Structural Assignments and Dynamics of the A Substates of MbCO: Spectrally Resolved Vibrational Echo Experiments and Molecular Dynamics Simulations

Kusai A. Merchant,[†] W. G. Noid,[‡] David E. Thompson,[†,§] Ryo Akiyama,[‡] Roger F. Loring,[‡] and M. D. Fayer*[,†]

Department of Chemistry, Stanford University, Stanford, California 94305, and Department of Chemistry and Chemical Biology, Baker Laboratory, Cornell University, Ithaca, New York 14853

Received: August 19, 2002; In Final Form: October 25, 2002

Spectrally resolved three-pulse stimulated vibrational echo experiments are used as the basis for structural assignments of the A_1 and A_3 spectroscopic substates in the IR spectrum of the carbon monoxide (CO) stretch of carbonmonoxymyoglobin (MbCO). The measured dephasing dynamics of these substates is compared to the dephasing dynamics of MbCO predicted from molecular dynamics (MD) simulations. We assign the A_1 and A_3 substates to different protein conformations on the basis of the agreement between the measured and computed vibrational echoes. In the A_1 substate, the $N_\epsilon-H$ proton and N_δ of His64 are equidistant from the ligand, whereas in the A_3 substate, the $N_\epsilon-H$ of His64 is oriented toward the CO.

Structural information is essential for understanding the functions of proteins. Techniques such as X-ray diffraction,[1−3] neutron diffraction,[4] and 2D-NMR [5] have been used to investigate the structures of many proteins and a wide variety of other biomolecules. Despite the power of these methods, the identification of conformational substates that interconvert rapidly is difficult because of the limited time resolution of these techniques. A long-standing problem of this type is the assignment of the A conformational substates of the protein carbonmonoxymyoglobin (MbCO).

The infrared (IR) spectrum of the CO stretching mode of MbCO has three absorption bands, denoted A_0 (∼1965 cm^{-1}), A_1 (∼1944 cm^{-1}), and A_3 (∼1930 cm^{-1}), as shown in Figure 1.[6−8] It has been suggested that different electrostatic environments in the heme pocket arising from distinct structures are largely responsible for the observed bands.[9] The distal histidine His64 plays a prominent role in determining the CO stretching frequencies, but the tautomerization and orientation of this residue remain controversial.[1−3,10,11] His64 has two titratable nitrogens, N_δ and N_ϵ, either of which can be oriented toward the ligand through rotation of the imidazole ring. This residue is also fairly mobile and has been observed at a wide variety of distances from the CO ligand.[2,3,12] At low pH, His64 is thought to be doubly protonated and has been observed in the low pH crystal structure rotated out of the heme binding pocket away from the CO ligand.[12] This conformer, with little interaction between His64 and the ligand, is thought to correspond to the A_0 substate, because at low pH the A_0 line is the most intense IR absorption band. In addition, mutations of His64 to apolar residues produce an A substate band at approximately the same frequency as the A_0 line.[9,13] At pH greater than 6, the A_1 and A_3 substates are the most populated.[7,14] Crystal structures at neutral pH indicate that His64 is rotated into the heme pocket and is much closer to the ligand than in the A_0 substate, but the exact orientation of His64 is uncertain.[2,3] Also, the tautomerization state of the singly protonated His64 at neutral pH is unclear.

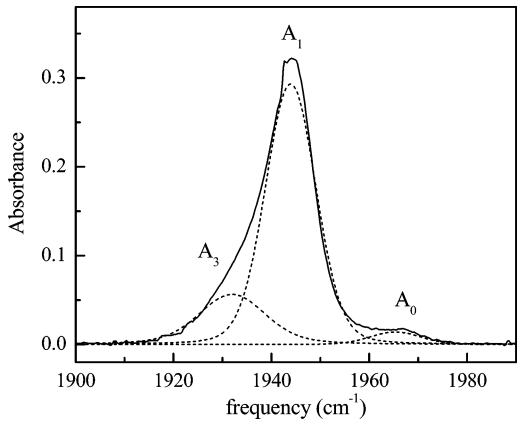

Figure 1. Background subtracted IR spectrum of horse heart MbCO at pH 7 and 298 K (solid curve) and deconvolution into the A_0, A_1, and A_3 substates (dotted curves) based on the line shapes and widths predicted from the MD simulations. The total calculated linear spectrum (not shown) has very good agreement with the measured spectrum.

