

Electrostatic DFT Map for the Complete Vibrational Amide Band of NMA

Tomoyuki Hayashi, Wei Zhuang, and Shaul Mukamel*

Department of Chemistry, University of California, Irvine, California 92697-2025

Received: May 4, 2005; In Final Form: August 16, 2005

An anharmonic vibrational Hamiltonian for the amide I, II, III, and A modes of *N*-methyl acetamide (NMA), recast in terms of the 19 components of an external electric field and its first and second derivative tensors (electrostatic DFT map), is calculated at the DFT(BPW91/6-31G(d,p)) level. Strong correlations are found between NMA geometry and the amide frequency fluctuations calculated using this Hamiltonian together with the fluctuating solvent electric field obtained from the MD simulations in TIP3 water. The amide I and A frequencies are strongly positively correlated with the C=O and N–H bond lengths. The C=O and C–N amide bond lengths are negatively correlated, suggesting the solvent-induced fluctuations of the contribution of zwitterionic resonance form. Sampling the global electric field in the entire region of the transition charge densities (TCDs) is required for accurate infrared line shape simulations. Collective electrostatic solvent coordinates which represent the fluctuations of the 10 lowest amide fundamental and overtone states are reported. Normal-mode analysis of an NMA–3H₂O cluster shows that the 660 cm⁻¹ to 1100 cm⁻¹ oscillation found in the frequency autocorrelation functions of the amide modes may be ascribed to the two bending vibrations of intermolecular hydrogen bonds with the amide oxygen of NMA.

I. Introduction

Ultrafast vibrational spectroscopy is a powerful tool for probing molecular structure and dynamics in the condensed phase.^{1–13} Vibrational frequencies are sensitive to the local solvent environment; the frequency shift with respect to the gas phase provides information about the average structure, whereas the line shapes depend on dynamics of fluctuations. The amide-I infrared absorption band of peptides which originates from the stretching motion of the C=O bond (coupled to N–H bending and C–H stretching) provides a useful indicator of secondary structure changes because of its sensitivity to hydrogen bonding, dipole–dipole interactions, and geometry of the peptide backbone. The anharmonicity and vibrational relaxation were probed in the first 2D IR measurements of peptides.¹⁴ Extensive studies were performed on a small and flexible peptide, trialanine, and yielded the Ramachandran angles in good agreement with NMR studies.^{8,9} For larger peptides the emphasis of 2D IR has been to relate certain patterns in the spectra to typical structural motifs as α -helix, 3_{10} -helix, and β -sheet.^{14–18} However, the cross-peak pattern of other amide bands can provide additional structural information.

A high-level vibrational Hamiltonian of *N*-methyl acetamide (NMA) which is the simplest model system for the peptide bond is crucial for constructing the Hamiltonian of larger proteins and simulating their vibrational dynamics. NMA is highly polarizable due to the contribution of the two resonance structures (Figure 1), and its structural changes in different solvent environments have been studied.^{19–21} The effects of hydrogen bonding with water on the structure were studied by ab initio geometry optimization of NMA–water clusters.¹⁹ H-bonding was found to increase the C=O bond length by 0.008–0.023 Å, indicating an increased contribution of the ionic resonance form. Similar results were found for liquid NMA.²⁰ Connecting the vibrations and the structure in solution is of great interest. Ab initio ground-state calculations have been compared with experimental infrared and Raman spectra in the gas phase and in water.^{22–27} MP2^{28–30} and several density functionals have been tested. The PW91³¹ and its hybrid functional with

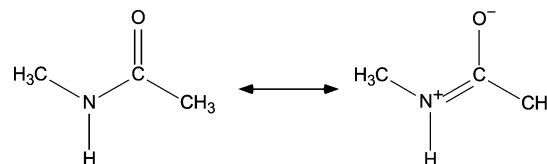


Figure 1. Resonance structures of NMA. Left: neutral; right: zwitterionic.

Becke88³² (BPW91) combined with 6-31G(d,p) and 6-31+G-(d) basis sets give better gas-phase amide harmonic frequencies of NMA and amide dimers than B3LYP^{33,34} and MP2.^{24,26} The PW91 functional provides good binding energies of hydrogen bond and van der Waals interactions in several dimers of small molecules (0.33 kcal/mol; exp 0.28 kcal/mol for Ar–Ar dimer), whereas B3LYP fails to have an energy minimum for the system.³⁵ Including three water molecules directly bonded to NMA explicitly, combined with the Onsager reaction field model^{36,37} was found crucial for reproducing the amide frequencies of NMA in water.³⁸ However it is still not clear whether these interactions are purely electrostatic or involve covalent bonding.

The amide I band is widely used as a marker for secondary structure, dynamics, and hydrogen bonding pattern.¹⁶ Electrostatic maps have been used to simulate the vibrational dephasing of amide I mode in NMA, reproducing both the peak shift and bandwidth.^{39,40} An empirical linear relation was assumed between the ab initio normal-mode frequency of NMA–water clusters and the electrostatic potentials at the 6 atom sites of NMA in the same clusters. The maps are limited to the fundamental frequency, and it is not clear whether they are generally transferable to other solvents. The amide II and III modes are known as marker bands and show a blueshift in polar solvents.³⁸ The diagonal and off-diagonal peaks of all amide modes can be used to characterize the structure of proteins. In linear spectra the amide A band in water is usually hidden underneath the broad O–H stretch band of water. However the amide A can be observed in 2D measurements as its off-diagonal peaks with other amide modes.

In this paper we construct a general ab initio electrostatic map that can simultaneously predict the fundamental and combination bands of several amide modes and is particularly suitable for the simulation of the nonlinear-infrared response. We have demonstrated this approach for the O–H stretch of HOD in D₂O.⁴¹ A fluctuating Hamiltonian was obtained by diagonalizing the anharmonic ab initio vibrational Hamiltonian of an isolated HOD molecule in a spatially nonuniform electrostatic potential expanded to second-order Cartesian coordinates at the center of charge.⁴¹ This approach has several notable advantages over the empirical ab initio maps which predict the line width with comparable accuracy. First, it shows that the electrostatic modeling can reproduce the experimental linear and third-order infrared spectra, suggesting that the electrostatic contributions are more important than covalent contributions to hydrogen bonding effects on high-frequency vibrational bands. Second, it includes geometry changes due to the solvent-induced polarization. The geometry of NMA is optimized in the presence of the electric field when the DFT map is created. Therefore the simulation includes geometry fluctuations in response to the solvent-induced electrostatic potential, which allow us to establish the connection between the frequency and geometry fluctuations. Third, the map incorporates higher (up to fourth order) derivatives of the electrostatic potential. Our calculations clearly demonstrate that the electric field gradient-tensor contributes significantly to the frequency fluctuations. Fourth, this map is not optimized to a particular solvent or solvent environment and is therefore transferable. Fifth, it provides the entire fluctuating Hamiltonian required for simulations of nonlinear spectra (fundamental and overtone frequency as well as the transition dipole magnitude and direction) from first principles. Collective solvent coordinates have been identified, offering a simple physical picture for the infrared bands and allowing high level line shape simulations using the stochastic Liouville equations.⁴²

The DFT electrostatic map was constructed for NMA at the BPW91/6-31G(d,p) level where the four amide (I, II, III, and A) fundamental frequencies and their combination and overtone frequencies are parametrized in terms of the electrostatic potential expanded to third order in Cartesian coordinate. The map is obtained by numerical derivatives of the vibrational eigenstates calculated for the sixth-order ab initio anharmonic vibrational potential with respect to the electric field, first and second electric field derivatives. The solute–solvent interactions are assumed to be purely electrostatic, and the map includes the electronic polarization of NMA and the solute–solvent dynamics. The complete fluctuating Hamiltonian required for the third-order vibrational response of NMA is calculated using the electric field and its first and second derivatives. These are obtained by least-squares fits to an electric field sampling grid from MD simulations. The linear-infrared spectra are calculated, including all 4 amide modes. A local sampling at 4 amide (C,O,N,H) atoms and 3 additional points (for out-of-plane gradients) fails to reproduce amide II and III solvent shifts, whereas a global sampling which spans the entire region of transition charge densities (TCDs) of amide modes predicts the amide I, III, and A peak shifts, in good agreement with experiment. The amide II shift is less accurate since its more delocalized TCD is not fully covered by our electrostatic grid. A collective solvent coordinate, a linear combination of 19 independent components of electric field and its derivatives, is constructed for each of the 10 amide states considered (the fundamentals, overtones, and combination bands).

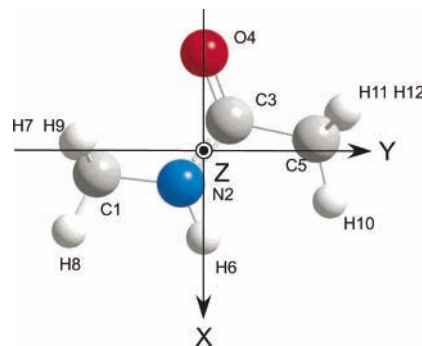


Figure 2. Atom labels and coordinate system used for the anharmonic force field and the electrostatic potential of NMA. Oxygen (red), carbon (gray), nitrogen (blue), and hydrogen (white). The four amide atoms (O4, C3, N2, and H6) are in the *x*-*y* plane. The origin is the middle point of the oxygen (O4) and hydrogen (H6).

Inhomogeneous electrostatic fluctuations for the amide modes obtained from MD simulations and the least-squares fit using the sampling point of large amide transition charge densities are shown in section II. Electrostatic DFT maps are constructed for the NMA structure, the 10 lowest amide frequencies (fundamental, overtone, and combination), and their transition dipole moments in section III. The infrared spectrum is simulated in section IV. An electrostatic collective coordinate for each of the 10 states is presented in section V. Finally, our results are discussed in section VI.

II. Inhomogeneous Electrostatic Fluctuations of the Amide Modes

The molecular charge density $\rho(\mathbf{r})$ can be expanded in the vibrational modes Q_i :⁴³

$$\rho(\mathbf{r}) = \rho^{(0)}(\mathbf{r}) + \sum_i \left(\frac{\partial \rho(\mathbf{r})}{\partial Q_i} \right) Q_i + \dots \quad (1)$$

The *transition charge density* (TCD) ($\partial \rho(\mathbf{r}) / \partial Q_i$) reflects the electronic structure change due to the Q_i vibration. We expect the electrostatic potential generated by the solvent in the TCD region to affect the optical response of that vibration. Atom labels and the molecular coordinate system are defined in Figure 2. Contour plots of the TCD for the various amide modes are presented in Figure 3. The amide I TCD is localized on the amide oxygen and carbon. For amide A it is localized on the N–H bond. For amide III it is delocalized over the four amide atoms (O4, C3, N2, and H6) and also the N2–C1 bond region, and for amide II it is most delocalized over the amide and methyl groups (C1, H7, H8, and H9). This suggests that while the electrostatic potential in the vicinity of four amide atoms (O4, C3, N2, and H6) should be most important to all amide modes, other areas such as the methyl group (C1, H7, H8, and H9) and the C3–H5 bond region are also relevant for amide II and III.

