in solution-phase experiments. Even very simplified empirical theories of peptide IR spectra[10,11] often capture approximately the spectral shapes for particular secondary structures if the coupling constants happen to be very large and of the easily manageable dipole−dipole type. However, such approaches do not claim to predict the correct spectral shifts from isolated molecule states, and they contain arbitrariness[7] in the choice of their line shapes, zero-order frequencies and fluctuations, prior to any coupling, for modes from different residues that might be experiencing very different hydrogen bonding or electric fields from the remaining secondary structure or the solvent. The computation of linear and 2D IR spectra of individual residues shifted by isotope replacement would constitute a much more stringent test of the theoretical methods.

2D IR methods use infrared pulse sequences to measure mode couplings and the dynamics of the mode frequency distributions.[12−14] The high mode selectivity of the methods arise from their double resonance character: the effects of driving one fundamental are sensed through the responses felt by other modes. Because the methods naturally separate the effects of molecular structure distributions from spontaneous relaxation processes, they are line narrowed compared with FTIR spectra, and they can be used to determine the spectral diffusion, the equilibrium dynamics, and the relaxation kinetics. They can access local structure in a unique manner to permit significant questions to be asked in protein dynamics that might not yet be answerable by other methods. The 2D IR spectroscopy is being used in a variety of fields as a result of significant theoretical, technical, and scientific advances from many different laboratories as illustrated in numerous very recent examples.[9,15−24] The signal field in these experiments is from a distribution of vibrational frequencies, representing the range of chemical environments for each vibrator with its own relaxation function,

† Part of the "Attila Szabo Festschrift".

* Corresponding author. E-mail: hochstra@sas.upenn.edu. Phone: 215-898-8203. Fax: 215-898-0590.

‡ University of Pennsylvania.

§ Present address: Institute of Chemistry, Chinese Academy of Sciences, Beijing 100080, China.

|| University of California Irvine.

10.1021/jp075683k CCC: $40.75
Published on Web 12/14/2007

dependent on its spatial location, chemical structure, and solvent induced dynamics. The 2D IR method greatly enlarges the range of biological applications of vibrational spectroscopy, but there is considerable scope for improvements, both theoretical and experimental, that will optimize its broad applicability. Of particular importance are improvements in the confidence level of spectra-structure relationships.

The 2D IR of the polypeptides has been based mainly on the structure sensitive amide-I and amide-II transitions which are carbonyl stretching and in-plane CN−H bending and the amide-A mode, the N−H or N−D stretches.[25] The potential energy surfaces (PES) of the amide motions in proteins and peptides depend on the electronic configuration of the solute and also on the solute−solvent interactions. Frequently, the solute−solvent interactions will be mainly electrostatic, so that modeling of the solvent influence by an inhomogeneous electrostatic field may often provide a good representation of solvent effects on vibrational spectra. The electrostatic effect on a polarizable molecule can also be approximated from the field and its gradients at a limited number of points. In any event there is a great need for accurate descriptions of the variation of properties such as the vibrational frequencies and transition dipoles as functions of the electrostatic potential and its gradients at a number of points within the solute molecule. Such maps can be used in combination with electric fields computed from classical molecular dynamics simulations to compute spectra. Correlations between these various spectral parameters associated with vibrations have recently been developed to describe the infrared spectra and dynamics of water[26,27] and of amide units. Recently, several electrostatic maps have been proposed for the amide-I vibration frequency based on a model peptide unit *N*-methyl acetamide (NMA, $CH_3CONHCH_3$), with electrostatic potential samplings at four to six atomic positions[28−31] or at certain locations in chemical bonds.[32] More recently, the electric field gradients on four atoms were considered in the construction of a map.[33] Another map by Skinner and co-workers[30] for the amide-I motions was based on the field samplings at four atoms (C, O, N, and H) of the NMA molecule. The same authors also constructed a map for HOD based on samplings at the locations of each of the atoms.[26] The correlation of the OH stretch in HOD with the electric-field component directed along the bond was also used to compute 2D IR responses.[27] In work that is most germane to the present contribution, an anharmonic vibrational Hamiltonian for the amide I, II, III, and A modes of NMA was recast in terms of 19 components of an external electric field and its first and second derivative tensors aimed at sampling the global field in the entire spatial region of the amide modes.[31]

Several currently available NMAD maps[28−31,33,34] for the amide-I modes in proteins and peptides were constructed on the basis of quantum chemical computations of NMAD in water clusters. The electrostatic potentials or fields can easily be computed from classical force fields that give rise to these molecular structures. In this manner, both the electrostatic and the chemical bonding effects are incorporated, but the results are sensitive to the details of the MD force field. In contrast, by introducing electric fields or point charges directly in the ab initio calculations, the pure electrostatic effects can be isolated. A density functional theory (DFT) map based on such a principle[31] has recently been established and is believed to be generally transferable. Nonelectrostatic interactions such as those arising from dispersion forces and chemical bonding are not accounted for in any of the NMAD maps; however, their effect on the frequency fluctuations can be separately evaluated as dictated by the situation.

In this paper, the simulated amide-I region linear IR and 2D IR spectra of the tryptophan zipper peptide, trpzip2[35] in a β-hairpin conformation, and its ^{13}C isotopomers are presented. Two representatives of the existing NMAD maps, namely, the electrostatic potential map proposed by Cho and co-workers[28] and the electrostatic field map proposed by Mukamel and co-workers,[31] have been used to compute the solution-phase local mode frequencies of the β-hairpin peptide. All spectral features including the ^{13}C-shifted regions have been compared with recent experimental results.[7] We also examined local conformational fluctuations of the peptide including backbone and side chain and solvent molecules via classical MD simulations. The peptide-solvent hydrogen bonding is examined from the radial distribution functions for selective locations on the hairpin. We then obtain instantaneous local mode frequencies using the two maps and compare them with a simple local mode picture.[7] The zero-order ^{13}C shift of the amide-I mode is evaluated by using ab initio DFT calculations.