Structural calculations have provided insight into the origins of the A_1 and A_3 substates. Rovira et al.[10] performed mixed quantum mechanical/molecular mechanical calculations for various possible conformers for the A substates. On the basis of shifts of the CO vibrational frequency, they concluded that His64 exists as the $N_\epsilon-H$ tautomer with the proton pointed toward the ligand. However, they were unable to assign specific structures to the A_1 and A_3 substates. Other computational studies[11] have proposed structural assignments for the A_1 and A_3 substates using the $N_\delta-H$ tautomer, which also reproduce the trends in the IR frequency shift. These studies illustrate that computations of CO vibrational frequencies associated with different structures cannot unambiguously identify the configurations associated with the spectroscopic substates.

Another route to distinguishing possible structures for the A substates is a comparison of their dynamics. Different conform-

* To whom correspondence should be addressed. E-mail: fayer@stanford.edu.
† Stanford University.
‡ Cornell University.
§ Permanent address: Department of Chemistry, Lawrence University, Appleton WI 54912.

10.1021/jp026793o CCC: $25.00 © 2003 American Chemical Society
Published on Web 12/11/2002

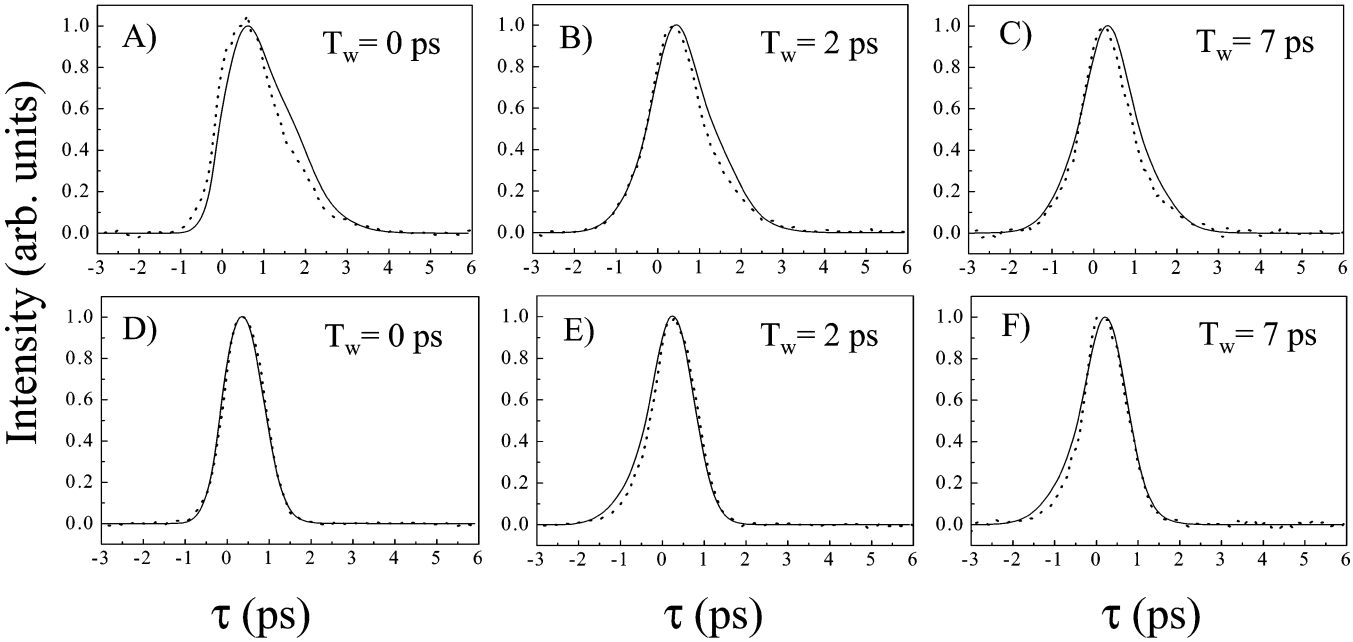

Figure 2. Comparison of the measured (dotted) and calculated (solid) spectrally resolved vibrational echo decays at 298 K for $T_w = 0$, 2, and 7 ps. The detection frequencies are 1946 cm^{-1} (A−C) and 1932 cm^{-1} (D−F).