MD simulations were performed using the CHARMM force field for NMA and TIP3 water at a constant temperature 300 K⁴⁴ and a constant volume (NVT) after a constant temperature 300 K and pressure 1 bar⁴⁵ (NPT) equilibration. A 30.75 Å cubic simulation box of one NMA with 1000 water molecules was created. The NMA and water molecules were kept rigid using the SHAPE and SHAKE⁴⁶ algorithms, respectively. A 12 Å cutoff distance for both electrostatic and Lennard-Jones forces was used. Long-range electrostatic interactions were accounted for with the Particle-mesh Ewald (PME) summation.⁴⁷ The electric field and gradients were obtained from a 1 ns trajectory with a 1 fs time step.

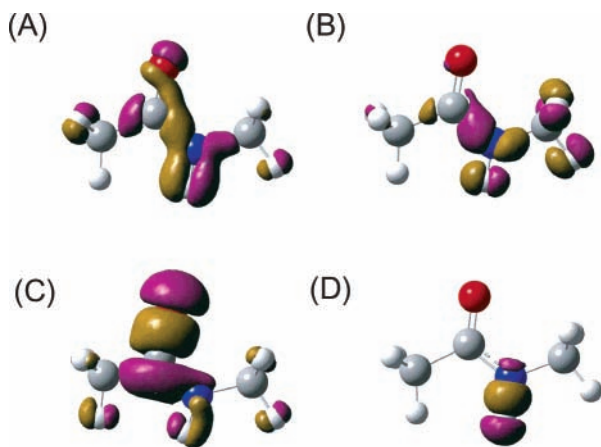


Figure 3. Transition charge densities (TCD) (eq 1) for the 4 amide modes of NMA. Shown is the 0.010 esu/Bohr contour of (A): amide III mode; (B): amide II mode; (C): amide III mode; (D): amide A mode. Violet and brown contours represent positive and negative values, respectively.

Contour plots of the average and variance of the electrostatic potentials and the electric fields (E_x and E_y) are shown in Figure 4. By symmetry, the average E_z vanishes in the x - y plane and is not shown. The average electrostatic potential has one maximum at $(-2, -1.3, 0)$ Å which corresponds to the H_2O hydrogen bound to the amide oxygen of NMA and a minimum at $(2, 1, 0)$ Å, corresponding to the oxygen of water bound to the amide hydrogen. The variances of electrostatic fields, E_x and E_y , are larger in the right and left side of the contour plots where the solvent is accessible. We defined the average correlation time of E_x and E_y :

$$t_\alpha(\mathbf{r}) = \int_0^\infty dt \langle E_\alpha(0; \mathbf{r}) E_\alpha(t; \mathbf{r}) \rangle / \langle E_\alpha(\mathbf{r})^2 \rangle, \alpha = x, y \quad (2)$$

Contour plots of t_α shown in Figure 4 indicate that in the region where the water hydrogen bonded to NMA is possible, the electric field varies more slowly due to the intermolecular vibrations and the exchange of water. However the fast time scales are not directly connected to intermolecular vibrations or exchange. It should be noted that the classical point charge approximation for the solvent overestimates the electric field in the area where the solvent is accessible due to neglecting the solvent charge distribution.

The solvent-induced electrostatic potential U was expanded to cubic order in Cartesian coordinates X_α ($\alpha = x, y, z$) around the midpoint between amide oxygen and hydrogen atoms of NMA (Figure 2)

$$U(\mathbf{X}) = U_0 - \sum_\alpha E_\alpha X_\alpha - \frac{1}{2} \sum_{\alpha\beta} E_{\alpha\beta} X_\alpha X_\beta - \frac{1}{6} \sum_{\alpha\beta\gamma} E_{\alpha\beta\gamma} X_\alpha X_\beta X_\gamma \quad (3)$$

Apart from the trivial factor U_0 , eq 3 has 19 independent parameters (note that $E_{\alpha\beta} = E_{\beta\alpha}$) which were arranged in a vector $\mathbf{C} = (E_x, E_y, E_z, E_{xx}, E_{yy}, E_{zz}, E_{xy}, E_{xz}, E_{yz}, E_{xxx}, E_{yyy}, E_{zzz}, E_{xyy}, E_{xxy}, E_{xxz}, E_{yzz}, E_{yyz}, E_{xyz}, E_{xyz})$. \mathbf{C} was obtained by a least-squares-fit of the actual electric field data obtained from the MD configurations. We used 67 sampling points consisting of two-dimensional 49 (7×7) grid points in the x - y plane within a rectangular area defined as $-1.4 < x < 1.4$ and $-0.8 < y < 0.8$ in Å, 6 additional points outside of the rectangular area in the x - y plane, and 10 points located at $(\pm 0.5, 0, 0)$, $(\pm 0.5, 0, \pm 0.5)$, and $(\pm 1, 0, 0)$, $(\pm 1, 0, \pm 0.5)$ in Å. This sampling, denoted S1, covers most of the areas of large TCD (larger than 0.01

esu/Bohr) for amide I, III, and A. For comparison, we also used a 7-point sampling (S2), 5 in the x - y plane at the positions of 4 amide atoms and the origin of the coordinate system and 2 out-of-plane located at $(x, y, z) = (0, 0, 0.2)$ and $(0, 0, -0.2)$. S2 was only used as a reference in a few calculations indicated below; all other calculations are based on S1.

Contour plots of the average E_x and E_y in the vicinity of 4 amide atoms directly calculated from MD simulation together with fits are shown in Figure 5. The root-mean-square-displacement (RMSD) between the MD simulation and least-squares fit was calculated as well and displayed in Figure 15 in the Supporting Information

$$E_\alpha^{\text{RMSD}}(\mathbf{r}) = \frac{1}{T} \int_0^T dt (E_\alpha^{\text{MD}}(\mathbf{r}, t) - E_\alpha^{\text{LSQ}}(\mathbf{r}, t))^2, \alpha = x, y \quad (4)$$

where $E_\alpha^{\text{MD}}(\mathbf{r}, t)$ and $E_\alpha^{\text{LSQ}}(\mathbf{r}, t)$ represent the electric field in $\alpha = x, y$ direction at time t , and position \mathbf{r} is calculated from MD simulation and from the least-squares fit. $T = 2$ ns is the simulation time. S2 gives a better field (smaller RMSD) at the 4 amide atoms. However E_y has a saddle point near the amide nitrogen, whereas S2 has a minimum. S1 has a larger RMSD at the 4 amide atoms but better reproduces the global electric fields E_x and E_y .

III. The DFT Electrostatic Map

Geometry optimization and normal mode calculations of *trans*-NMA with C_s symmetry in the gas phase³⁸ were performed using Gaussian 03 at the BPW91/6-31G(d,p) level.⁴⁸ The BPW91 functional with 6-31G(d) basis set is known to reproduce experimental amide I, II, and III frequencies.³⁸ We used the 6-31G(d,p) basis set with extra p-functions on the hydrogen atoms. The resulting anharmonic vibrational potential $V(\mathbf{Q}; \mathbf{C}(t))$ was expanded in 5 normal coordinates involving the amide C=O stretch, N-H bending, and N-H stretch (Q_1 : mixture of N-H bending and methyl deformation, Q_2 : amide III mode, Q_3 : amide II mode, Q_4 : amide I mode, and Q_5 : amide A mode). The harmonic frequencies are listed in Table 1. $V(\mathbf{Q}; \mathbf{C}(t))$ fluctuates with time, following the solvent-induced electrostatic potential $U(\mathbf{X})$ parametrized by $\mathbf{C}(t)$. The potential energy of NMA was expanded to sixth order in the 5 normal coordinates

$$V(\mathbf{Q}; \mathbf{C}) = \sum_i f_i^{(1)}(\mathbf{C}) Q_i + \sum_{ij} f_{ij}^{(2)}(\mathbf{C}) Q_i Q_j + \sum_{ijk} f_{ijk}^{(3)}(\mathbf{C}) Q_i Q_j Q_k + \sum_{ijkl} f_{ijkl}^{(4)}(\mathbf{C}) Q_i Q_j Q_k Q_l + \sum_{ijklm} f_{ijklm}^{(5)}(\mathbf{C}) Q_i Q_j Q_k Q_l Q_m + \sum_{ijklmnp} f_{ijklmnp}^{(6)}(\mathbf{C}) Q_i Q_j Q_k Q_l Q_m Q_n \quad (5)$$

where $f_{i_1 \dots i_n}^{(n)} \equiv (\partial^n V / \partial Q_{i_1} \dots \partial Q_{i_n}) / n!$. The geometry was optimized with respect to these 5 normal coordinates, holding all other degrees of freedom frozen to vanish the linear coefficients $f_i^{(1)}(\mathbf{C})$ in the presence of nonuniform external electric fields. The anharmonic force constants were then calculated by numerical derivatives of the analytical energy gradients in the same external fields, using our modified Gaussian 03 code.⁴¹

We have parametrized the internal coordinates at the minimal energy geometry with \mathbf{C}

$$R^a = R_{\text{gas}}^a + \mathbf{R}^{a(1)\dagger} \mathbf{C} + \frac{1}{2} \mathbf{C}^\dagger \mathbf{R}^{a(2)} \mathbf{C} \quad (6)$$

where R^a is either a bond length, a bond angle, or a dihedral

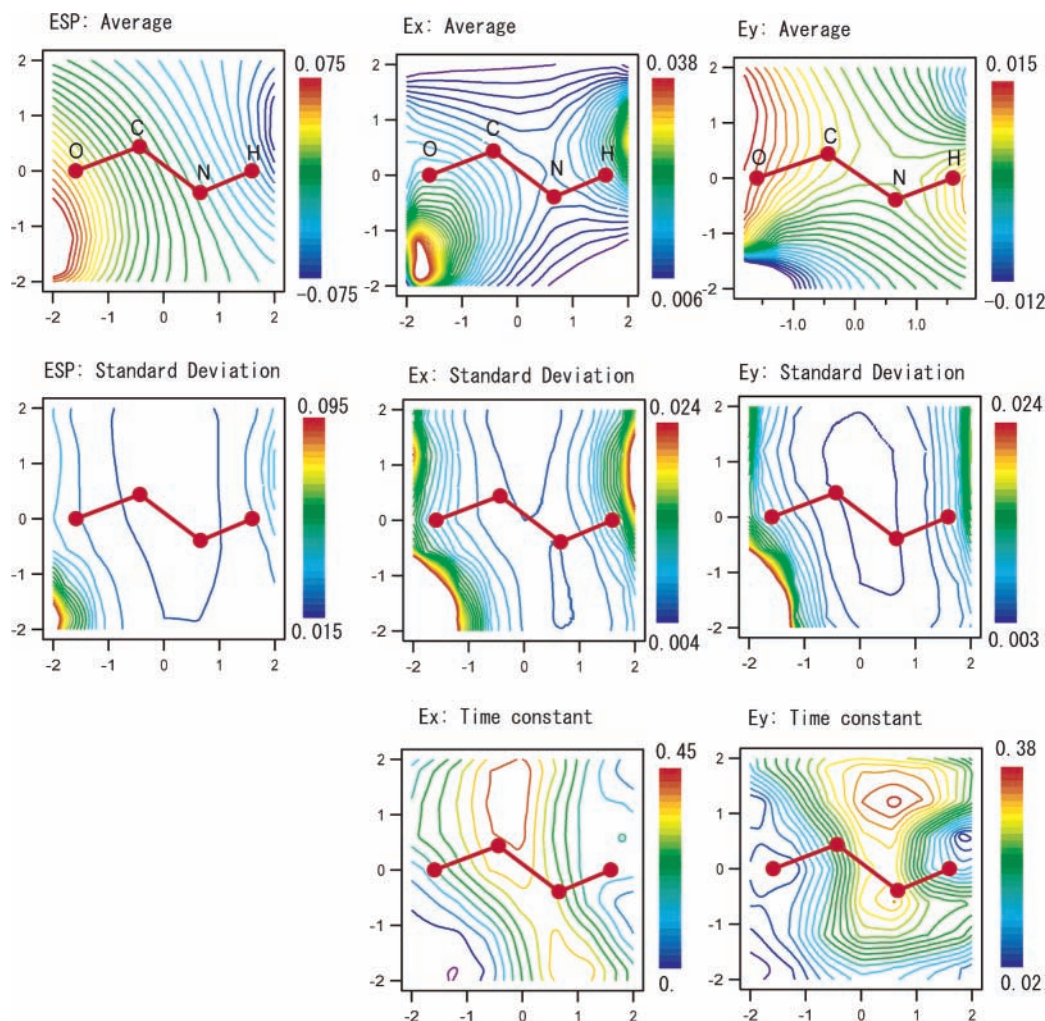


Figure 4. Top row: the average electrostatic potential (ESP) in atomic unit and the electric field components (E_x and E_y) in atomic units obtained from the MD simulation of NMA in H_2O . Red circles represent the four amide atoms (O4, C3, N2, and H6). Middle row: standard deviation of the same quantities using the same units as the top panels. Bottom row: The correlation time scale (eq 2) for E_x and E_y (in ps).