Experimental Section

A description of the sample and its FTIR and 2D IR spectra at 278 K in the 6 μm frequency region have been reported in a recent paper where all of the experimental details can be found.[7] Briefly, the trpzip2 β-hairpin, having the sequence SWTWENGKWTWK and consisting of 12 residues but 13 amide units was synthesized by AnaSpec (San Jose, CA). The 13th amide unit is located on the side chain of the Asn6 residue. Two peptide isotopomers, having $^{13}C{=}O$ on Trp2 (denoted as L2) and $^{13}C{=}O$ on Gly7 (denoted as L7), were prepared along with the unlabeled one (denoted as UL). The peptides were dissolved in D_2O at a concentration of approximately 10 mM (pD = 2.5). The 2D IR experiments were performed with IR laser pulses having width of 74 fs. The time domain signal $S(\tau, T = 0, t)$ is obtained from a discrete inverse Fourier transform of the raw data in the frequency domain.[36] A double Fourier transform along the coherence time τ and the detection time t axes generates the rephasing $S_R(-\omega_\tau, \omega_t)$ and non-rephasing $S_{NR}(\omega_\tau, \omega_t)$ spectra, and their sum yields the 2D IR correlation spectrum $S(\omega_\tau, \omega_t)$ shown in the figures. The 2D IR correlation spectra reported here correspond to the $\langle zzzz \rangle$ tensor element of the third-order nonlinear response.[37]

Theoretical Description

Molecular Dynamics Simulations. The MD simulations of trpzip2 in explicit solvent water have been carried out using the NAMD2.5 package.[38] The initial structure of trpzip2 was chosen as the first of the 20 reported NMR structures.[35] The system includes 3369 explicit water molecules of TIP3P type. It also includes two chloride ions to keep the entire system neutral. The cubic solvent box is 48 × 48 × 48 Å in dimension. The CHARMM27 force field[39] was used to compute all interactions. The cutoff of 12 Å was used for the nonbonded interactions. The long-range electrostatic interaction was evaluated via the Ewald sum approach.[40] The system was first energy minimized at 0 K, followed by a slow heating procedure (1 fs step for 20 ps). The system was then equilibrated at 278 K for 1 ns at a time step of 2 fs. The data were taken at constant number-pressure-temperature condition at a time step of 2 fs for 1 ns. The total number of snapshots is 5×10^5, in which 500 evenly distributed snap shots were used to compute the one-dimensional (1D) and 2D IR spectra of the trpzip2 ^{13}C isotopomers.

Spectral Simulations. The computational package Spectron was used to calculate the 1D IR and 2D IR spectral signals. The protocol has been described recently.[41] Here, we give a brief description. We start with the vibrational exciton Hamiltonian:

$$H = \hat{H}_S + \hat{H}_F$$

where

$$\hat{H}_S = \sum_m \epsilon_m \hat{B}_m^+ \hat{B}_m + \sum_m^{m\neq n} J_{mn} \hat{B}_m^+ \hat{B}_n - \frac{1}{2}\sum_m \Delta_m B_m^+ \hat{B}_m^+ \hat{B}_m \hat{B}_m$$

is the system Hamiltonian and $\hat{H}_F$ is the interaction with the optical field, $E(t)$:

$$\hat{H}_F = -E(t)\cdot\sum_m \mu_m (\hat{B}_m^+ + \hat{B}_m)$$

and $\hat{B}_m^+$ ($\hat{B}_m$) is the Boson creation (annihilation) operator for the mth amide-I mode, with frequency ϵ_m, anharmonicity Δ_m, and transition dipole moment μ_m. J_{mn} is the harmonic inter-mode couplings. All parameters of $\hat{H}_S$ fluctuate because of conformational changes of backbone, side-chain, and solvent dynamics.

The frequency of mode m subjected to by a time dependent electric field is given by $\epsilon_m(E) = \epsilon - \delta\epsilon_m(E)$. $\epsilon = 1717$ cm^{-1} is the frequency of an isolated NMA in the gas phase.[31] The vibrational frequency shift $\delta\epsilon_m(E)$ and the anharmonicity $\Delta_m(E)$ can be obtained from an ab initio map which relates the electric field to the fundamental and overtone frequencies of the amide-I mode. Recently developed ab initio maps[28−31,33,34] of amide-I mode frequency can be separated into two kinds by their strategies: (I) quantum chemistry computation is carried out for an ensemble of NMA−solvent molecule clusters to obtain the frequencies while the electric fields at certain points of an NMA molecule (for instance C, O, N, or H atoms) are calculated for each cluster using a point charge model for the solvent molecules. Then a least-square fitting procedure is used to obtain a general analytical relation between the NMA frequency and the electric field. (II) The quantum chemistry computation is carried out directly for a NMA molecule in different electric fields. The electric field can be incorporated by adding an analytical term in the energy calculation (Mukamel model)[31] or by surrounding NMA with a certain distribution of point charges (Knoester model).[33] Finally, a general relation between the NMA frequency and the electric field can be derived in a way similar to strategy I.

We use two different electrostatic interaction maps for the evaluation of local mode frequencies in the 1D and 2D IR spectra simulation of trpzip2: the map by Cho and co-workers,[28] which is an electrostatic potential map we refer to as the HKLC map; and the map by Mukamel and co-workers,[31] which is an electrostatic field map, referred to as HZM. Figure 1 is an illustration showing how the potential and field models work for a given peptide. The peptide main chain, side chain, and solvent molecules have been taken into account in calculating the amide-I mode frequency shift on the basis of electrostatic interactions. Basically, the local potential or field is calculated for each of the amide units in trpzip2, at several reference points. For example, a single reference point is chosen to be at the midpoint of the C═O bond in the field map, but four reference points are chosen to be at C, O, N, and H atoms in the potential model. The calculated electrostatic contributions from the backbone, side chain, and solvent of the solvated peptide system

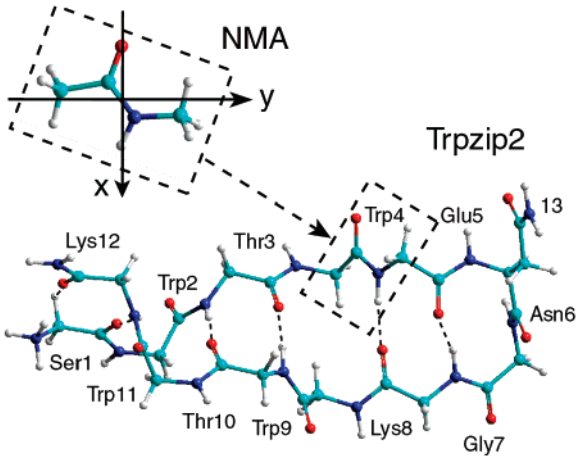

Figure 1. Illustration of the peptide unit (NMA) whose field maps are used to compute the electrostatic field at different sites (visualized in the figure for Trp4) of the tryptophan zipper-2 peptide.