ers of a protein undergo characteristic motions, which lead to distinct time-dependent dynamic vibrational line shapes. The time-dependent dynamic line shape of a protein conformational substate therefore acts as a signature of the underlying structure. Experimental techniques such as ultrafast infrared vibrational echoes[15−17] and multidimensional vibrational echoes[8,18−20] are capable of measuring the fast structural dynamics of proteins through the effects of these fast structural motions on dynamic vibrational line shapes. Unlike NMR techniques, vibrational echoes probe dynamics on the subpicosecond time scale. Such time scales are readily accessed with molecular dynamics (MD) simulations. The combination of the two techniques enables a direct comparison between the dynamics predicted from MD for a given structure and the experimentally measured dynamics.

The vibrational echo technique has been described in detail previously.[21,22] In the spectrally resolved stimulated vibrational echo experiments reported here, an IR pulse (center frequency = 1940 cm^{-1}, bandwidth = 95 cm^{-1}, pulse duration = 185 fs) is beam split into three pulses with wave vectors $\vec{k}_1$, $\vec{k}_2$, and $\vec{k}_3$ with variable delay time τ between the pulses with $\vec{k}_1$ and $\vec{k}_2$ and with variable delay time T_w between pulses with $\vec{k}_2$ and $\vec{k}_3$. The beams were crossed and focused (spot size ~ 150 μm) in the sample, and the vibrational echo signal was detected in the $\vec{k}_s = \vec{k}_2 + \vec{k}_3 - \vec{k}_1$ phase-matched direction. The vibrational echo signal was dispersed in a 0.5 m monochromator (210 lines/mm, resolution of 1.5 cm^{-1}) and detected with a HgCdTe detector. The ~20 mM MbCO sample was prepared by dissolving horse heart Mb (Aldrich) in 0.1 M phosphate pH 7 buffer and reducing with dithionite under a CO atmosphere. The sample was centrifuged prior to reduction and filtered before loading into a 50 μm gastight custom IR sample cell with CaF_2 windows.

Molecular dynamics simulations were performed using the MOIL software package[23] on one molecule of sperm whale MbCO, 2627 rigid water molecules, one SO_4^{2-} present in the crystal structure, and two Na^+ added to ensure electroneutrality.[24] Echo calculations reported here for the N_ϵ−H structure are based on 39 production trajectories totaling 12.7 ns in duration. The initial N_ϵ−H structure was generated from an equilibrated N_δ−H structure by moving the proton, which required modification of the original force field. Where possible, modifications were taken from the AMBER[25] and OPLS[26] force fields. All modifications were consistent with MOIL. Vibrational echo experiments measure the dephasing rate of a spectroscopic transition, which is a combination of energy relaxation (lifetime contributions) and pure dephasing (adiabatic fluctuations of the energy levels).[21,22] The lifetime contribution can be measured independently with pump−probe experiments and has been shown to contribute negligibly to dephasing in MbCO at 298 K.[8] The CO transition frequency is highly sensitive to the electric field in the heme pocket.[9,13] The pure dephasing dynamics were therefore modeled as a time-dependent Stark effect perturbation of the transition frequency arising from the fluctuating electric field at the CO caused by protein and solvent motions.[8,27] The frequency fluctuations take the form $\delta\omega(t) = \lambda[\vec{E}(t)\cdot\vec{\mu}(t) - \langle\vec{E}\cdot\vec{\mu}\rangle]$, with $\vec{\mu}(t)$ the electric dipole moment of CO, $\vec{E}(t)$ the instantaneous electric field at the midpoint of the CO bond, and λ the Stark coupling constant, discussed further below. We calculated $\vec{E}(t)$ from MD trajectories using Coulomb's law in a vacuum and atomic partial charges. The frequency-frequency correlation function (FFCF) $\langle\delta\omega(t)\,\delta\omega(0)\rangle$ was used to calculate the linear absorption spectrum and the spectrally resolved three-pulse vibrational echo signal. The linear absorption spectrum for each substate was calculated from the FFCF by treating each transition as a quantum two-level system coupled to a classical solvent. The third-order nonlinear response function was calculated by treating each substate as a quantum three-level system interacting with a classical solvent.[28] The nonlinear response function for each substate was computed from the eight relevant Feynman diagrams.[22,28] Each nonlinear response function was convolved with the temporal profiles of the excitation fields, and the total nonlinear polarization was computed from a weighted sum of polarizations associated with each conformer. The power spectrum was then computed as a function of the delay times τ and T_w and evaluated at the detection frequencies for comparison with vibrational echo data.