angle, $\mathbf{R}^{(1)}$ is a 19 component row vector

$$(\mathbf{R}^{a(1)\dagger})_i \equiv P_i \frac{\partial R^a}{\partial E_i} \quad (7)$$

and $\mathbf{R}^{(2)}$ is a 19×19 matrix

$$(\mathbf{R}^{a(2)})_{ij} \equiv P_i P_j \left(\frac{\partial^2 R^a}{\partial E_i \partial E_j} \right) \quad (8)$$

P_i represents the number of possible permutation of the indices of E_i , e.g. P_i is 2 for $E_i = E_{xy}$. The maps ($\mathbf{R}^{a(1)}$) of the C=O, C–N, and N–H bond lengths are given in Table 8 of the Supporting Information.

Fluctuations of the structure were calculated using eq 6. The static distributions of the C=O, C–N, and N–H bond lengths and the C–N–H angle are shown in Figure 6. The statistics of these internal coordinates are summarized in Table 2. The distribution of the N–H bond length is asymmetric with a tail for a longer bond length. Both average C=O and N–H bond lengths in water are significantly longer than in the gas phase (0.014 and 0.005 Å, respectively), and the C–N stretch in water is shorter by 0.007 Å. This is in qualitative agreement with the earlier study of the Hartree–Fock level calculations of NMA+2- (or 3)H₂O¹⁹ (C=O: 0.0122–0.0227 Å, C–N: –0.0158 ~ –0.0299 Å, and N–H: 0.0080–0.0092 Å depending on cluster configurations). The C–N–H angle has the scatter plots of C=

O versus C–N, and the N–H bond lengths are shown in Figure 7. The C=O and C–N bond lengths are negatively correlated, suggesting that the relative contribution of the two resonance structures $\text{N}2\text{--C}3\text{=O}4$ and $\text{N}2^+\text{=C}3\text{--O}4^-$ change depending on the solvent environment (Figure 1). The C=O and N–H bond lengths have a positive correlation. This can be rationalized since hydrogen bonding to either O4 or H6 causes a positive electric field E_x making both C=O and N–H bond lengths longer due to the larger molecular dipole–electric field interactions.

The vibrational Hamiltonian was recast in a normally ordered form (Appendix C of ref 49), and the eigenstates were obtained by diagonalizing the Hamiltonian using the Implicit Restarted Arnoldi Method (IRAM)^{49–51} using a Hartree product basis set made of 12 harmonic oscillator eigenfunctions per one normal mode. Basis states where the total number of excitations $n_T \equiv n_1 + n_2 + n_3$ are larger than 12 are neglected. The basis set size was 6183. The gas-phase vibrational eigenstates (eigenfrequencies and eigenvectors) are tabulated in Table 9 of the Supporting Information.

We focused on the 10 eigenstates which dominate the linear and the third-order infrared spectra: the 4 amide I, II, III, and A fundamentals which have large transition dipole moments to the ground state, the 3 first overtone of amide I, II, and III modes, and the 3 combination bands. The 6 overtone and combination states have large transition dipole moments to the

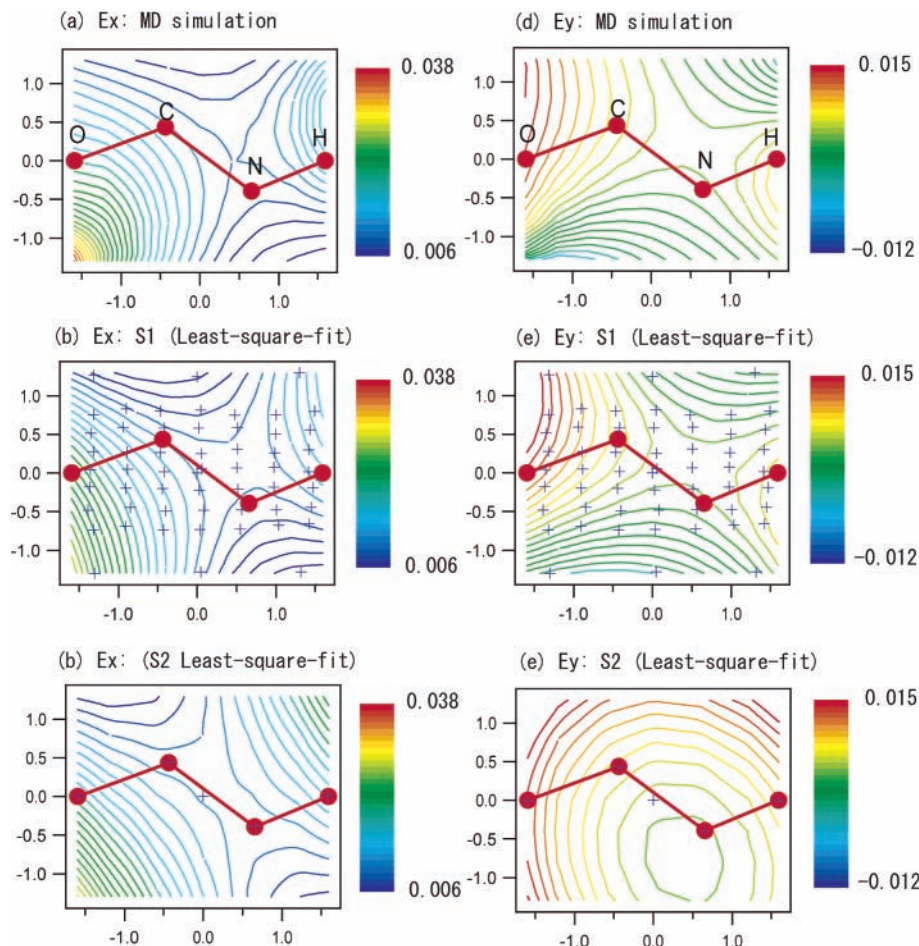


Figure 5. Left column: top – the average electric field E_x of NMA in H_2O from the MD simulation; middle – the least-squares-fit to the MD simulation data using sampling 1 (S1); bottom – the least-squares-fit to the MD simulation data using sampling 2 (S2); right column: same quantities for E_y . Red circles represent the four amide atoms (O4, C3, N2, and H6). The sampling points (S1) are shown by the blue crosses in panels (b) and (e).

TABLE 1: Calculated Gas-Phase Normal Modes and Harmonic Frequencies Used To Calculate the Vibrational Hamiltonian of NMA

mode no.	harmonic frequency	description
1	1081.6	N–H bend + methyl def
2	1241.2	amide III
3	1506.6	amide II
4	1732.8	amide I
5	3582.8	amide A

3 amide I, II, and III fundamentals. The transition moments between the other states are 1–2 orders of magnitude smaller. The frequencies and transition dipole moments of these 10 states are summarized in Table 3. The amide I, II, and III frequencies are in good agreement with experiment (without scaling) to within 3–27 cm^{-1} . The first overtone of amide I mode is at 3436 cm^{-1} (experiment 3440 cm^{-1}) and its anharmonicity is 10.6 cm^{-1} (experiment 16 cm^{-1}). Compared with Table 1, the fundamental frequencies are around 10 cm^{-1} lower than their harmonic frequencies (9 cm^{-1} for amide I).

The vibrational transition frequency from the ground state to state ν and the transition dipole moments between states ν and ν' were parametrized using the electrostatic vector \mathbf{C}

$$\omega_{\text{map}}^{\nu} = \omega_{\text{gas}}^{\nu} + \mathbf{O}^{\nu(1)\dagger} \mathbf{C} + \frac{1}{2} \mathbf{C}^{\dagger} \mathbf{O}^{\nu(2)} \mathbf{C} \quad (9)$$

$$\mu_{\alpha}^{\nu\nu'} = \mu_{\text{gas}\alpha}^{\nu\nu'} + \mathbf{M}_{\alpha}^{\nu\nu'(1)\dagger} \mathbf{C} + \frac{1}{2} \mathbf{C}^{\dagger} \mathbf{M}_{\alpha}^{\nu\nu'(2)} \mathbf{C} \quad (\alpha = x, y, z) \quad (10)$$

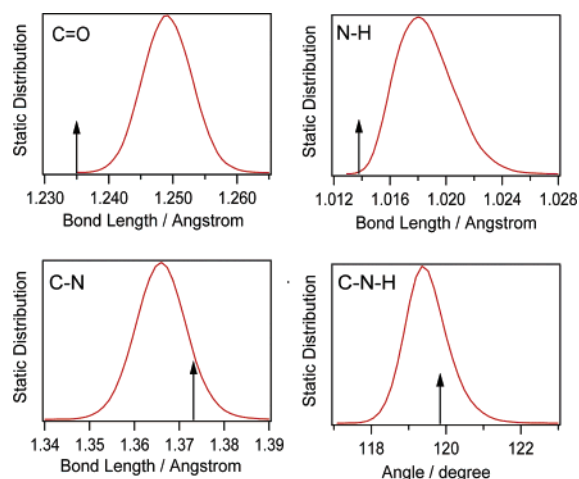


Figure 6. Static distribution of the equilibrium bond lengths of C3=O4, C3–N2 and N2–H6, and the equilibrium C3–N2–H6 angle. The arrows represent the gas-phase bond lengths and the C–N–H angle.

where $\mathbf{O}^{(1)}$ and $\mathbf{M}^{\nu\nu'(1)}$ are 19 component row vector and $\mathbf{O}^{(2)}$ and $\mathbf{M}^{\nu\nu'(2)}$ are 19×19 matrix defined in the same way as eqs 7 and 8. The gas-phase frequencies $\omega_{\text{gas}}^{\nu}$ were taken from experiment.⁵² The linear terms of the DFT map, $\mathbf{O}^{(1)}$ and $\mathbf{M}^{(1)}$, are given in Tables 10 and 11 of the Supporting Information, respectively.