significantly shift a local mode frequency from the gas (NMA) to the solution by about 60−90 cm^{-1}. In order to roughly fit experimental linear IR and 2D IR, we set $\epsilon = 1717$ cm^{-1} and $\epsilon = 1728$ cm^{-1} for the HKLC and HZM maps, respectively. By using MD trajectories, the calculated frequency fluctuation is instantaneous, and the calculated solution-phase local mode frequency can be used either to compute the 1D and 2D spectra or to characterize the time behavior of the local amide-I modes. The ^{13}C shift of the amide-I mode is taken to be 44 cm^{-1} in both 1D and 2D IR spectral simulation. In the simulation, a half width at half-maximum due to the lifetime broadening was chosen as 8 cm^{-1}.[12] By using Spectron, the so-called rephasing spectra and nonrephasing spectra were computed using the sum-over-states approach.[38] The correlation (or absorptive) spectra were the summation of equally weighted rephasing and nonrephasing spectra.

Vibrational Couplings. The couplings of different amide modes were assumed to depend solely on the peptide backbone structure. For nearest covalently bonded modes, we used an available ab initio map.[42] However, there are other ab initio maps[36,43,44] that are more accurate which will be incorporated into the simulation package in the future. The couplings can be evaluated conveniently by using an electrostatic transition dipole coupling (TDC) scheme, proposed by Krimm and his co-workers,[25,45] by an electrostatic transition charge coupling with charge fluxes,[43,46] or by a transition charge density derivative distribution approach proposed recently.[47] For two nearby amide units, the transition charge-based approaches have advantages over the TDC because the dipole approximation is inaccurate in this case. However, as the inter-amide distance increases, all of the approaches tend to give the same answer. In our simulation, all the non-nearest neighbor couplings were calculated using the TDC approach, which is currently available in the Spectron code. The TDC is computed by using the following formula:

$$J_{mn} = \frac{0.1A}{\epsilon}\frac{(\mu_m\cdot\mu_n) - 3(\mu_m\cdot e_{mn})(\mu_n\cdot e_{mn})}{r_{mn}^3}$$

where μ_m is the transition dipole in D Å^{-1} u$^{-1/2}$ units, r_{mn} is the distance between dipoles (in angstroms), e_{mn} is the unit vector connecting mth and nth vibrator, and $\epsilon = 1$ is the dielectric constant. The angle between the transition dipole and the C═O bond is chosen as 10°. This angle has been used as an adjustable parameter in transition dipole coupling models.[10] We found that

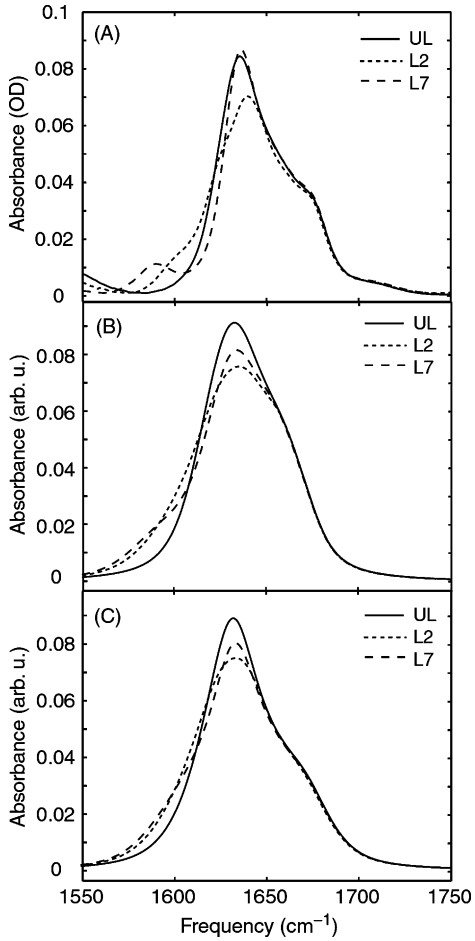

Figure 2. Experimental linear IR (A) and simulated IR spectra of trpzip2 ^{13}C isotopomers using the HKLC electrostatic potential model (B) and the HZM electrostatic field model (C). Solid lines, UL; dashed lines, L2; dotted lines, L7.

10° reproduces best the observed FTIR shapes, but the results of the paper are not changed by choosing the angle within the generally accepted range from 10 to 20°. The conversion factor $A = 848619/1650$ gives the coupling energy in cm^{-1}.

Ab Initio Calculations of the ^{13}C Shift of the Amide-I Mode. Since the electrostatic contributions to isotopical bands are being compared, it is necessary to know the isotope shifts. Ab initio DFT calculations have been used to evaluate the ^{13}C isotope effect on the amide-I mode frequency in a model β-hairpin by using the Gaussian03 program.[48] A six unit model hairpin with a type I′ turn was constructed as Ac-GlyAsnGlyGly-NMe in the conformation of the first solution structure (PDB code 1LE1 in the protein data bank[49]) determined by NMR.[35] The normal-mode frequency calculation was performed at the level of B3LYP/6-31+G*, after partial geometry optimization with fixed dihedral angles. Local mode frequencies were obtained by the wave function demixing of the harmonic normal modes.[43] The ^{13}C═O was placed on the first or fifth amide unit of the model peptide to mimic the situation of L2 or L7 of the trpzip2. The zero-order isotope shift was estimated as a difference between the local mode frequency with and that without ^{13}C substitution.

Results and Discussion

1D IR and 2D IR Results: Experimental and Simulation. Figure 2 shows experimental linear IR spectra of trpzip2 ^{13}C isotopomers in the amide-I region in panel A.[7] The linear IR shows two main transitions: the low frequency, stronger transition peaked at ~1640 cm^{-1} is presumably due to the across-chain in-phase and along-chain out-of-phase of the C═O motions, whereas the high frequency, weaker transition peaked at ~1675 cm^{-1} is mainly due to the across-chain out-of-phase and along-chain in-phase C═O motions. Two isotopomers show different ^{13}C effects: the ^{13}C-shifted band is ~10 cm^{-1} lower in L7 than in L2 (1590 vs 1600 cm^{-1}). The simulated spectra using the HKLC model and the HZM model are shown in panels B and C, respectively. Although the high-frequency component is slightly stronger in the field model, two models generally reproduce the main spectral feature of the unlabeled β-hairpin. In addition, both the potential model and the field model predict a small difference in the observed ^{13}C shifts in L2 and L7, and the difference is slightly larger in the field model than in the potential model.