Figure 2 presents spectrally resolved three-pulse vibrational echo decay data at 298 K for horse heart MbCO. Measured echo decays are shown by dotted lines at two detection wavelengths, 1946 and 1932 cm^{-1}, for $T_w = 0$, 2, and 7 ps. As

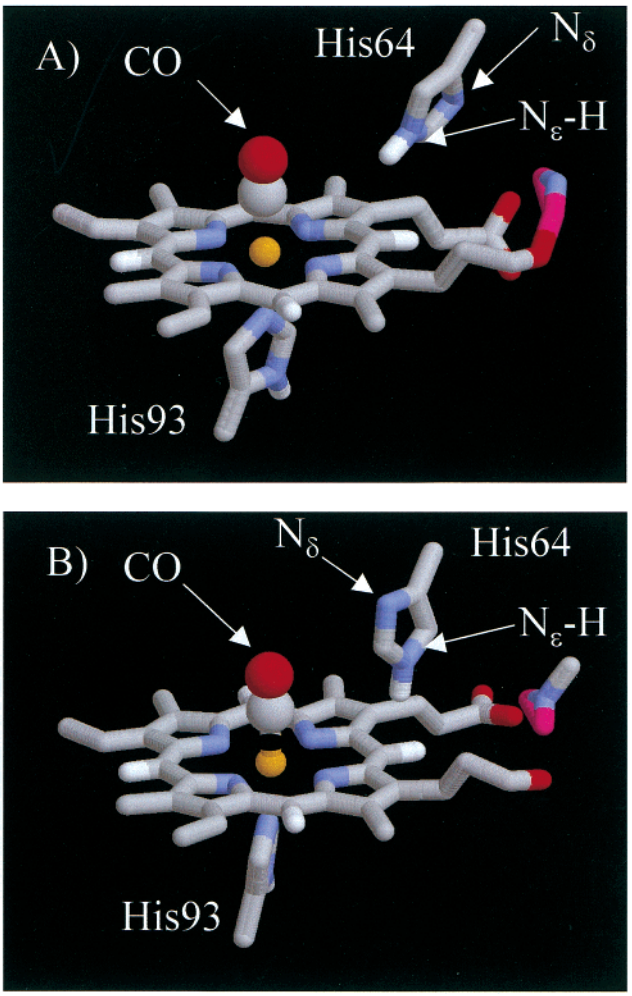

Figure 3. MD snapshots of the R and B conformational substate structures of the distal pocket. The heme group, CO ligand, distal His64, and proximal His93, are shown. (A) The R substate has $N_\epsilon-H$ oriented toward the CO ligand and N_δ oriented toward the solvent. (B) The B substate has $N_\epsilon-H$ and N_δ both approximately equidistant from the ligand. Both $N_\epsilon-H$ and N_δ are in the interior of the heme pocket. Color scheme: C, gray; N, blue; O, red; H, white; Fe, orange.

T_w is increased, the echo decay increasingly reflects the effects of slower dynamics.[17,29] The signal at 1932 cm^{-1} decays more rapidly than the signal at 1946 cm^{-1}, and the signals at the two detection wavelengths change differently as the value of T_w is increased. The time-dependent (T_w-dependent) dynamic vibrational line shape is the Fourier transform of the vibrational echo decay.[28,30] The different vibrational echo decay line shapes and distinct T_w dependences at the two detection wavelengths indicate that the conformational substates that give rise to the vibrational echo signal at these detection frequencies have distinct fast structural dynamics, as sensed by the CO ligand bound at the active site.