The amide frequency fluctuations are calculated using eq 9 with the S1 sampling. The static distributions of the four

TABLE 2: Statistics of the Internal Coordinates (C=O, C–N, and N–H Bond Lengths and C–N–H Angle)^a

	gas	in water	
		average	variance
C=O	1.2350	1.2495	0.0038
C–N	1.3732	1.3667	0.0058
N–H	1.0138	1.0187	0.0020
C–N–H	119.84	119.59	0.5860

^a Bond lengths are in Å and C–N–H angle is in degrees. The average and variance of the internal coordinates of NMA in water are obtained from eq 6 and a MD trajectory. The gas-phase values are from the optimized structure of an isolated NMA with no external field at the same computational level (BPW91/6-31G(d,p)).

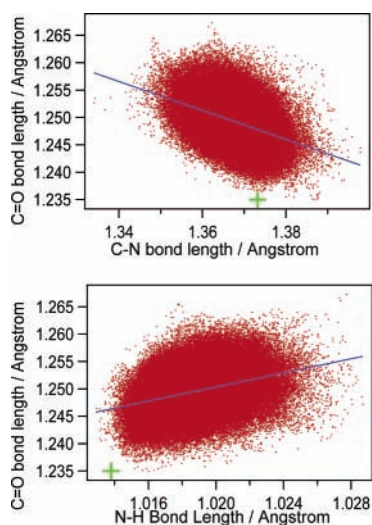


Figure 7. Scatter plot of C3=O4 versus the C3–N2 and N2–H6 bond lengths of NMA in water. The best optimal fits (blue lines) are $R_{CO} = 1.6106 - 0.26419R_{C-N}$ and $R_{CO} = 0.59198 + 0.64543R_{N-H}$. The gas-phase values are marked by green crosses.

fundamental frequencies are shown in Figure 8. The amide A distribution is asymmetric with a long red tail. The other modes have more symmetric, Gaussian profiles. Scatter plots of the amide I and A fundamental frequencies versus C=O and N–H bond lengths are displayed in Figure 9. The strong correlations suggest the importance of structural changes of NMA caused by the solvent-induced polarization. Optimal linear fits are $\omega = 6549.2 - 3905R_{CO}$ (amide I vs the C=O bond length) and $\omega = 24259 - 20426R_{NH}$ (amide A vs the N–H bond length). The scatter plots of the amide II and III fundamental frequencies

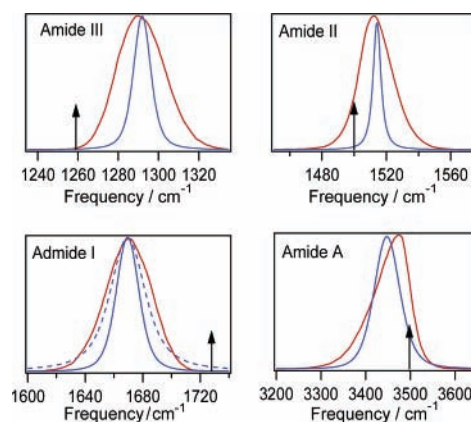


Figure 8. Red: static distributions of the amide I, II, III, and A fundamental frequencies obtained from the electrostatic map, blue solid line: calculated infrared line shape. For amide I, we also show the infrared line shapes convoluted with a Lorentzian of 11.8 cm⁻¹ fwhm corresponding to the experimental vibrational relaxation rate¹⁴ (blue dashed line). Arrows represent the gas-phase frequencies.

versus the N–H bond length are also displayed in Figure 9. The two bending frequencies are positively correlated with the N–H bond length, which is ascribed to the fact that the hydrogen bonding to H6 atom causes the longer N–H bond length and making the potential energy surface along the bending modes (amide II and III) deeper by stabilizing more the parallel N2–H6···OH₂ structure. Optimal linear fits are $\omega = -3066 + 4278R_{NH}$ (amide III vs the N–H bond length) and $\omega = -2768 + 4204R_{NH}$ (amide II vs the N–H bond length).

The correlations between the internal coordinates and the amide fundamental frequencies are shown in Table 7. The correlation is defined as

$$r_{vv'} = \langle X^v X^{v'} \rangle / \sqrt{\langle X^{v2} \rangle \langle X^{v'2} \rangle}, \quad (X^v = R^v, \omega^v) \quad (11)$$

r_{ij} varies between 1 (full correlation), 0 (no correlation), or -1 (full anticorrelation). The C–N–H angle has the strongest positive correlation with the C–N bond length (0.328), suggesting that the longer C–N bond length causes the smaller C–N–H angles. The two (amide II and III) bending modes also have significant correlations with the C–N–H angle (II: -0.280 and III: 0.326); however, the amide A has almost no correlation. This might be because the change in the minimum of the C–N–H angle shifts the bending frequencies due to the anharmonicity of their potential.

TABLE 3: Frequencies (in cm⁻¹) and Transition Dipole Moments (in Debye) of the 10 States of the 4 Amide Modes of NMA in the Gas Phase^c

state		frequency		transition dipole moment			
index	mode	calc	exp	0	1	2	3
0		0.0					
1	III	1229.2	(1258 ^a , 1255 ^b)	4.6 (–1.1, 4.5)			
2	II	1495.4	(1500 ^a , 1499 ^b)	7.9 (4.0, –6.9)	0.3 (0.2, –0.2)		
3	I	1723.9	(1728 ^a)	9.0 (9.0, –0.5)	0.7 (–0.6, 0.4)	0.5 (0.4, –0.3)	
4	A	3428.1	(3498 ^a)	2.4 (2.4, 0.0)	0.2 (0.1, –0.2)	0.7 (0.2, –0.7)	0.7 (–0.7, –0.1)
5	III × 2	2451.5	(2504 ^a)	0.0 (0.0, –0.0)	6.4 (–1.5, 6.2)	0.0 (0.0, 0.0)	0.1 (–0.1, 0.0)
6	II × 2	2981.1		0.2 (0.2, 0.1)	0.1 (0.0, –0.1)	11.3 (–5.6, 9.8)	0.1 (0.0, 0.1)
7	I × 2	3436.8	(3440 ^a)	0.5 (–0.5, –0.1)	0.0 (–0.0, –0.0)	0.1 (0.0, 0.1)	12.8 (–12.7, –0.7)
8	III+II	2716.6	(2758 ^a)	0.4 (0.4, –0.0)	8.0 (–4.0, 7.0)	4.3 (1.0, –4.2)	0.1 (–0.1, 0.0)
9	III+I	2955.2	(2971 ^a)	0.2 (–0.2, 0.1)	9.0 (9.0, 0.5)	0.1 (0.0, 0.1)	4.6 (1.1, –4.5)
10	II+I	3222.2		0.2 (–0.2, 0.1)	0.0 (0.0, 0.0)	9.0 (–9.0, –0.5)	7.9 (4.0, –6.9)

^a Experimental frequencies are given in parentheses: from ref 52. ^b Experimental frequencies are given in parentheses: from ref 38. ^c One number and the three numbers in the parentheses at column i and row index j represent the magnitude $|\mu|$ and the two components (μ_x, μ_y) of the transition dipole moment between states i and j . All z components μ_z are zero. Transition dipole moments between the 2 exciton states are very small (not shown).

TABLE 4: Correlation between the Internal Coordinates and the Amide Frequencies (Eq 11)

	internal coordinate				amide fundamental frequency			
	C=O	C-N	N-H	C-N-H	III	II	I	A
C=O	1.0000	-0.4063	0.3408	-0.1241	0.0823	-0.1630	-0.6728	0.0977
C-N		1.0000	-0.3034	-0.3276	-0.1341	0.0368	0.3422	-0.0916
N-H			1.0000	-0.0282	0.2795	0.1921	-0.4367	-0.7110
C-N-H				1.0000	0.3262	-0.2798	-0.2197	-0.0705

IV. The Infrared Absorption

The linear response was calculated using the cumulant expansion of Gaussian fluctuations (CGF).^{53,54} The quantum frequency correlation functions of the four amide fundamentals calculated using the Wiener-Khinchin relation⁵⁵ with the harmonic quantum correction⁵⁶ are displayed in Figure 10. Amide II has the fastest decay. All four modes have an oscillatory behavior between 0 and 100 fs. The real part of the correlation function was fitted to a three exponentials overdamped and two damped oscillating functions corresponding to a high-frequency underdamped Brownian oscillator in the high-temperature limit (eq 8.69 in ref 56) (Table 5):

$$\langle \delta\omega(t)\delta\omega(0) \rangle / \langle \delta\omega^2 \rangle = \sum_{i=1}^3 A_i \exp(-t/T_i) + \sum_{i=4}^5 A_i [\cos(\Omega_i t) + 1/(T_i \Omega_i) \sin(\Omega_i t)] \exp(-t/T_i) \quad (12)$$

The fastest exponential decay of the 4 modes is between 16 and 30 fs, which is much faster than that found in water (HOD in D₂O).⁴¹ This suggests that hydrogen bonding between NMA and water is less stable than in pure water. The amide I, III, and A modes have a slow exponential 1.2–1.9 ps decay which can be ascribed to a global solvent fluctuation. The oscillation frequencies of four modes Ω and Ω' are between 660 cm⁻¹ and 1100 cm⁻¹. To trace their origin we further performed a normal mode calculation for the NMA-3H₂O cluster in essentially the same geometry as reported in ref 38 at the same level (BPW91/6-31G(d,p)) and found the 2 bending frequencies 674 and 726 cm⁻¹ of hydrogen bondings between 2 H₂O molecules and an amide oxygen of NMA. The bending motions

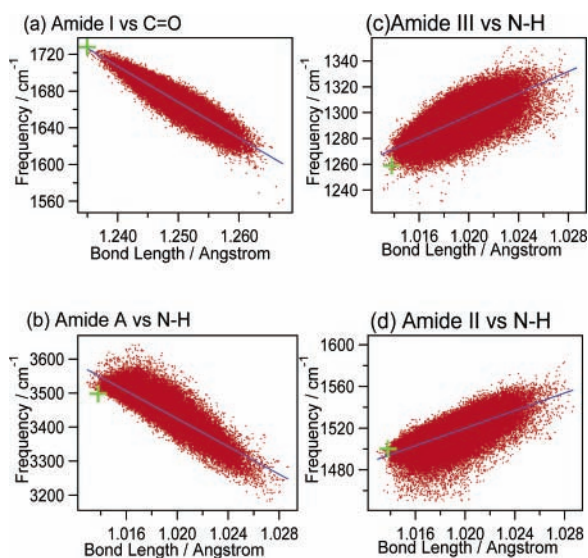


Figure 9. Scatter plots of amide frequencies versus bond lengths. Linear fits (blue lines) are $\omega = 6549.2 - 3905R_{CO}$ (amide I vs the C=O bond length), $\omega = 24259 - 20426R_{NH}$ (amide A vs the N-H bond length), $\omega = -3066 + 4278R_{NH}$ (amide III vs the N-H bond length), and $\omega = -2768 + 4204R_{NH}$ (amide II vs the N-H bond length). The gas-phase values are marked by green crosses.