Figure 3 shows experimental 2D IR spectra of the trpzip2 ^{13}C isotopomers in panels A−C. For the spectrum in panel A, the diagonal signals in red color are due to 0−1 transitions, and those in blue are due to 1−2 transitions. Two main 0−1 transitions have their peak maxima in agreement with those of FTIR, as can be seen when a projection of the 2D spectrum is made onto the $|\omega_t|$ axis. The off-diagonal cross-peaks appear because of pairwise vibrational couplings among local amide-I modes.[7] The spectral features, including the diagonal and off-diagonal peaks, change upon ^{13}C labeling as shown in panels B and C. The simulated spectra using the HKLC potential model are shown in panels D−F, and the results using the HZM field model are given in panels G−I. The 2D IR characteristics of the UL, L2, and L7 are reasonably computed in using both models.

Local Solvation of the β Hairpin. To examine the local solvation of the β hairpin, we use MD simulations. MD simulations with explicit solvent molecules allow us to examine the distribution of water molecules surrounding the peptide. Here, we focus on two sites, namely, site 2 and site 7. Figure 4 shows two fragments of the β hairpin which emphasize the amide unit of either site 2 or site 7. The snap shot is taken at 20 ps. Water molecules within 3.5 Å from the O atom of the carbonyl group are shown along with amide units and peptide backbone structures. However, bulky side chains were removed to simplify the picture. Both of the C═O groups are solvent faced and are not intervolved in internal hydrogen bonding. However, they are located at two different regions of the hairpin peptide: site 2 is near the end of the strand, whereas site 7 is near the type-I′ turn region. It is shown that, in each site, there are two nearby water molecules. In the case of site 2, one hydrogen bond is formed between the peptide C═O and a solvent water molecule in the snap shot of Figure 4. Site 2 is a tryptophan residue having a bulky hydrophobic side chain. At the same time, two solvent−water hydrogen bonds are formed at site 7, which is glycine and has the least bulky side chain. The shortest O−H distances are also somewhat shorter in the case of site 7, indicating a slightly stronger H bond formed between glycine C═O and solvent water molecules. Although both of the backbone C═O groups face the solvent, the number of hydrogen bonds formed between the C═O and the solvent water molecules are often different. This difference can be attributed to the steric effect of the side chain and also to the peptide secondary structure. This snap shot is typical of the average results from the radial distribution discussed below.

Figure 5 shows the radial distribution function $g(r)$ of the hydrogen atoms of water and the amide oxygen atom of site 2 (dashed line) and site 7 (solid line). It is evident from the figure that the externally hydrogen-bonded O−H distances are com-

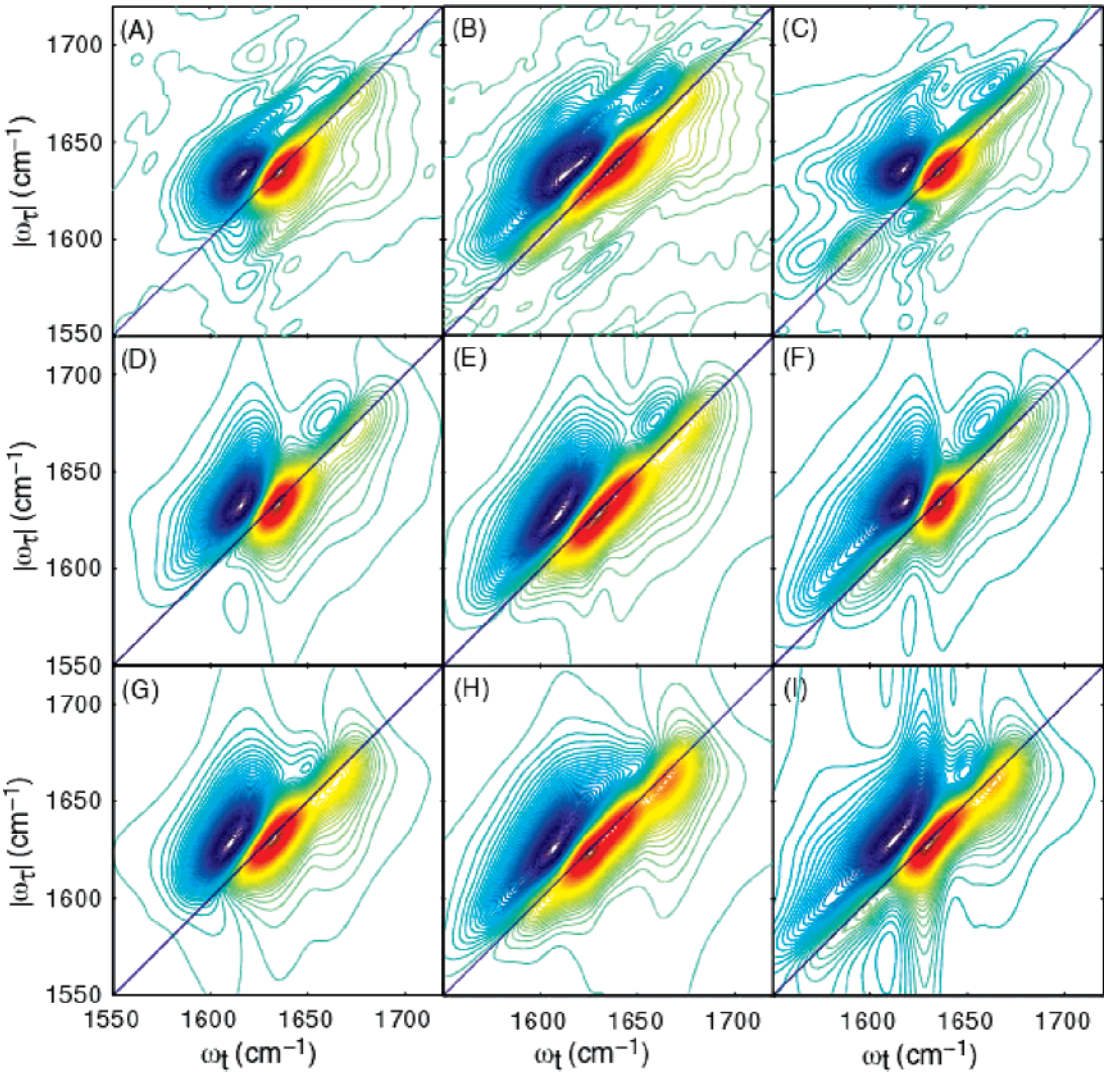

Figure 3. Experimental 2D IR and simulated IR spectra of trpzip2 ^{13}C isotopomers. (A−C) experimental results; (D−F) simulated results using the HKLC model; (G−I) simulated results using the HZM model. A, D, and G are unlabeled; B, E, and H are L2, and C, F, and I are L7.

