Probing Polar Solvation Dynamics in Proteins: A Molecular Dynamics Simulation Analysis

Andrei A. Golosov[†] and Martin Karplus*[,†,‡]

Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138, and Institut de Science et d'Ingénierie Supramoléculaires, Université Louis Pasteur, 67000 Strasbourg, France

Received: August 24, 2006; In Final Form: October 27, 2006

Measurements of time-resolved Stokes shifts on picosecond to nanosecond time scales have been used to probe the polar solvation dynamics of biological systems. Since it is difficult to decompose the measurements into protein and solvent contributions, computer simulations are useful to aid in understanding the details of the molecular behavior. Here we report the analysis of simulations of the electrostatic interactions of the rest of the protein and the solvent with 11 residues of the immunoglobulin binding domain B1 of protein G. It is shown that the polar solvation dynamics are position-dependent and highly heterogeneous. The contributions due to interactions with the protein and with the solvent are determined. The solvent contributions are found to vary from negligible after a few picoseconds to dominant on a scale of hundreds of picoseconds. The origin for the latter is found to involve coupled hydration and protein conformational dynamics. The resulting microscopic picture demonstrates that a wide range of possibilities have to be considered in the interpretation of time-resolved Stokes shift measurements.

I. Introduction

Understanding the dynamics of biological systems is an important step toward gaining insight into the relationship between their structures and functions. Recent advances in ultrafast optical spectroscopy allow one to probe the dynamics of biological systems from femtoseconds to nanoseconds,[1,2] a time scale directly accessible by molecular dynamics simulations.[3] One such approach involves measurements of the time-dependent fluorescence Stokes shift (TDSS).[2,4−8] The TDSS is dominated by the response of the polar and charged groups in the environment to changes in the charge distribution of chromophore probe; this has been referred to as "polar solvation dynamics".[1] Zewail and co-workers, for example, have applied TDSS spectroscopy to study DNA−drug binding,[9] protein unfolding,[10] and enzymatic activity of proteins,[11] and Cohen et al. used the technique to investigate conformational changes.[7] The potential importance of such applications makes it essential to have a full understanding of the mechanism involved in the relaxation. An experimental approach to the mechanism is the work of Cohen et al.[6] who measured the TDSS at several sites of immunoglobulin binding domain B1 of protein G (GB1); an alanine derivative of the fluorophore 6-dimethylamino-2-acylnaphthalene (Aladan) was used as the fluorescent probe. The TDSS obtained for four sites showed bimodal relaxation with ultrafast decay (within a few picoseconds) followed by a much slower decay on the time scale of 100−200 ps. The results indicated that the time-dependent response is very heterogeneous and depends strongly on the location in the protein.

A fundamental difficulty with the analysis of TDSS measurements is that they probe the collective polar solvation response of the solvent and the solute. Thus, it has been impossible by the experiments to separate the protein and solvent contributions to the relaxation, although heuristic models have been proposed.[12] Molecular dynamics simulations[13,14] are ideally suited for this purpose because they can provide a detailed atomistic picture of the contributions of various groups (e.g., parts of the protein or solvent) to the relaxation. A pioneering study of solvation effects for interpreting the TDSS of simple solutes was made by Ladanyi and co-workers.[15,16] Since the time scales of the experiments, which are on the order of 1−10 ns, are now easily accessible by molecular dynamics simulations for moderately sized proteins or nucleic acids (up to a few hundreds residues), several molecular dynamics simulations have, in fact, been performed in relation to TDSS measurements.[17−19] Vivian and Callis[17] suggested that ordered water molecules beyond 20 Å contributed significantly to the TDSS; their results were based on a single 30 ps trajectory. Bagchi and co-workers[18] reported that the solvation dynamics are faster with increasing surface exposure of the probe and that they are sensitive to the secondary structure. Of most interest is a recent molecular dynamics simulation by Nilsson and Halle.[19] They studied the polar solvation relaxation of the single tryptophan in the protein monellin, a system for which TDSS data are available.[20] The computed results for the time scales of the relaxation are in reasonable agreement with the experimental measurements. They decomposed the TDSS relaxation data into protein and solvent contributions and found that the protein component exhibited a long time relaxation (74 ps), while the water contribution did not (1.5 ps). This led them to assign the long time relaxation to solvation by protein, in contrast to the slow exchange of bound water.[12] Their conclusion is consistent with magnetic relaxation dispersion (MRD) measurements, which found that the majority of waters on a protein surface are slowed only by a factor of 2−3 in comparison to bulk waters.[21] Also, Cohen et al.[6] had shown that the TDSS of a probe in the hydrophobic core exhibits a slow relaxation on the 100−200 ps time scale, suggesting that solvation by the protein environment can be significant.