TABLE 5: Fit Parameters (Eq 12), Variance, Correlation Time Constant τ , and Line Broadening Parameter κ for the Frequency Correlation Functions Shown in Figure 10^a

	amide III	amide II	amide I	amide A
A_1/cm^{-2}	0.481	0.584	0.503	0.521
T_1/fs	17.65	30.87	16.36	27.79
A_2/cm^{-2}	0.273	0.172	0.273	0.367
T_2/fs	155.7	297.7	210.6	250.2
A_3/cm^{-2}	0.122		0.151	0.038
T_3/fs	1888		1820	4277
A_4/cm^{-2}	0.170	0.250	0.218	0.081
Ω_4/fs^{-1}	0.183	0.187	0.213	0.158
T_4/fs	22.64	17.44	21.29	32.92
A_5/cm^{-2}	0.040	0.114		0.054
Ω_5/fs^{-1}	0.224	0.125		0.094
T_5/fs	59.48	38.96		37.04
$\langle \delta\omega^2 \rangle$	140.5	146.6	266.3	2153.6
τ/fs	275.3	101.5	431.9	186.6
κ	1.62	4.32	0.75	0.61

^a Numbers in parentheses represent the oscillation frequencies in cm⁻¹.

change the distance between the hydrogens of water and carbonyl oxygen thus affecting the electrostatic potential. The oscillation in the frequency correlations can be attributed to the two intermolecular hydrogen bond bending vibrations. The oscillation of the frequency autocorrelation of amide I was not observed clearly in an earlier study.³⁹ The variance $\langle \delta\omega^2 \rangle$, the correlation time constant $\tau \equiv \int_0^\infty dt \langle \delta\omega(0)\delta\omega(t) \rangle / \langle \delta\omega^2 \rangle$, and the motional narrowing parameter⁵⁶ $\kappa \equiv (\tau \sqrt{\delta\omega^2})^{-1}$ are shown in Table 5 as well. Amide I has the longest correlation time constant 432 fs, and the amide II has the shortest 101.5 fs, suggesting that the amide II bending is modulated mainly by the local solvent fluctuation whereas the amide I is more sensitive to the global solvent dynamics. The variance is smaller for amide II and III compared to amide I and A. κ is largest for amide II (4.32) and smallest for amide A (0.61). The static distributions of the absolute transition dipole moments between the ground state to amide III, II, I, and A fundamental states are shown in Figure 11. All distributions are nearly Gaussian. The transition dipole moments for the two bending (amide III and II) modes in water are similar to the gas phase; however, the transition dipole moments for the two stretching modes (amide I and A) are significantly enhanced in solution. The amide A modes changes from 4.92 D (gas phase) to 11.94 D (average in water).

The average transition dipole moments of the NMA/water trajectory were used in the calculation of infrared peak intensities. The infrared band shape for each mode is displayed together with the static frequency distribution in Figure 8. The line shapes are intermediate between Gaussian and Lorentzian ($\kappa = 0.61 \sim 4.32$) and narrower than the static distributions (motional narrowing). The calculated solvent shift and line width of the infrared band obtained from our two electric field samplings are given in Table 6 together with data from the empirical ab initio map.³⁹ The experimental infrared peaks are obtained by the multipole Lorentzian fit (Figure 16 in the Supporting Information) to the infrared spectra (Figure 1 in ref 38). Two

TABLE 6: Infrared Peak Positions in Gas Phase, Solvent Peak Shift, fwhm, and Intensity for 4 Amide Modes of NMA in H₂O Solution^g

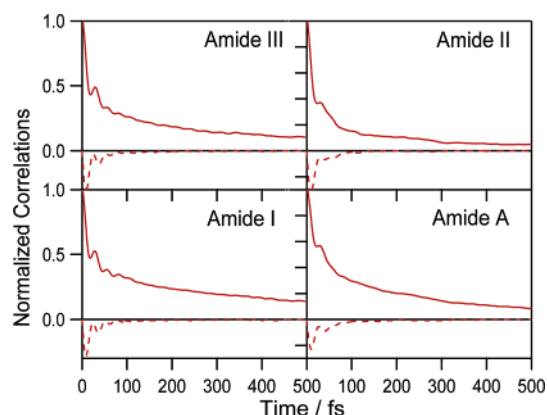
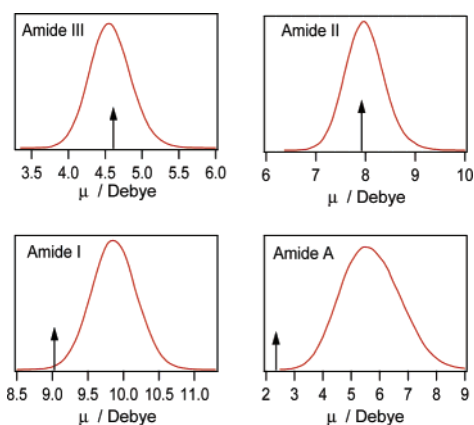
	amide I				amide II				amide III				amide A			
	gas	shift	fwhm	int	gas	shift	fwhm	int	gas	shift	fwhm	int	gas	shift	fwhm	int
S1	1724	-59	29	1.0	1497	+14	19	0.656	1232	+33	22	0.215	3428	-51	77	0.338
S2	1724	-57	28	1.0	1497	-16	37		1232	+10	33		3428	-72	116	
EM ^a	-	-78	25													
NM ^b	1738	-102			1533	+47			1240	+92						
exp.	1728 ^c	-87 ^{c,d}	29	1.0	1500 ^c	+76 ^{c,d}	36	0.803	1259 ^c	+53 ^{c,d}	32	0.205	3498 ^e	-78 ^f		

^a Empirical ab initio map based on the normal mode ab initio calculations of several different NMA in H₂O clusters.³⁹ ^b Normal mode ab initio calculations of the NMA + 3H₂O cluster in dielectric medium.³⁸ ^c From ref 52. ^d From ref 23. Experimental peak shift, fwhm, and intensity of amide III, II, and I modes are obtained from the multiplex Lorentzian fit to the infrared spectra³⁸ shown in Figure 16 in the Supporting Information. Intensities are normalized with the frequency integrated intensity of amide I band. ^e Experimental amide A frequency shift is from ref 27. ^f The columns represent the gas-phase frequency (gas), solvent peak shift (shift), fwhm (fwhm), and intensity (int) of the band in solution.

TABLE 7: Correlation Coefficients of Pairs of Frequencies and Anharmonicities (Eq 13)

	ω^1	ω^2	ω^3	ω^4	Δ^5	Δ^6	Δ^7	Δ^8	Δ^9	Δ^{10}	
ω^1	(III)	1.00	0.60	-0.71	-0.53	-0.30	0.65	-0.10	0.84	0.63	0.11
ω^2	(II)		1.00	-0.16	-0.69	0.19	0.35	-0.33	0.38	0.09	-0.29
ω^3	(I)			1.00	0.15	0.22	-0.51	0.06	-0.74	-0.46	-0.07
ω^4	(A)				1.00	-0.04	-0.15	0.59	-0.24	0.10	0.52
Δ^5	(III × 2)					1.00	0.03	0.07	-0.06	-0.16	-0.23
Δ^6	(II × 2)						1.00	0.23	0.73	0.64	0.27
Δ^7	(I × 2)							1.00	0.08	0.61	0.75
Δ^8	(III+II)								1.00	0.79	0.32
Δ^9	(III+I)									1.00	0.77
Δ^{10}	(II+I)										1.00

peaks are assigned to the broad 1628 cm⁻¹ line (1635 cm⁻¹ amide I and 1619 cm⁻¹ H-O-H bending of water²³). Two sharp peaks around 1400 cm⁻¹ were assigned to the CCH₃

**Figure 10.** Normalized correlation functions of the 4 amide fundamental states. Solid lines – real part, dashed lines – imaginary part.**Figure 11.** Static distribution of the transition dipole moments for the fundamental transitions. Arrows represent the gas-phase transition dipole. Average dipoles are 4.58 (amide III), 8.01 (amide II), 9.89 (amide I), and 5.74 D (amide A). Variances are 0.282 (amide III), 0.410 (amide II), 0.315 (amide I), and 1.012 (amide A).

symmetric bend (1377 cm⁻¹) and the NCH₃ symmetric bend (1416 cm⁻¹)²³ which are not included in our calculation. Since the amide A mode is hidden underneath the broad O-H stretch band of water, different experimental values of solvent shift have been reported from Raman spectra (-184 cm⁻¹)⁵⁷ and the infrared spectra of NMA in the D₂O/HOD mixture (-78 cm⁻¹).²⁷ Both S1 and S2 predict amide I peak shifts (-59 and -57 cm⁻¹) and line widths (29 and 28 cm⁻¹) in good agreement with experiment -80 cm⁻¹ and 29 cm⁻¹. Both samplings also give the similar amide A frequency shifts (S1: -51, S2: -72 cm⁻¹). However S2 fails to reproduce the experimental amide II peak shift (calc: +10 cm⁻¹; exp: +53 cm⁻¹), and only S1 gives a good amide III peak shift (+33 cm⁻¹). Note that S2 reproduces the electric field at amide 4 atoms better than S1, whereas S1 better reproduces the global electric field structure. This suggests that the amide I solvent peak shift only depends on the local electric field structure around the 4 amide atoms, but the amide III is dominated by the global electric field structure. This is consistent with the fact that the TCD of amide I (Figure 3) is localized on the C=O bond region, but for amide III it is more delocalized over the amide group and their bonds to carbons in the 2 methyl groups. The infrared band is thus determined by the global electric field structure in the area where TCD is large. Amide II is most delocalized over the amide and 2 methyl groups. Even though S1 gives a better solvent peak shift (+14 cm⁻¹) than S2 (-16 cm⁻¹) compared with experiment (+76 cm⁻¹), neither is very good. The calculated amide I, II, and III peak intensities are in very good agreement with experiment (Table 6). The amide A intensity is strongly enhanced in water (1.42 × 10⁻² Debye²) compared to the gas phase (2.42 × 10⁻³ Debye). The effect is much smaller for the amide I intensity, implying that the N-H bond is more polarized with the N-H stretch vibration in solution.

V. The Collective Solvent Coordinates

We have defined the following anharmonicity parameters for the overtones and combination bands

$$\Delta^v \equiv \omega^{v_1} + \omega^{v_2} - \omega^v \quad (13)$$

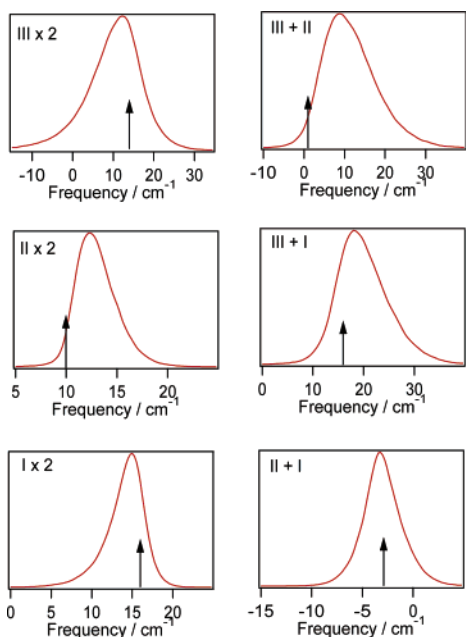


Figure 12. Static distribution of the anharmonicity (eq 13) of 3 overtone and 3 combination states. Arrows represent the gas-phase anharmonicity.