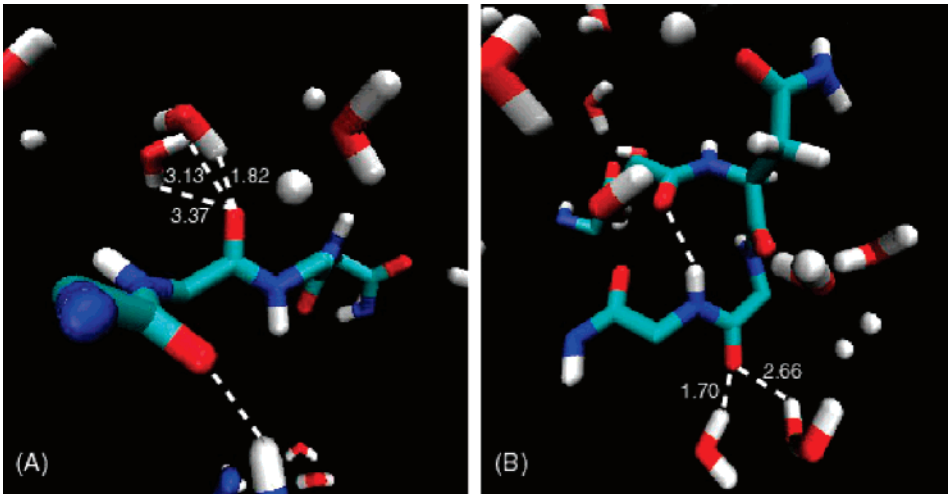

Figure 4. Trpzip2 β-hairpin configuration fragment with nearby water molecules. (A) Site 2. (B) Site 7. The snap shot is taken at 20 ps. Bulky side chains were removed to more clearly reveal the carbonyl/water configurations.

parable in both cases (ca. 1.8 Å). However, the average number of hydrogen-bonded solvent H atoms is close to 1.5 in site 7, indicating a hydrophilic environment in the vicinity of the amide C═O group that is located in the turn region. The number drops to 1 for the case of site 2, which is consistent with the fact that the environment of site 2 is somewhat hydrophobic near the end of the peptide chain where there are two tryptophan residues on each side of the hairpin. In addition, the peaks are narrower, the valleys have smaller r values, and the $g(r)$ function has lower values at longer distances in the case of site 2, suggesting a thinner water layer and fewer water molecules as r increases. This further supports the argument that indeed the hydration

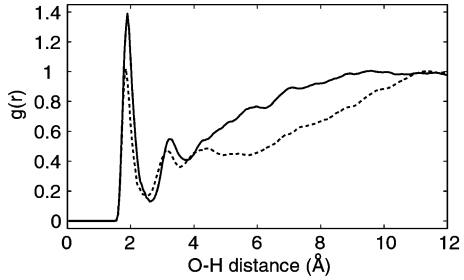

Figure 5. Radial distributions of the hydrogen atoms in solvent water molecules and the oxygen atom in the amide unit at site 2 (dashed line) and site 7 (solid line).

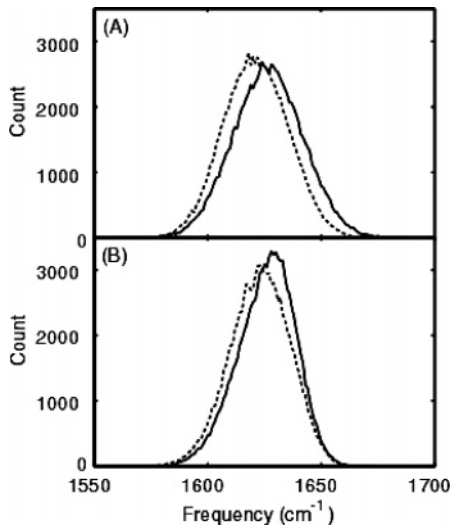

Figure 6. Local mode frequency distribution for site 2 (solid curves) and site 7 (dotted curves) by using the HKLC model (A) and HZM model (B). With respect to the peak maximum of site 2, that of site 7 is 3 cm^{-1} blue-shifted in A and 6 cm^{-1} blue-shifted in B.

state of the amide unit at site 2 is influenced by its hydrophobic environment.

Amide-I Local Mode Frequency Distributions of the β Hairpin. In principle, the local mode frequencies can be obtained by the ab initio computations-based wave function demixing approach as described earlier.[43] In the present study, the local mode frequency distributions were obtained for all 13 amide-I local modes. Figure 6 shows the results for site 2 and site 7. The statistics were obtained for 1 ns MD simulations with a step of 2 fs. The results using the HKLC map are given in panel A, and those of the HZM map are given in panel B. The two panels compare the local mode distribution of site 2 (solid curve) with that of site 7 (dashed curve). The peak location for site 7 is at a significantly lowered frequency than that of site 2. The mean frequencies of site 2 and site 7 show a greater than 3 cm^{-1} difference in Figure 6A for the potential map and a greater than 6 cm^{-1} difference in Figure 6B for the field map. These results suggest that electrostatic interactions including backbone, side chain, and solvent molecules lead to a lower local mode frequency for site 7 than site 2. Further analysis showed that the contribution from solvent interactions dominates this frequency shift (data not shown). Although the vibrational excited states are excitonic in nature, the local mode properties are very useful especially when a heavy atom isotope is introduced into the vibrator to lower its frequency. The results verify that one can use the isotopically labeled 2D IR spectra to examine the local mode characteristics. By examining the characteristics of the 2D spectra of L7 compared with that of L2, the local environment of specific peptide backbone sections