* Author to whom correspondence should be addressed. E-mail: marci@tammy.harvard.edu.
[†] Harvard University.
[‡] Université Louis Pasteur.

10.1021/jp065493u CCC: $37.00
Published on Web 01/24/2007

In this paper we report the results of a molecular dynamics simulation of the immunoglobulin binding domain B1 of protein G, which was stimulated by the experiments of Cohen et al.[6] Since the paper of Nilsson and Halle[19] was published while we were doing the analysis, we focus this report on the aspects of our results that complement their work. Their study was done using a single probe, while we focus on the heterogeneity of TDSS relaxation by studying a series of different sites. We find that a simple decomposition into a short time component due to water and a long time component due to the protein cannot always be made. Moreover, we find cases in which the solvent contribution to relaxation is dominant and slow (>100 ps), and for these cases we investigate whether there is a connection with the hydration dynamics. Although we used the same force field (CHARMM[22]) as Nilsson and Halle, we treated the system with periodic boundary conditions and used the particle-mesh Ewald summation for long-range electrostatics, while Nilsson and Halle used the stochastic boundary method.[23] Thus, we can determine whether qualitatively similar results are obtained with different treatments of the long-range electrostatics. This comparison is important because of the essential role played by electrostatics in the TDSS. We also test the connection of relaxation rates with such simple descriptors such as probe surface exposure to the solvent and the relation to secondary structural elements and determine the radius of interaction that is being probed by the TDSS relaxation function.

II. Methods

A. System and Simulation Details. The system that we use is the B1 domain of protein G (GB1) (PDB code 1PGB), the protein studied by Cohen et al.[6] However, we did not examine the relaxation of the artificial chromophore used by them because the protein structure is not available. Instead we used the set of 11 threonine residues as probes for studying the relaxation of the polar solvation throughout the protein. Threonines were chosen because they are the most abundant residue in GB1, they are located in the different secondary structural elements (α-helices, β-strands, and loops), and they have varying degrees of exposure to the solvent (Figure 1 and Table 1). Hence, even though the threonines are not photoactive, they are good candidates for probing the heterogeneity of solvation dynamics by simulations. To determine the solvation dynamics, we calculate the time dependence of the electrostatic interaction energy of the whole residue (main chain and side chain) with the rest of the system, both protein and explicit water. Standard CHARMM charges[22] were used for the protein, and TIP3P water was used to represent the solvent.[24] We use these results to calculate correlation functions related to the TDSS, as described in section B.

GB1 is a small thermostable protein (56 residues). This makes it a very attractive candidate for molecular dynamics simulations in explicit solvent. Although computationally demanding, an explicit solvent treatment is necessary for an accurate representation of the solvation effects and the time scales of the relaxation. The system was prepared by immersing the GB1 protein with its crystal waters in a preequilibrated 68 Å × 58 Å × 53 Å water box; water molecules that overlap within 2.8 Å with the protein (and its crystal waters) were removed. Water molecules with the lowest (highest) electrostatic potentials were replaced by 21 Na^+ (17 Cl^- ions) for electroneutrality at an ionic strength of 0.15 M. Ions were placed at least 8 Å apart and from any protein atom to avoid "pseudo-crystalization" of the salt. After equilibration at constant pressure and temperature conditions (NPT), the system with 6205 water molecules had reached a box size with 66 Å × 56 Å × 51 Å dimensions.