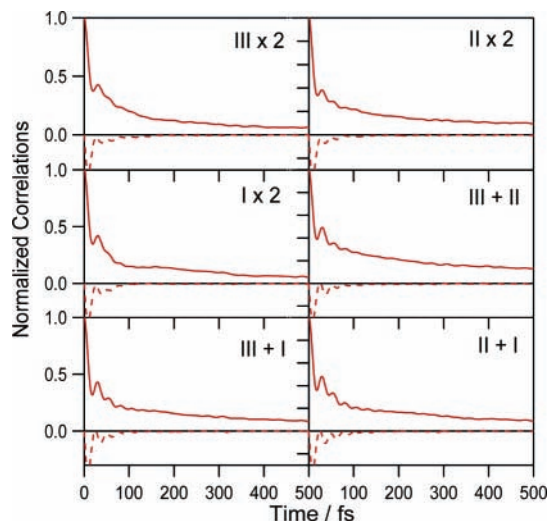


Figure 13. Normalized correlation functions of 3 overtone and 3 combination states. Solid lines – real part, dashed lines – imaginary part.

where ν_1 and ν_2 are the two fundamental states for the overtone ($\nu_1 = \nu_2$) or combination states ($\nu_1 \neq \nu_2$). The static distributions of the anharmonicities of 3 overtone and 3 combination states are displayed in Figure 12. All states have a narrow (fwhm: 5–15 cm^{-1}) asymmetric distributions. Anharmonicities of the amide III and I overtones are a few cm^{-1} smaller in water than in the gas phase, whereas the amide II overtone anharmonicity becomes larger in water. The correlation functions of these 6 anharmonicities are displayed in Figure 13. Scatter plots of the anharmonicities versus their fundamental frequencies are given in Figure 14. The correlation between the two defined in the same way as eq 11 is very weak, as shown in Table 7. The amide I anharmonicity and fundamental have almost no correlation 0.06.

The 10 amide frequencies depend quadratically on the 19 electrostatic components \mathbf{C} in the DFT map (eq 9). We can extract the linear relationship between the frequency and electrostatic fluctuations by linearizing the map around the

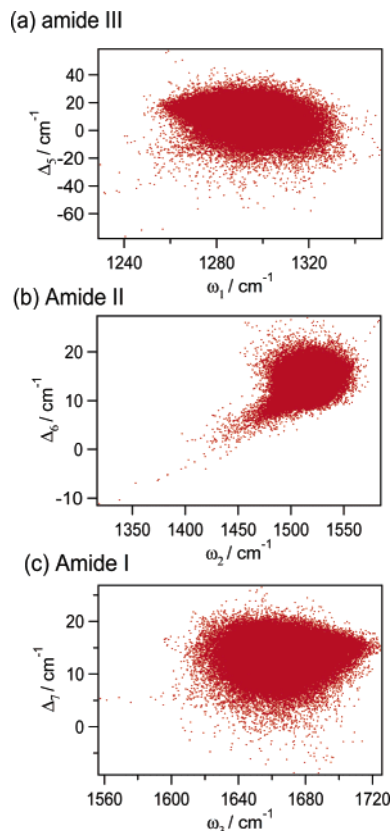


Figure 14. Scatter plot of the amide III, II, and I anharmonicity versus the fundamental frequency.

equilibrium electrostatic components $\langle \mathbf{C} \rangle$. A collective solvent coordinate Ω_ν was constructed to describe the fluctuations of each of the 10 states. These are given by a linear combination of the fluctuating part of the 19 electrostatic components

$$\Omega_\nu \equiv \sum_{i=1}^{19} L_i^\nu \delta E_i \quad (14)$$

where $\delta \mathbf{C} = (\delta E_x, \dots, \delta E_{yz})$ and $\mathbf{L}^\nu = (L_1^\nu, \dots, L_{19}^\nu)$ are defined as

$$\delta \mathbf{C} = \mathbf{C} - \langle \mathbf{C} \rangle \quad (15)$$

$$\mathbf{L}^\nu = \mathbf{O}^{\nu(1)} + \mathbf{O}^{\nu(2)\dagger} \langle \mathbf{C} \rangle \quad (16)$$

The collective coordinates offer a simple physical picture for solvent fluctuations and allow the high level simulation of line shapes.⁴² The DFT map frequencies ω_{am}^ν in eq 9 for the 4 fundamentals are approximated by a quadratic polynomial in the collective coordinate Ω_ν

$$\omega_{\text{cc}}^\nu = \omega_{\text{eq}}^\nu + \Omega_\nu + k\Omega_\nu^2 \quad (17)$$

where the coefficient of Ω_ν is 1 by definition of the collective coordinate (eqs 14–16). The contribution of the electrostatic component E_i to the solvent shift is approximated as

$$\omega_{\text{ss}}^{\nu_i} = L_i^\nu \langle E_i \rangle \quad (18)$$

The anharmonicities Δ^ν for the 3 overtone and the 3 combination states are also approximated as

$$\Delta_{\text{cc}}^\nu = \Delta_{\text{eq}}^\nu + \Omega_\nu' + k\Omega_\nu'^2 \quad (19)$$

TABLE 8: Collective Solvent Coordinates for All 10 Amide States (4 Fundamentals, 3 Overtones, and 3 Combinations) in cm^{-1} ^a

<i>i</i>	L_i^v				$L_i^{v'}$					
	III	II	I	A	III × 2	II ×	I × 2	III+II	III+I	II+I
<i>x</i>	2045	208	-2685	-3376	-624	234	-152	1491	825	188
<i>y</i>	-52	-313	-2422	-1482	271	-205	-287	-112	-702	-386
<i>z</i>	0	0	0	0	0	0	0	0	0	0
<i>xx</i>	378	-55	2392	-8436	-107	-229	-479	486	-19	-225
<i>yy</i>	-718	-111	-3182	1438	-132	-197	-671	291	-753	-162
<i>zz</i>	876	1643	341	-798	-988	-193	18	651	593	251
<i>xy</i>	-1866	3671	3129	-6945	4515	-357	-703	68	-1022	-275
<i>xz</i>	236	435	-35	-112	-33	-195	-20	39	204	-103
<i>yz</i>	138	-1179	-223	-1288	-258	-384	15	81	-96	57
<i>xxx</i>	3933	2560	-6978	-26844	-783	186	-1113	337	-994	-771
<i>yyy</i>	-840	-4189	-3896	-3931	-1835	-925	-614	-1367	-852	-7
<i>zzz</i>	0	0	0	0	0	0	0	0	0	0
<i>xyy</i>	-7080	-3793	5368	616	250	-773	-2428	-1840	-4179	-845
<i>xyx</i>	3711	1248	-3905	-16396	-979	-305	-1747	-719	-1595	-709
<i>xxz</i>	0	0	0	0	0	0	0	0	0	0
<i>xzz</i>	796	-355	-2632	7115	-1	33	-136	-70	-131	17
<i>yyz</i>	-547	-4675	-2692	1480	-123	-67	34	-153	33	42
<i>yyz</i>	0	0	0	0	0	0	0	0	0	0
<i>xyz</i>	0	0	0	0	0	0	0	0	0	0

^a Linear coefficients L_i^v (eq 14) are shown for fundamentals. $L_i^{v'}$ (eq 20) are tabulated for overtones and combinations.

TABLE 9: Contribution of Each Electrostatic Component to the Solvent Shift of 4 Amide Fundamental States^a

<i>i</i>	$\langle E_i \rangle^a$	amide III		amide II		amide I		amide A	
		L_i^b	$\omega_{ss}^{i,c}$	L_i^b	$\omega_{ss}^{i,c}$	L_i^b	$\omega_{ss}^{i,c}$	L_i^b	$\omega_{ss}^{i,c}$
<i>x</i>	0.0132	2045	64.9	208	26.2	-2685	48.5	-3376	51.7
<i>y</i>	0.0076	-52	-1.0	-313	-22.6	-2422	25.2	-1482	12.6
<i>xx</i>	-0.0013	378	-1.2	-55	0.7	2392	4.3	-8436	-12.8
<i>yy</i>	0.0011	-718	-2.0	-111	-1.2	-3182	5.0	1438	-1.9
<i>zz</i>	-0.0001	876	-0.3	1643	-2.2	341	0.1	-798	-0.2
<i>xy</i>	-0.0008	-1866	3.5	3671	-26.6	3129	3.3	-6945	-6.1
<i>xxx</i>	0.0013	3933	12.5	2560	32.0	-6978	12.6	-26844	41.1
<i>yyy</i>	-0.0017	-840	3.4	-4189	66.6	-3896	-8.9	-3931	-7.6
<i>xyy</i>	-0.0008	-7080	12.9	-3793	27.2	5368	5.5	616	0.6
<i>xyx</i>	0.0010	3711	8.8	1248	11.8	-3905	5.3	-16396	18.8
<i>xzz</i>	-0.0005	796	-1.0	-355	1.8	-2632	-1.8	7115	4.4
<i>yyz</i>	0.0003	-547	-0.4	-4675	-13.1	-2692	1.1	1480	-0.6

^a The units are a.u.⁻¹. ^b The units are cm^{-1} ·a.u.. ^c These columns give ω_{ss}^i normalized as $\sum_i \omega_{ss}^i = 100$. Noncontributing electrostatic components (E_z , E_{xz} , etc.) are not shown.

where $\Omega_v' = \Omega_{v_1} + \Omega_{v_2} - \Omega_v$ is the collective solvent coordinate for the anharmonicity Δ^v . Here v_1 and v_2 are the two fundamental states for the overtone or combination state. The resulting collective coordinate frequency for the 3 overtone and 3 combination states are $\omega_{cc}^v = \omega_{cc}^{v_1} + \omega_{cc}^{v_2} - \Delta_{cc}^v$. The contribution of each electrostatic component to the solvent shift of the anharmonicity Δ_{ss}^v is approximated as

$$\Delta_{ssi}^v = L_i^{v'} \langle E_i \rangle \quad (20)$$

where $L_i^{v'} = L_i^{v_1} + L_i^{v_2} - L_i^v$. The collective coordinates L_i for the 4 fundamentals, 3 overtones, and 3 combination states are given in Table 8. The contribution of each electrostatic component to the solvent shift ($\Delta\omega_i$) is shown in Table 9. The quadratic coefficients in eq 17 of all modes are tabulated in Table 10. E_x contributes most (49, 65, and 52%) to the solvent peak shift of amide I, III, and A. This is attributed to the fact that the hydrogen bonding to either O4 or H6 atoms is parallel to the *x* axis causing the positive E_x (corresponding to the large $\langle E_x \rangle$), and the transition dipole moments in the *x* direction are large especially for amide I and A modes (corresponding to the large coefficients L_x in collective coordinates). E_y contributes 25% and 13% to the amide I and A solvent shift. The second-order derivative tensors (E_{xxx} , E_{yyy} , and E_{xyy}) make strong

TABLE 10: Quadratic Polynomials of the Collective Coordinates for 10 Amide States: $\omega_{cc} = \omega_{eq} + \Omega + k\Omega^2$ for the 4 Fundamental States and $\Delta_{cc} = \Delta_{eq} + \Omega' + k\Omega'^2$ for the Anharmonicity of 3 Overtone and 3 Combination States

state	$\omega_{eq} (\Delta_{eq})$	k_2	shift ^a	fwhm ^b
1	1292.7	0.005301	+1.3	100.1
2	1516.2	-0.002990	+0.7	101.6
3	1670.0	-0.003862	+1.2	100.5
4	3449.9	-0.006801	-0.5	104.8
5	10.7	-0.014946		
6	13.0	0.054489		
7	14.3	-0.026074		
8	11.0	0.019615		
9	19.7	-0.023374		
10	-3.1	-0.053952		

^a The differences of the solvent shifts in the infrared spectrum between collective coordinates and the DFT map are shown in cm^{-1} .