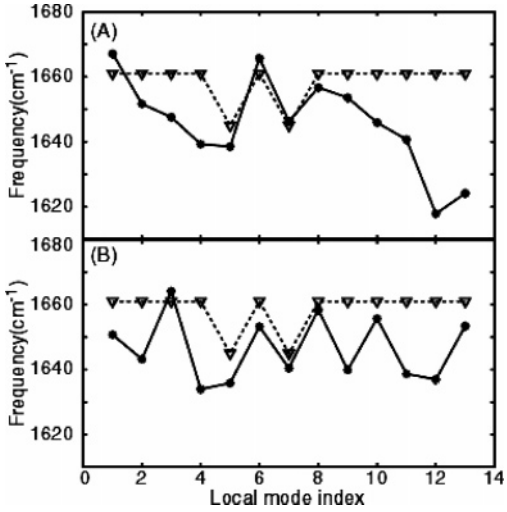

Figure 7. Calculated mean local mode frequency at each site using the HKLC model (panel A, solid curves) or the HZM model (panel B, solid curves), in comparison with an empirical search discussed in the text (dashed curves with down triangles in both panels).

could be evaluated experimentally. These characteristics include peak shape, intensity, orientation, and aspect ratio as was pointed out earlier.[7] Spectral simulations using the two maps that are electrostatic interaction in nature reproduce the main feature of both linear IR and 2D IR spectra, including the isotopic labeling induced spectral regions in L2 and L7 (see Figure 2), suggesting that it is reasonable to conclude that isotopically shifted bands are very sensitive to the environmental electrostatic interactions. In other words, when the solvent interaction dominates, the 2D IR technique could be used to characterize peptide−solvent electrostatic interactions at the residue level or serve as spectroscopic sensors of the peptide environment in water.

These arguments are further supported by the observation that the solvent inhomogeneity can be characterized from the local mode frequency distributions. For example, the frequency distribution is less symmetric for the amide-I mode in site 2 than that in site 7 (see Figure 6), suggesting that the frequency is likely to be more inhomogeneously distributed in site 2 as a result of solvent interactions. This is also supported by the radial distribution functions shown in Figure 5. The solvent density differs in sites 2 and 7 so different electric potentials and field strengths act at the two sites. In addition, the potential model predicts a less symmetric mean frequency distribution for the amide-I units on the two β-hairpin strands: the mean frequencies are somewhat lower for units 11 and 13 than those of units 1−3. However, a more or less symmetric frequency distribution for the two strands is maintained through out the entire 13 amide units for the field model (Figure 7, panel B).

In a recent study of the 1D and 2D IR spectra of trpzip2 using traditional exciton theory and incorporating empirical frequency shifts, the local mode or zero-order frequencies were systematically varied to generate a reasonable agreement with the experimental band locations.[7] The results of this exercise are compared with the present computations in Figure 7, and they are quite similar, particularly in the turn region (amide unit 5−8). For example, the local frequency of site 5 is predicted by the two maps to be similar to that of site 7, which is also in agreement with the empirical modeling.

Comparison of Figure 2B,C shows that the HKLC model computation provides evidence of a smaller apparent exciton splitting. Figure 6 shows that HKLC contains more diagonal disorder which, although it leads to linear spectral widths that are very similar to those of HZM, the peptide secondary

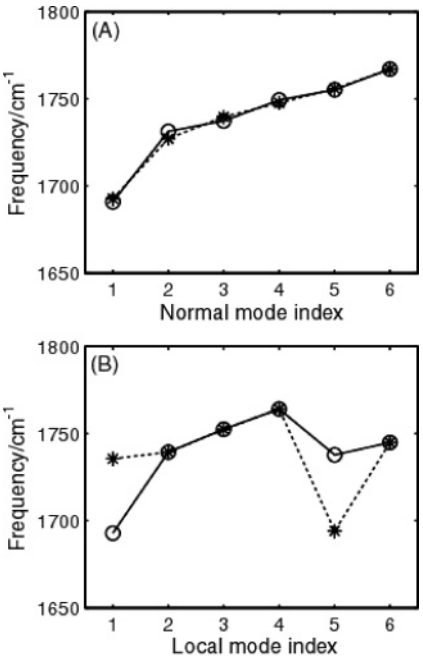

Figure 8. Calculated normal mode and local mode frequencies of a penta-peptide Ac-GlyAsnGlyGly-NMe in the conformation of the β-hairpin with ^{13}C labeled at first (circle, solid line) or fifth (star, dashed line) amide unit: (A) normal-mode frequencies; (B) local mode frequencies. The resultant ^{13}C shift is 42.7 cm^{-1} for site 1 and 43.7 cm^{-1} for site 5.

structure marker is smeared out as normally occurs with excitons in more disordered media.

^{13}C Shift of the Amide-I Modes. Since the simulated spectra using both maps do not fully reproduce the observed spectral shifts of L2 and L7, it was decided to examine the possible variations in the ^{13}C-isotopic shift that are expected because the force fields are different. A significant question is whether the ^{13}C-isotopic shifts vary significantly for different amide-I modes corresponding to different residues in a polypeptide chain. Ab initio DFT computations were used to evaluate the normal mode and local mode frequencies of a pentapeptide, Ac-GlyAsnGlyGly-NMe in a structure identical to the NMR solution structure of trpzip2 from residue 4 to 8. A ^{13}C label was placed either at the first or at the fifth amide unit to mimic the case of L2 or L7 of trpzip2. The normal-mode frequencies given in Figure 8A were first computed. The local mode frequencies of the peptide are shown in Figure 8B. The case when the ^{13}C label is on the first amide unit is shown as solid curves, whereas that on the fifth amide unit is shown as dashed curves. The difference between the two curves in panel B gives the ^{13}C shifts. It is found that the ^{13}C shift is 42.7 cm^{-1} for site 1 and 43.7 cm^{-1} for site 5, suggesting that the ^{13}C shift is essentially the same for these two sites in a hairpin conformation. We also calculated the isotopic shifts at other sites, for example, site 6, and the result is within 1 cm^{-1} of these two values. The calculations were also performed at the HF/6-31+G* level of theory yielding a ^{13}C shift of approximately 46.1 cm^{-1} at all sites. These results indicate that the calculated ^{13}C shift of ~44 cm^{-1} in a local amide-I mode picture is more or less independent of the peptide conformation for these force fields. The value is very close to the one used in our earlier 2D IR spectral simulations.[50] The effects caused by ^{13}C in natural abundance are omitted from consideration in this work, although they are well-known in the spectra of proteins and large peptides. The heterogeneity of the site environments and their couplings causes the peak from naturally abundant ^{13}C molecules to be diffuse and contribute less at each frequency than the 13% abundance of labeled molecules would suggest.