MD trajectories of the instantaneous Stark shift $\delta\omega(t)$ for the $N_\epsilon-H$ structure show pronounced two-state behavior with a red transition frequency state R and a blue transition frequency state B characterized by Stark shifts fluctuating about mean values separated by 9.4λ in units of cm^{-1}, with λ the value of the Stark coupling constant expressed in conventional units of $cm^{-1}/(MV\ cm^{-1})$. States R and B differ primarily in the orientation of His64 relative to the ligand. Representative configurations of both states taken from the simulations are shown in Figure 3. In state R (Figure 3A), the $N_\epsilon-H$ bond points toward the CO ligand with a mean distance from H to CO midpoint of 3.2 Å. The mean distance from N_δ to CO is 5.8 Å. In the B state (Figure 3B), His64 has rotated about the $C_\beta-C_\gamma$ bond relative to the R state, such that both the N_ϵ proton and N_δ are approximately the same distance from the ligand, with mean distances of 4.6 and 4.8 Å, respectively. In the B state, the $N_\epsilon-H$ bond is not oriented directly toward the ligand, and both $N_\epsilon-H$ and N_δ are in the interior of the protein.

To identify the simulated states R and B with the spectroscopic substates A_3 and A_1, respectively, we have calculated separate FFCFs for the R and B states. The line shape for each conformer was calculated and used to deconvolve the linear absorption spectrum of horse MbCO (Figure 1, dotted lines). The splitting between the A states and their relative populations were allowed to vary, producing a 12 cm^{-1} difference between the A_1 and A_3 transition frequencies. No FFCF was available for the A_0 substate from the MD simulations. The B state FFCF was therefore used to approximate the A_0 substate dynamics. This approximation introduces a negligible error in our analysis of the A_1 and A_3 substates because the A_0 substate ($\sim$1965 cm^{-1}) is minimally populated ($\sim$5%) at this pH and has only a small effect on the linear absorption and vibrational echo data.

Figure 2 compares the measured (dotted) and calculated (solid) spectrally resolved three-pulse vibrational echo decays at two wavelengths and three values of T_w. The calculated curves use the FFCFs for the B and R substates derived from the MD simulations, and transition frequencies and relative populations derived from the linear spectrum deconvolution. The only adjustable parameter in the echo calculation is the Stark coupling constant λ, whose optimum value was found to be 1.9 $cm^{-1}/(MV\ cm^{-1})$. This quantity has been measured by Park et al.[31] from vibrational Stark effect experiments on an aqueous MbCO sample. Park et al. [31] report $\lambda = 2.4/f\ cm^{-1}/(MV\ cm^{-1})$, with f a correction factor obeying $f > 1$ that accounts for the difference between the local and externally applied electric fields in a dielectric medium. The appropriate value for f is dependent on the model of the dielectric medium.[32] The calculation reproduces the more rapid dephasing decay at 1932 cm^{-1} relative to 1946 cm^{-1}, as well as the T_w dependence, which reflects the dynamics of slower spectral diffusion processes.[17,29] The predicted and measured dynamics for the A_1 and A_3 substates in Figure 2 agree remarkably well. In addition, with a value of $\lambda = 1.9$ $cm^{-1}/(MV\ cm^{-1})$, the difference in mean frequencies of the R and B states determined from MD simulation is 18 cm^{-1}, similar to the 12 cm^{-1} splitting of the A_1 and A_3 lines in the absorption spectrum. The extent of the agreement between the predicted and measured dynamics is strong evidence for the assignment of the B and R structures with the $N_\epsilon-H$ tautomer to the A_1 and A_3 conformational substates of MbCO.

The R configuration that we assign to A_3 is similar to the structures observed in many crystallographic studies,[1−4,12] and also resembles the predominant structure proposed by Rovira et al.[10] Rovira et al. also considered a structure (structure IV of Figure 3 in that work[10]) that, like our B state, has the His64 rotated so that the $N_\epsilon-H$ points away from the ligand, but with N_δ closer to the ligand than in the B structure. They proposed that their structure could contribute to the A_0 line. Our experiments and calculations show that the B structure is consistent with A_1. The B configuration that we assign to A_1 has not been reported crystallographically. However, whereas the A_1 substate is the most populated substate in solution, the A_3 substate is most populated in pure MbCO crystals.[33] Also, rotamers of His64 might be difficult to distinguish in an X-ray structure, especially if the residue is relatively flexible. There is still no consensus from crystallographic data[2,3] on the exact

configuration of His64. The interconversion time between A_1 and A_3 states in solution has been estimated by Johnson et al.[34] to be ~1 ns, whereas the interconversion time between A_1/A_3 and A_0 is estimated to be ~1 μs. These estimates are consistent with our having observed in the MD simulations occasional transitions between R and B states on the nanosecond time scale, but not having observed a state that could be identified as A_0.