^b The ratios of the fwhms from collective coordinates to the fwhms from the DFT map are shown in %.

contributions (32, 67, and 27%) to the amide II solvent shift. They are also significant for amide I, III, and A. The diagonal electric field gradient in the out-of-plane direction E_{zz} is negligible for all modes, which has a significant contribution in HOD in the D₂O case.⁴¹ In pure water the two out-of-plane hydrogen bonds to the HOD oxygen cause the diagonal electric

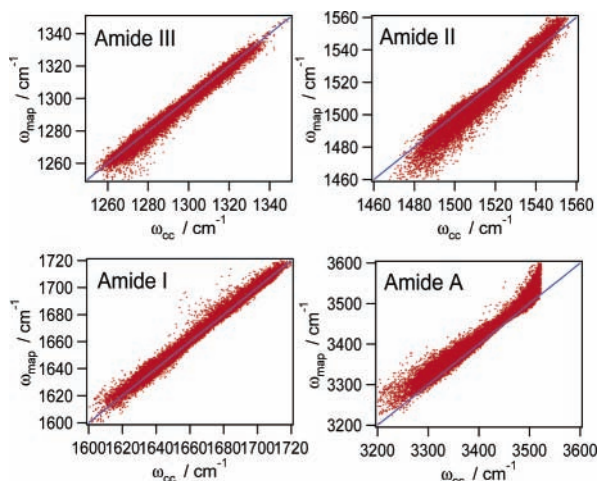


Figure 15. Scatter plots of the fundamental frequencies from DFT map ω_{map} versus the approximate collective coordinate expressions ω_{cc} (eq 17). The blue lines have a unit slope $\omega_{\text{map}} = \omega_{\text{cc}}$.

field gradient in the out-of-plane direction and lower the O–H frequency. In NMA, three hydrogen bonds to the carbonyl oxygen and the amino hydrogen are all in the C_s symmetry plane resulting in no significant field gradient E_{zz} . The contributions of the autocorrelation and crosscorrelation of electrostatic components to the frequency variance $\langle \delta\omega^2 \rangle$ are approximated as

$$\langle \delta\omega^2 \rangle \approx \sum_{i,j} L_i L_j \langle \delta E_i \delta E_j \rangle \quad (21)$$

The contributions to the frequency variance of amide I, III, and A are shown in Table 12 of the Supporting Information. $\langle E_x E_x \rangle$ contributes most (37% and 48%) to amide I and III. $\langle E_y E_y \rangle$ has the second largest contribution to the amide I. The second-order derivatives ($\langle E_x E_{xy} \rangle$ and $\langle E_{xy} E_{xy} \rangle$) are more important for amide III.

Scatter plots of ω_{map} vs ω_{cc} for the 4 fundamental states and anharmonicities of 3 overtone and the 3 combination states are given in Figures 15 and 16. Frequencies of all 4 fundamental states and anharmonicities⁴¹ from the collective coordinate agree very well with the DFT map. The infrared spectrum calculated

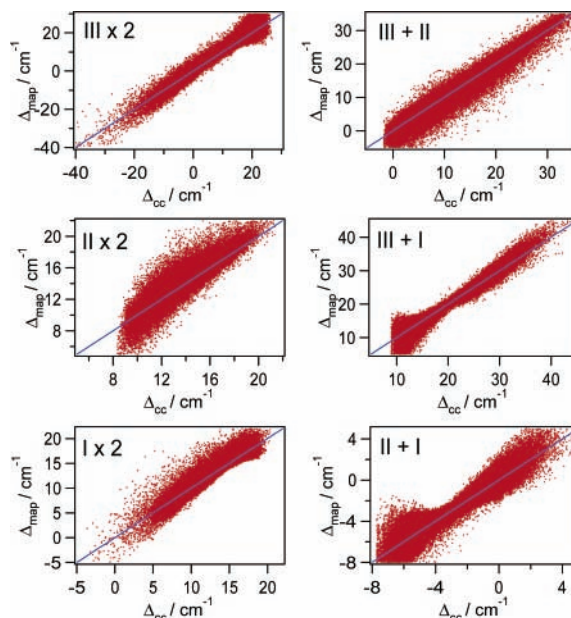


Figure 16. Scatter plots of anharmonicities from the DFT map (Δ_{map}) versus the approximate collective coordinate expressions (Δ_{cc} in eq 19) for 3 overtone and 3 combination states. The blue lines have a unit slope $\Delta_{\text{map}} = \Delta_{\text{cc}}$.

using the collective coordinate using eq 17 is displayed in Figure 17 of the Supporting Information. The difference of the solvent peak shift of the infrared band between the collective coordinate and DFT map calculations and the ratio of the collective coordinate fwhm with respect to the DFT map fwhm are reported in Table 10.

The collective solvent coordinates for the internal coordinates are also defined in the same way as eq 14. The contributions of each electrostatic component to the solvent shifts of the internal coordinates are listed in Table 11.

VI. Discussion

We have analyzed the solvent effects on the infrared spectrum of NMA using a DFT electrostatic map constructed at the BPW91/6-31G(d,p) level. The vibrational anharmonic Hamil-

TABLE 11: Collective Solvent Coordinates for the Internal Coordinates

i	$\langle E_i \rangle^a$	C=O		C–N		N–H		C–N–H	
		$L_i'^b$	$R_{ss}^i{}^c$	$L_i'^b$	$R_{ss}^i{}^c$	$L_i'^b$	$R_{ss}^i{}^c$	$L_i'^b$	$R_{ss}^i{}^c$
x	0.0132	0.811	71.3	−0.674	140.3	0.257	58.4	9.35	−48.6
y	0.0073	0.173	8.4	0.877	−101.4	0.037	4.6	−24.24	70.0
z	0.0004	−0.000	−0.0	−0.000	0.0	0.000	0.0	0.06	−0.0
xx	−0.0013	−0.523	4.5	0.495	10.2	0.177	−3.9	70.25	36.2
yy	0.0012	0.611	4.7	−0.289	5.2	−0.052	−1.0	−78.81	35.8
zz	−0.0002	−0.034	0.0	−0.245	−0.6	−0.123	0.3	8.12	0.5
xy	−0.0008	−0.621	3.1	0.539	6.4	0.188	−2.4	−349.40	−103.7
xz	−0.0001	−0.000	0.0	−0.000	−0.0	−0.000	0.0	−0.04	−0.0
yz	0.0000	0.000	−0.0	0.000	0.0	0.000	0.0	−0.01	−0.0
xxx	0.0013	1.305	11.4	−0.272	5.6	0.666	15.1	66.58	−34.6
yyy	−0.0017	0.509	−5.6	2.624	68.8	−0.255	7.3	102.12	66.9
zzz	0.0001	0.000	0.0	0.000	0.0	0.000	0.0	0.00	0.0
xyy	−0.0007	−0.734	3.6	−0.551	−6.5	−0.407	5.2	118.99	35.2
xyx	0.0010	0.150	0.9	0.334	−5.2	0.604	10.2	−54.36	21.1
xxz	0.0000	0.000	0.0	0.000	0.0	0.000	0.0	0.00	0.0
xzz	−0.0005	0.905	−3.1	−0.980	−8.1	−0.675	6.1	10.33	2.1
yzx	0.0003	0.231	0.4	3.021	−14.7	0.000	0.0	−155.13	18.9
yyz	0.0000	0.000	0.0	0.000	0.0	0.000	0.0	0.00	0.0
xyz	0.0000	0.000	0.0	0.000	0.0	0.000	0.0	0.00	0.0

^a The units for columns are a.u.^{−1}. ^b The units for columns are cm^{−1}·a.u.. ^c These columns represent R_{ss}^i normalized as $\sum_i R_{ss}^i = 100$. Noncontributing electrostatic components (E_z , E_{xz} , etc.) are not shown.

tonian is expanded to sixth order in 5 normal coordinates located on the amide bond (a mixture of N–H bend and methyl deformation, amide III, II, I, and A modes) in the presence of the external electric field and its gradient-tensor. Calculations were carried out using Gaussian 03 modified to allow geometry optimizations in a spatially inhomogeneous electric field (see Appendix A of ref 41). The vibrational eigenstates were calculated using the Arnoldi algorithm for diagonalization of the Hamiltonian using the harmonic basis set.

The electrostatic potential generated by the solvent is expanded to third order in Cartesian displacements, and the ab initio electrostatic map is constructed by expanding the frequencies of amide I, II, III, and A modes to second order in the 19 electrostatic components. The coefficients are calculated by numerical derivatives of the vibrational eigenstates of the sixth-order anharmonic vibrational Hamiltonian constructed around the equilibrium geometries in the presence of spatially nonuniform external electric fields. Therefore the DFT map implicitly includes the geometry changes of NMA¹⁹ due to the solvent-induced polarization.

The equilibrium geometry fluctuations of NMA were calculated. The C=O and N–H stretch bond lengths are longer than in the gas phase and the C–N stretch bond length is shorter. Negative correlation between the C=O and C–N bond lengths is observed. This suggests that the relative contribution of the resonance $N2^+=C3-O4^-$ form becomes larger and the $N2-C3=O4$ form becomes smaller in water than in the gas phase. The amide I and A frequencies are strongly correlated with the C=O and N–H stretch bond lengths, respectively.

The calculated amide I, II, and III gas-phase frequencies are in good agreement with experiment to within 3–27 cm^{-1} with no scaling. The calculated amide A mode is 70 cm^{-1} lower than experiment. The fundamental amide I, II, and III frequencies are lower by 9–12 cm^{-1} compared to the harmonic frequencies. The fundamental amide A frequency is 155 cm^{-1} lower than the harmonic frequency. The BPW91 functional has been known to give good harmonic frequencies of the NMA amide. The anharmonic effects on the amide fundamental frequencies are comparable to the difference between the calculated frequency and experiment, suggesting the importance of anharmonic effects when high level electron correlation contributions are included. The fidelity of DFT functionals for vibrational spectroscopy should therefore be tested by comparing not only the harmonic frequencies but also the actual vibrational eigenstates with experiment.