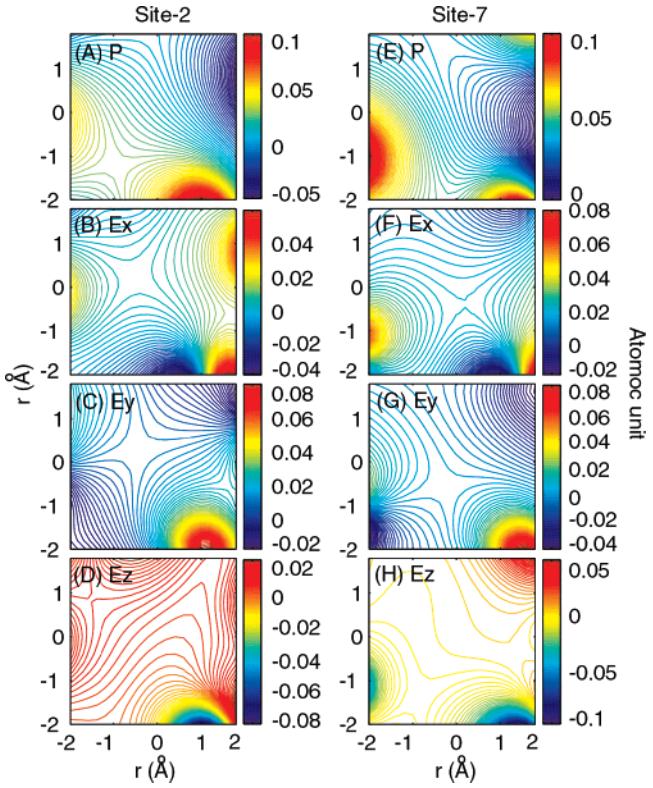

Figure 9. Electrostatic potential and electrostatic field components in atomic units due to peptide backbone, side chain and solvent, calculated for site 2 amide unit (left panels) and site 7 amide unit (right panels). The data are from a snap shot at 20 ps of a 1 ns trajectory. The origin was set to the midpoint of the C═O bond. (A and E) Potentials. (B and F) E_x component. (C and G) E_y component. (D and H) E_z component.

Electrostatic Interaction Pictures. Figure 9 shows the electrostatic potential and field components for site 2 and site 7 amide units for a specific snapshot of a 1 ns all-atom MD trajectory. The potential and fields are in atomic units. The origin was assumed to be the midpoint of the C═O bond in these panels. The electrostatic force fields are found to be significantly different for site 2 (left panels) and site 7 (right panels). By comparing panel A and E, it is observed that the strength of the potential ranges from −0.05 to +0.1 for site 2 but from 0 to +0.1 for site 7; therefore, on average, the potential felt by the amide-I mode is generally stronger for the latter site. For the electrostatic field components, it is found that the E_y component is similar for site 2 and site 7; however, the E_x and E_z components are both larger for site 7, indicating a stronger field in the neighborhood of site 7. These potentials or fields at the reference point(s) are the driving forces of the resultant local amide-I mode frequency difference between these two amide units. These results illustrate what have turned out to be large influences on the frequencies, the frequency dynamics, and line-shapes of the local vibrational modes and hence strongly predict that infrared spectroscopy of all types should reveal the site dependent solvent−solute electrostatic interactions in condensed phases. These effects apparently dominate the variations in amide vibrational modes arising from chemical differences between residues and small variations in backbone structure.

Conclusion

Two types of electrostatic maps have been used to calculate the Hamiltonians of the trpzip2 system: the HKLC[28] map based

on the electric potential and the HZM[31] map based on the electric field. According to comparisons with the experiments, both maps can capture the main experimental FTIR and 2D IR features. Isotope labeling of a residue at the β strand and turn segments generate bands with different frequencies, consistent with what is experimentally observed. Both maps are unsuccessful in fully reproducing the accurate peak positions and splitting of site 2 and site 7 isotope bands. This result suggests a possible new benchmark for future improvement of the electrostatic maps. Single ^{13}C or $^{13}C{=}^{18}O$ isotopic substitutions are proposed as probes of the electrostatic field heterogeneity experienced by different residues on a peptide chain for proteins and peptides in condensed phases. Although not reported here, the same isotope replacement approaches are expected to reveal the fluctuations in the electric fields at different sites directly from measurements of the spectral diffusion by 2D IR.

Acknowledgment. This research was supported by grants from NIH (GM12592 and RR01348) and NSF to R.M.H. This research was also supported by grants from NIH (GM59230-05) and NSF (CHE-0446555) to S.M. Computational support was mainly provided by the National Science Foundation CRIF Program, Grant CHE-013113.