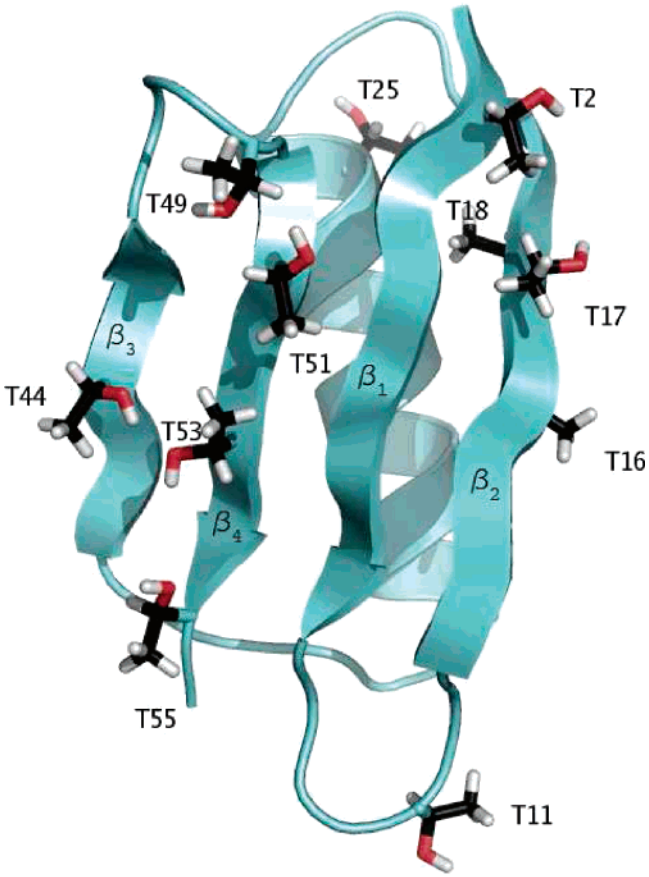

Figure 1. Ribbon diagram of GB1 showing the positions of the threonine residues. The N-terminus of the protein is at the end of the β_1 strand in the upper part of the picture, while the C-terminus is at the end of the β_4 strand in the lower part of the picture. All threonines are shown: β_1 strand, T2; β_2 strand, T16, T17, T18; loops, T11, T49; α-helix, T25; β_3 strand, T44; β_4 strand, T49, T51, T53, T55. The image was prepared using PYMOL (DeLano Scientific, http://pymol.sourceforge.net).

TABLE 1: Solvent Coverage/Secondary Structure Information on Probed Residues

residue	solvent coverage[b] (%)	secondary structure	dominant contributions[a]
T2	27 (±3)	β_1	a
T11	45 (± 4)	loop	a
T16	17 (± 3)	β_2	w
T17	29 (± 2)	β_2	a
T18	12 (± 3)	β_2	w
T25	22 (± 3)	α	a
T44	30 (± 3)	β_3	w
T49	35 (± 2)	loop	a
T51	5 (± 2)	β_4	a
T53	11 (± 2)	β_4	w
T55	20 (± 3)	β_4	a

[a] The protein, solvent, and protein + solvent (all) contributions are represented by p, w, and a. [b] Defined as the ratio (in the protein/isolated) of the accessible surface areas for a given residue. The Lee−Richards accessible surface area was calculated with a radius of 1.4 Å of the probe with the CHARMM program.

Molecular dynamics (MD) simulations at 300 K of GB1 with periodic boundary conditions were performed with the NAMD (version 2.5)[25] program. The trajectories obtained with the NAMD program were analyzed with the CHARMM[26] program. The CHARMM22[22] parameter set and TIP3P[24] water model consistent with it were used. Nonbonded interactions were calculated using a 13.5 Å list that was updated every 10 steps of dynamics. Van der Waals interactions were truncated at 12 Å

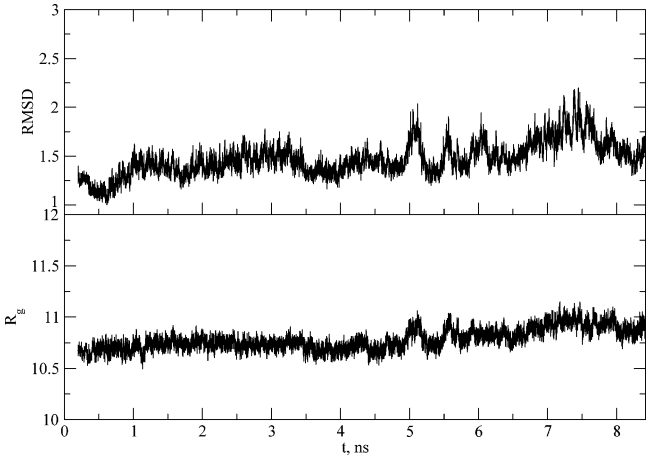

Figure 2. Root-mean-square deviations from the crystal structure and R_g in Å as function of time.