Infrared absorption of the amide modes was calculated by combining the ab initio electrostatic map and the trajectory of the 19 electrostatic components obtained by least-squares fits. Two sampling strategies for the inhomogeneous electric field obtained from the MD trajectory were employed to determine these components. S1 uses 65 sampling points which cover the dominant areas of the transition charge densities (TCDs) of amide modes. S2 uses 7 points (4 amide atoms, the origin of the coordinate, and 2 out of plane). S1 reproduces the global electric field structure, whereas S2 gives the electric field locally at 4 atom positions in the amide group. Both S1 and S2 reproduce the experimental solvent shift and line width of amide I and give a similar solvent shift and line width of amide A. However only S1 reproduces the experimental amide III solvent shift and predicts the amide II blueshift, consistent with experiment (S2 gives a redshift). Overall the global sampling (S1) performs better than the atom position sampling (S2). The amide II and III modes cannot be parametrized by the electrostatic potential at the atom positions alone. This can be

rationalized since the amide III TCD is more delocalized over the amide group and their bonds to carbons in 2 methyl groups. The global electric field structure in the area where the TCDs of amide vibrations are significant is important to the infrared response. This map can be applied to larger peptides. The electrostatic potential is spatially nonuniform, and higher order derivatives in the vicinity of NMA can improve the results for all amide modes. The amide modes are sensitive to intermolecular hydrogen bonding vibrations. The 660 cm^{-1} to 1100 cm^{-1} frequencies of the damped oscillatory profiles in the frequency autocorrelation functions of amide modes are close to the two frequencies 674 cm^{-1} and 726 cm^{-1} bending vibrations of hydrogen bonding to amide oxygen of NMA obtained from the normal mode calculation of the NMA–3H₂O cluster.

The scatter plots of fluctuating anharmonicity of amide III, II and I modes versus their fundamental frequencies shown in Figure 14 suggest that the amide III and I mode anharmonicities have almost no correlation with their fundamental frequencies, whereas the amide II anharmonicity has a weak positive correlation with its fundamental frequency.

The absence of correlation between the anharmonicities and fundamentals may be traced to their different dependence on the electrostatic components of the solvent collective coordinates (Table 8). For example the anharmonicity and fundamental amide I frequencies depend on E_x and E_y with the coefficient having the same sign but depend on some of the electric field gradient components with different signs. Therefore the electric field shifts the anharmonicity and fundamental frequency in the same direction (positive correlation), but some electric field gradient components cause opposite shifts (negative correlation), resulting in almost no correlation when all electrostatic components are included. The positive correlation due to the constant electric field can be rationalized using the Morse potential model in the presence of the dipole–constant external electric field interaction. By neglecting the mechanical anharmonicity, the stabilization energy $-\mu \cdot E$ becomes linear in the coordinate, which shifts the energy minimum and shifts the fundamental and anharmonicity in the same direction. The interactions of the molecule with the electric field derivatives are complex and affect the fundamental and anharmonicity in different ways.

The 4 amide fundamental frequencies and the anharmonicities of 3 overtones and 3 combination states can be reproduced by a quadratic dependence on 10 collective solvent coordinates. The frequencies from the DFT map and the collective coordinates are very close, and the linear-infrared spectra from the DFT map and the collective coordinates are virtually identical, suggesting the collective solvent coordinates provided a good description of frequency fluctuations. Correlations between the anharmonicity of 6 overtone and combination states and their fundamental frequencies are found to be very weak. E_x contributes primarily to the collective coordinates for amide I and III. The second-order derivatives of electric field (E_{xxx} , E_{yyy} , and E_{yyy}) contribute significantly to all modes, especially the amide II.

Finally, we summarize the procedure for calculating the infrared spectrum of NMA amide modes and generating the diagonal frequency fluctuations of protein local amide modes. MD simulations should be performed for more than 1 ns to get enough statistics. The electric field and its derivatives are calculated using least-squares fitting with sampling points S1 described in section II. Collective coordinates for 4 amide fundamental states can be calculated using eq 14. The collective coordinates for the anharmonicity of the 3 overtone and 3

combination states are calculated by eq 20. All coefficients of the collective coordinates are presented in Table 8. Each amide frequency is calculated as a quadratic polynomial of the collective coordinate (eq 17 for fundamental and eq 19 for anharmonicities). The quadratic expansion coefficients are shown in Table 10. The spectrum of NMA can be calculated using CGF approximation (see eqs 20–23 of ref 41).

Acknowledgment. The support of the National Institutes of Health Grant No. (2 RO1 GM059230-05) and the National Science Foundation Grant No. (CHE-0446555) is gratefully acknowledged. We wish to thank Prof. Timothy A. Keiderling for providing his experimental data.

Supporting Information Available: Tables 8–14 and Figures 15–17. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References and Notes

- Asbury, J. B.; Steinel, T.; Stromberg, C.; Gaffney, K. J.; Piletic, I. R.; Goun, A.; Fayer, M. D. *Chem. Phys. Lett.* **2003**, *374*, 362.
- Bredenbeck, J.; Hamm, P. *J. Chem. Phys.* **2003**, *119*, 1569.
- Khalil, M.; Demirdoven, N.; Tokmakoff, A. *J. Phys. Chem. A* **2003**, *107*, 5258.
- Krummel, A. T.; Mukherjee, P.; Zanni, M. T. *J. Phys. Chem. B* **2003**, *107*, 9165.
- Woutersen, S.; Pfister, R.; Hamm, P.; Mu, Y.; Kosov, D. S.; Stock, G. *J. Chem. Phys.* **2002**, *117*, 6833.
- Golozka, O.; Khalil, M.; Demirdoven, N.; Tokmakoff, A. *Phys. Rev. Lett.* **2001**, *86*, 2154.
- Thompson, D. E.; Merchant, K. A.; Fayer, M. D. *J. Chem. Phys.* **2001**, *115*, 317.
- Woutersen, S.; Hamm, P. *J. Chem. Phys.* **2001**, *115*, 7737.
- Woutersen, S.; Hamm, P. *J. Chem. Phys.* **2001**, *114*, 2727.
- Zanni, M. T.; Asplund, M. C.; Hochstrasser, R. M. *J. Chem. Phys.* **2001**, *114*, 4579.
- Piryatinski, A.; Tretiak, S.; Chernyak, V.; Mukamel, S. *J. Raman Spectrosc.* **2000**, *31*, 125.
- Merchant, K. A.; Noid, W. G.; Akiyama, R.; Finkelstein, I. J.; Goun, A.; McClain, B. L.; Loring, R. F.; Fayer, M. D. *J. Am. Chem. Soc.* **2003**, *125*, 13804.
- Tanimura, Y.; Mukamel, S. *J. Chem. Phys.* **1993**, *99*, 9496.
- Hamm, P.; Lim, M. H.; Hochstrasser, R. M. *J. Phys. Chem. B* **1998**, *102*, 6123.
- Scheurer, C.; Mukamel, S. *Bull. Chem. Soc. Jpn.* **2002**, *75*, 989.
- Moran, A. M.; Park, S.-M.; Mukamel, S. *J. Chem. Phys.* **2003**, *118*, 9971.
- Cheatum, C. M.; Tokmakoff, A.; Knoester, J. *J. Chem. Phys.* **2004**, *120*, 8201.
- Demirdoven, N.; Cheatum, C. M.; Chung, H. S.; Khalil, M.; Knoester, J.; Tokmakoff, A. *J. Am. Chem. Soc.* **2003**, *126*, 7981.
- Guo, H.; Karplus, M. *J. Phys. Chem.* **1992**, *96*, 7273.
- Ludwig, R.; Weinhold, F.; Farrar, T. C. *J. Phys. Chem. A* **1997**, *101*, 8861.
- Ludwig, R.; Weinhold, F.; Farrar, T. C. *J. Chem. Phys.* **1997**, *107*, 499.
- Torii, H.; Tatsumi, T.; Tasumi, M. *J. Raman Spectrosc.* **1998**, *29*, 537.
- Chen, X. G.; Schweitzerstener, R.; Asher, S. A.; Mirkin, N. G.; Krimms, S. *J. Phys. Chem.* **1995**, *99*, 3074.
- Papamokos, G. V.; Demetropoulos, I. N. *J. Phys. Chem. A* **2004**, *108*, 7291.
- Renugopalakrishnan, V.; Madrid, G.; Cuevas, G.; Hagler, A. T. *Proc. - Indian Acad. Sci., Chem. Sci.* **2000**, *112*, 35.
- Cuevas, G.; Renugopalakrishnan, V.; Madrid, G.; Hagler, A. T. *Phys. Chem. Chem. Phys.* **2002**, *4*, 1490.
- Zhang, R.; Li, H. R.; Lei, Y.; Han, S. J. *J. Mol. Struct.* **2004**, *693*, 17.
- Moller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.
- Headgordon, M.; Pople, J. A.; Frisch, M. J. *Chem. Phys. Lett.* **1988**, *153*, 503.
- Frisch, M. J.; Headgordon, M.; Pople, J. A. *Chem. Phys. Lett.* **1990**, *166*, 281.
- Perdew, J. P.; Chevary, J. A.; Vosko, S. H.; Jackson, K. A.; Pederson, M. R.; Singh, D. J.; Fiolhais, C. *Phys. Rev. B* **1992**, *46*, 6671.
- Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
- Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.
- Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
- Tsuzuki, S.; Luthi, H. P. *J. Chem. Phys.* **2001**, *114*, 3949.
- Onsager, L. *Phys. Rev.* **1938**, *54*, 554.
- Wong, M. W.; Frisch, M. J.; Wiberg, K. B. *J. Am. Chem. Soc.* **1991**, *113*, 4776.
- Kubelka, J.; Keiderling, T. A. *J. Phys. Chem. A* **2001**, *105*, 10922.
- Kwac, K.; Cho, M. H. *J. Chem. Phys.* **2003**, *119*, 2247.
- Kwac, K.; Cho, M. H. *J. Chem. Phys.* **2003**, *119*, 2256.
- Hayashi, T.; Jansen, T. I. C.; Zhuang, W.; Mukamel, S. *J. Phys. Chem. A* **2005**, *109*, 64.
- Jansen, T. I. C.; Hayashi, T.; Zhuang, W.; Mukamel, S. *J. Chem. Phys.* In press.
- Moran, A. M.; Mukamel, S. *Proc. Natl. Acad. Sci.* **2004**, *101*, 506.
- Nose, S. *J. Chem. Phys.* **1984**, *81*, 511.
- Berendsen, H. J. C.; Postma, J. P. M.; Gunsteren, W. F. v.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
- Andersen, H. C. *J. Comput. Phys.* **1983**, *52*, 24.
- Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
- Frisch, M. J. et al. Gaussian 03, revision c. 01, 2003.
- Hayashi, T.; Mukamel, S. *J. Phys. Chem. A* **2003**, *107*, 9113.
- Lehoucq, R. B.; Sorensen, D. C.; Yang, C. *ARPACK Users Guide: solution of Large Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods*, 1998.
- Hayashi, T.; Mukamel, S. *Bull. Korean Chem. Soc.* **2003**, *24*, 1097.
- Mayne, L. C.; Hudson, B. *J. Phys. Chem.* **1991**, *95*, 2962.
- Mukamel, S. *Phys. Rev. A* **1983**, *28*, 3480.
- Mukamel, S.; Abramavicius, D. *Chem. Rev.* **2004**, *104*, 2073.
- Allen, M. P.; Tildesley, D. *Computer Simulation of Liquids*; Oxford University Press: Oxford, 1987.
- Mukamel, S. *Principles of Nonlinear Optical Spectroscopy*; Oxford University Press: New York, 1995.
- Herrebout, W. A.; Clou, K.; Desseyn, H. O. *J. Phys. Chem. A* **2001**, *105*, 4865.