References and Notes

(1) Gnanakaran, S.; Hochstrasser, R. M. *J. Am. Chem. Soc.* **2001**, *123*, 12886.
(2) Scheurer, C.; Piryatinski, A.; Mukamel, S. *J. Am. Chem. Soc.* **2001**, *123*, 3114.
(3) Decatur, S. M. *Acc. Chem. Res.* **2006**, *39*, 169.
(4) Walsh, S. T. R.; Cheng, R. P.; Wright, W. W.; Alonso, D. O. V.; Daggett, V.; Vanderkooi, J. M.; DeGrado, W. F. *Protein Sci.* **2003**, *12*, 520.
(5) Manas, E. S.; Getahun, Z.; Wright, W. W.; DeGrado, W. F.; Vanderkooi, J. M. *J. Am. Chem. Soc.* **2000**, *122*, 9883.
(6) Fang, C.; Hochstrasser, R. M. *J. Phys. Chem. B* **2005**, *109*, 18652.
(7) Wang, J.; Chen, J.; Hochstrasser, R. M. *J. Phys. Chem. B* **2006**, *110*, 7545.
(8) Mukherjee, P.; Kass, I.; Arkin, I.; Zanni, M. T. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 3528.
(9) Smith, A. W.; Tokmakoff, A. *J. Chem. Phys.* **2007**, *126*, 045109/1.
(10) Torii, H.; Tasumi, M. *J. Chem. Phys.* **1992**, *96*, 3379.
(11) Wang, J.; Hochstrasser, R. M. *Chem. Phys.* **2004**, *297*, 195.
(12) Hamm, P.; Lim, M.; Hochstrasser, R. M. *J. Phys. Chem. B* **1998**, *102*, 6123.
(13) Hamm, P.; Lim, M.; DeGrado, W. F.; Hochstrasser, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 2036.
(14) Asplund, M. C.; Zanni, M. T.; Hochstrasser, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8219.
(15) Zhuang, W.; Abramavicius, D.; Mukamel, S. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 18934.
(16) Lee, C.; Park, K.-H.; Kim, J.-A.; Hahn, S.; Cho, M. *J. Chem. Phys.* **2006**, *125*, 114510/1.
(17) Li, S.; Schmidt, J. R.; Piryatinski, A.; Lawrence, C. P.; Skinner, J. L. *J. Phys. Chem. B* **2006**, *110*, 18933.
(18) Kolano, C.; Helbing, J.; Kozinski, M.; Sander, W.; Hamm, P. *Nature (London)* **2006**, *444*, 469.
(19) Krummel, A. T.; Zanni, M. T. *J. Phys. Chem. B* **2006**, *110*, 13991.
(20) Zheng, J.; Fayer, M. D. *J. Am. Chem. Soc.* **2007**, *129*, 4328.
(21) Kurochkin, D. V.; Naraharisetty, S. R. G.; Rubtsov, I. V. *J. Phys. Chem. A* **2005**, *109*, 10799.
(22) Larsen, O. F. A.; Bodis, P.; Buma, W. J.; Hannam, J. S.; Leigh, D. A.; Woutersen, S. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13378.
(23) Cowan, M. L.; Bruner, B. D.; Huse, N.; Dwyer, J. R.; Chugh, B.; Nibbering, E. T. J.; Elsaesser, T.; Miller, R. J. D. *Nature (London)* **2005**, *434*, 199.
(24) Maekawa, H.; Toniolo, C.; Broxterman, Q. B.; Ge, N.-H. *J. Phys. Chem. B* **2007**, *111*, 3222.
(25) Krimm, S.; Bandekar, J. *Adv. Protein Chem.* **1986**, *38*, 181.
(26) Corcelli, S. A.; Lawrence, C. P.; Skinner, J. L. *J. Chem. Phys.* **2004**, *120*, 8107.
(27) Fecko, C. J.; Eaves, J. D.; Loparo, J. J.; Tokmakoff, A.; Geissler, P. L. *Science* **2003**, *301*, 1698.
(28) Ham, S.; Kim, J.-H.; Lee, H.; Cho, M. *J. Chem. Phys.* **2003**, *118*, 3491.
(29) Bour, P.; Keiderling, T. A. *J. Chem. Phys.* **2003**, *119*, 11253.
(30) Schmidt, J. R.; Corcelli, S. A.; Skinner, J. L. *J. Chem. Phys.* **2004**, *121*, 8887.
(31) Hayashi, T.; Zhuang, W.; Mukamel, S. *J. Phys. Chem. A* **2005**, *109*, 9747.
(32) Watson, T. M.; Hirst, J. D. *Mol. Phys.* **2005**, *103*, 1531.
(33) la Cour Jansen, T.; Knoester, J. *J. Chem. Phys.* **2006**, *124*, 044502/1.
(34) Choi, J.-H.; Ham, S.; Cho, M. *J. Phys. Chem. B* **2003**, *107*, 9132.
(35) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5578.
(36) Kim, Y. S.; Wang, J.; Hochstrasser, R. M. *J. Phys. Chem. B* **2005**, *109*, 7511.
(37) Hochstrasser, R. M. *Chem. Phys.* **2001**, *266*, 273.
(38) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Klaus, S. *J. Comput. Chem.* **2005**, *26*, 1781.
(39) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(40) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
(41) Zhuang, W.; Abramavicius, D.; Hayashi, T.; Mukamel, S. *J. Phys. Chem. B* **2006**, *110*, 3362.
(42) Torii, H.; Tasumi, M. *J. Raman Spectrosc.* **1998**, *29*, 81.
(43) Wang, J.; Hochstrasser, R. M. *Chem. Phys.* **2004**, *297*, 195.
(44) Gorbunov, R. D.; Kosov, D. S.; Stock, G. *J. Chem. Phys.* **2005**, *122*, 224904.
(45) Cheam, T. C.; Krimm, S. *Chem. Phys. Lett.* **1984**, *107*, 613.
(46) Hamm, P.; Lim, M.; DeGrado, W. F.; Hochstrasser, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 2036.
(47) Moran, A.; Mukamel, S. *PNAS* **2004**, *101*, 506.
(48) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Vreven, J. T.; K. N. Kudin; J. C. Burant; J. M. Millam; S. S. Iyengar; J. Tomasi; V. Barone; B. Mennucci; M. Cossi; G. Scalmani; N. Rega; G. A. Petersson; H. Nakatsuji; M. Hada; M. Ehara; K. Toyota; R. Fukuda; J. Hasegawa; M. Ishida; T. Nakajima; Y. Honda; O. Kitao; H. Nakai; M. Klene; X. Li; J. E. Knox; H. P. Hratchian; J. B. Cross; C. Adamo; J. Jaramillo; R. Gomperts; R. E. Stratmann; O. Yazyev; A. J. Austin; R. Cammi; C. Pomelli; J. W. Ochterski; P. Y. Ayala; K. Morokuma; G. A. Voth; P. Salvador; J. J. Dannenberg; V. G. Zakrzewski; S. Dapprich; A. D. Daniels; M. C. Strain; O. Farkas; D. K. Malick; A. D. Rabuck; K. Raghavachari; J. B. Foresman; J. V. Ortiz; Q. Cui; A. G. Baboul; S. Clifford; J. Cioslowski; B. B. Stefanov; G. Liu, A.; Liashenko; P. Piskorz; I. Komaromi; R. L. Martin; D. J. Fox; T. Keith; M. A. Al-Laham; C. Y. Peng; A. Nanayakkara; M. Challacombe; P. M. W. Gill; B. Johnson; W. Chen; M. W. Wong; C. Gonzalez; Pople, J. A., Eds. *Gaussian 03, Revision B.05*; Gaussian, Inc.: Pittsburgh, PA, 2003.
(49) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235.
(50) Fang, C.; Wang, J.; Kim, Y. S.; Charnley, A. K.; Barber-Armstrong, W.; Smith, A. B., III; Decatur, S. M.; Hochstrasser, R. M. *J. Phys. Chem. B* **2004**, *108*, 10415.