Previously reported simulations of MbCO with protonated N_δ[8,27] yielded two species with distinct FFCFs: one in which one water molecule resides within the heme pocket, and another in which no water molecule is located within the pocket. Structures in which the distance from the CO to the N_δ proton exceeded 6.5 Å were assigned to the A_0 substate[11] and were not further analyzed. Because neither species displayed spectroscopically distinct states that could be identified as A_1 and A_3, we do not further consider the $N_\delta-H$ structure here. A complete discussion of the $N_\delta-H$ structure will be given elsewhere.

The vibrational echo experiments reported here were performed on horse heart MbCO, whereas the MD simulations used to calculate the dephasing dynamics employed a model of sperm whale MbCO. The primary structures of sperm whale MbCO and horse heart MbCO differ at 20 residues.[35] The substitutions are generally conservative, and none of them is near the distal heme pocket. The ligand binding kinetics and heme affinity[35] (the dominant factor in determining folding stability[36]) for the two holoproteins are very similar. Nonetheless, minor differences in the peak shapes and positions have been noted in the IR spectra of MbCO from different species.[9,13] We assume that these small spectral differences do not affect the interpretation of the structural origins of the A substates. A quantitative study on the species dependence of MbCO substate dynamics is currently in progress.

The computations reported here differ from prior simulations of MbCO in the calculation of the nonlinear vibrational response, which we compare directly to experiments that probe the dynamics of the CO ligand in different protein conformational states. We have used this comparison to assign the A_1 and A_3 spectroscopic substates of MbCO to two conformers observed in MD simulations of the $N_\epsilon-H$ tautomer of His64. The comparison of MD simulations and spectrally resolved stimulated vibrational echo experiments permits atomic level interpretation of protein dynamics measurements and provides a rigorous test for the predictions made in MD simulations.

Acknowledgment. K.A.M., D.E.T., and M.D.F. acknowledge the National Institutes of Health (1R01-GM61137) for support of this research. R.F.L., W.G.N., and R.A. acknowledge support from the National Science Foundation (CHE-0105623) and the Petroleum Research Fund of the American Chemical Society. The molecular dynamics portion of this research was carried out using the resources of the Cornell Theory Center, which receives funding from Cornell University, New York State, federal agencies, and corporate partners. R.A. is a Research Fellow of the Japan Society for the Promotion of Science (2000). W.G.N. acknowledges fellowship support from Cornell's IGERT program in nonlinear systems, funded by NSF Grant DGE-9870681. K.A.M. was partially supported by an Abbott Laboratories Stanford Graduate Fellowship. We thank Ileana Stoica and Prof. Ron Elber for helpful discussions regarding MOIL.

References and Notes

(1) Kuriyan, J. W.; Karplus, M.; Petsko, G. A. *J. Mol. Biol.* **1986**, *192*, 133.
(2) Kachalova, G. S.; Popov, A. N.; Bartunik, H. D. *Science* **1999**, *284*, 473.
(3) Vojtechovsky, J.; Chu, K.; Berendzen, J.; Sweet, R. M.; Schlichting, I. *Biophys. J.* **1999**, *77*, 2153.
(4) Cheng, X.; Schoenborn, B. P. *J. Mol. Biol.* **1991**, *220*, 381.
(5) Theriault, Y.; Pochapsky, T. C.; Dalvit, C.; Chiu, M. L.; Sligar, S. G.; Wright, P. E. *J. Biomol. NMR* **1994**, *4*, 491.
(6) Ansari, A.; Beredzen, J.; Braunstein, D.; Cowen, B. R.; Frauenfelder, H.; Hong, M. K.; Iben, I. E. T.; Johnson, J. B.; Ormos, P.; Sauke, T.; Schroll, R.; Schulte, A.; Steinback, P. J.; Vittitow, J.; Young, R. D. *Biophys. Chem.* **1987**, *26*, 337.
(7) Müller, J. D.; McMahon, B. H.; Chen, E. Y. T.; Sligar, S. G.; Nienhaus, G. U. *Biophys. J.* **1999**, *77*, 1036.
(8) Merchant, K. A.; Thompson, D. E.; Xu, Q.-H.; Williams, R. B.; Loring, R. F.; Fayer, M. D. *Biophys. J.* **2002**, *82*, 3277.
(9) Phillips, G. N., Jr.; Teodoro, M. L.; Li, T.; Smith, B.; Olson, J. S. *J. Phys. Chem. B* **1999**, *103*, 8817.
(10) Rovira, C.; Schulze, B.; Eichinger, M.; Evanseck, J. D.; Parrinello, M. *Biophys. J.* **2001**, *81*, 435.
(11) Schulze, B. G.; Evanseck, J. D. *J. Am. Chem. Soc.* **1999**, *121*, 6444.
(12) Yang, F.; Phillips, G. N., Jr. *J. Mol. Biol.* **1996**, *256*, 762.
(13) Li, T. S.; Quillin, M. L.; Phillips, G. N., Jr.; Olson, J. S. *Biochemistry* **1994**, *33*, 1433.
(14) Shimada, H.; Caughey, W. S. *J. Biol. Chem.* **1982**, *257*, 1893.
(15) Zimdars, D.; Tokmakoff, A.; Chen, S.; Greenfield, S. R.; Fayer, M. D.; Smith, T. I.; Schwettman, H. A. *Phys. Rev. Lett.* **1993**, *70*, 2718.
(16) Rector, K. D.; Engholm, J. R.; Rella, C. W.; Hill, J. R.; Dlott, D. D.; Fayer, M. D. *J. Phys. Chem. A.* **1999**, *103*, 2381.
(17) Hamm, P.; Lim, M.; Hochstrasser, R. M. *Phys. Rev. Lett.* **1998**, *81*, 5326.
(18) Zanni, M. T.; Asplund, M. C.; Hochstrasser, R. M. *J. Chem. Phys.* **2001**, *114*, 4579.
(19) Merchant, K. A.; Thompson, D. E.; Fayer, M. D. *Phys. Rev. Lett.* **2001**, *86*, 3899.
(20) Golonzka, O.; Khalil, M.; Demirdoven, N.; Tokmakoff, A. *Phys. Rev. Lett.* **2001**, *86*, 2154.
(21) Fayer, M. D. *Annu. Rev. Phys. Chem.* **2001**, *52*, 315.
(22) Hamm, P.; Hochstrasser, R. M. Structure and Dynamics of Proteins and Peptides: Femtosecond Two-Dimensional Infrared Spectroscopy. In *Ultrafast Infrared and Raman Spectroscopy*; Fayer, M. D., Ed.; Marcel Dekker: New York, 2001; Vol. 26, p 273.
(23) Elber, R.; Roitberg, A.; Simmerling, C.; Goldstein, R.; Li, H.; Verkhivker, G.; Keaser, C.; Zhang, J.; Ulitsky, A. *Comput. Phys. Commun.* **1994**, *91*, 159.
(24) Meller, J.; Elber, R. *Biophys. J.* **1998**, *74*, 789.
(25) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765.
(26) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1666.
(27) Williams, R. B.; Loring, R. F.; Fayer, M. D. *J. Phys. Chem. B* **2001**, *105*, 4068.
(28) Mukamel, S. *Principles of Nonlinear Optical Spectroscopy*; Oxford University Press: New York, 1995.
(29) Berg, M.; Walsh, C. A.; Narasimhan, L. R.; Littau, K. A.; Fayer, M. D. *J. Chem. Phys.* **1988**, *88*, 1564.
(30) Berg, M. A.; Rector, K. D.; Fayer, M. D. *J. Chem. Phys.* **2000**, *113*, 3233.
(31) Park, E. S.; Andrews, S. S.; Hu, R. B.; Boxer, S. G. *J. Phys. Chem. B* **1999**, *103*, 9813.
(32) Bottcher, C. J. F. *Theory of Electric Polarization*; Elsevier: Amsterdam, 1973; Vol. I.
(33) Makinen, M. W.; Houtchens, R. A.; Caughey, W. S. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 6042.
(34) Johnson, J. B.; Lamb, D. C.; Frauenfelder, H.; Müller, J. D.; McMahon, B.; Nienhaus, G. U.; Young, R. D. *Biophys. J.* **1996**, *71*, 1563.
(35) Scott, E. E.; Paster, E. V.; Olson, J. S. *J. Biol. Chem.* **2000**, *275*, 27129.
(36) Hargrove, M. S.; Olson, J. S. *Biochemistry* **1996**, *35*, 11310.