with a smooth switching function applied starting from 10 Å. All electrostatic interactions beyond 12 Å were computed with the particle-mesh Ewald summation.[27] Covalent bonds to hydrogens were constrained using SHAKE,[28] which allowed a 2 fs step for propagation. The equilibration started from the system that had been subjected to a 300 step conjugate gradient minimization. The system was equilibrated for 320 ps in the constant pressure and temperature ensemble (NPT) using the Langevin piston Nosé−Hoover method for pressure control[29] with 1 atm target pressure, 100 fs barostat oscillation time period, and a 50 fs decay time scale. The temperature was maintained at 300 K by coupling all non-hydrogen atoms to a Langevin bath with 5 ps^{-1} friction. (The solute with its crystal structure water molecules were held fixed during the first 20 ps to avoid distortions in the structure because of partial solvation and/or sudden pressure changes due to filling voids resulting from the deletion of waters during the immersion of the protein in the water box.)

After equilibration, a 8.1 ns production run was performed in a constant volume and energy ensemble (NVE) to avoid affecting the time scale by coupling to an external thermal bath and/or barostat (energy drifted 0.8% after 8.1 ns run). During the simulation, the coordinates of the system were stored every 10 steps (0.02 ps) for subsequent analysis of the last 8 ns of the production run. The radius of gyration of the protein reached a value near 10.9 Å, which is very close to the original 10.7 Å. The all-atom mass-weighted root-mean-square deviation (rmsd) fluctuated around 1.5 Å (Figure 2). This indicated that the structure was stable during the simulation run.

B. Analysis for the Determination of the TDSS. In TDSS studies, the chromophore is excited from a ground state with a system Hamiltonian H_g to the excited state with a system Hamiltonian H_e. One measures the shift of the fluorescence spectrum peak $\Delta\nu(t)$ as a function of time. This can be expressed as[30,31]

$$\Delta\nu(t) = \frac{1}{h} Tr[U\, e^{-iH_e t/\hbar}\rho_g\, e^{iH_e t/\hbar}] \tag{1}$$

where the collective "solvation" coordinate $U = H_e - H_g$ represents the energy gap between excited and ground states. The limiting values $\Delta\nu(0) = (1/h)TrU\rho_g$ and $\Delta\nu(\infty) = (1/h)TrU\rho_e$ correspond to the absorption and stationary fluorescence emission peaks, respectively, where ρ_g and ρ_e are the ground and excited-state density matrices (assumed to be normalized so that $Tr\rho_g = Tr\rho_e = 1$).

Standard analyses use the linear response approximation,[31] which is valid under assumptions of either Gaussian or small energy gap fluctuations. (In the former, cumulants higher than second one are all rigorously zero, while in the latter all cumulants beyond second order can be neglected.[31]) With this assumption and the high-temperature limit, the connection between the equilibrium classical energy gap correlation function, $C(t)$,

$$C(t) = \langle\Delta E(t)\Delta E(0)\rangle - \langle\Delta E\rangle\langle\Delta E\rangle \tag{2}$$

and the nonequilibrium measured $\Delta\nu(t)$ is given by[30]

$$\Delta\nu(t) - \Delta\nu(\infty) = \frac{C(t)}{hk_B T} \tag{3}$$

where $\Delta E(t)$ is the time-dependent energy gap between the excited state and the ground state of the chromophore. The variation in the energy gap is assumed to be dominated by the electrostatic interactions of the probe with the environment. The value of $C(t)$ can be obtained by using a thermal average over either the ground state or the excited state in eq 2.[30]

Equation 3 allows one to relate the TDSS relaxation data (i.e., the function $\Delta\nu(t)$) to the ground-state solvation dynamics. The measured TDSS relaxation is usually given in its normalized form

$$S(t) = \frac{\Delta\nu(t) - \Delta\nu(\infty)}{\Delta\nu(0) - \Delta\nu(\infty)} = \frac{C(t)}{C(0)} \tag{4}$$

which follows from eq 3.

We are interested in calculating the contributions of the probe interaction from different parts of the environment to $S(t)$. By decomposing the excited- and ground-state Hamiltonians as $H_e = \sum_\mu H_{\mu,e}$ and $H_g = \sum_\mu H_{\mu,g}$, the contributions to the TDSS can be expressed, in analogy to eq 1, as

$$\Delta\nu_\mu(t) = \frac{1}{h} Tr[(U_\mu\, e^{-iH_e t/\hbar})\rho_g\, e^{iH_e t/\hbar}] \tag{5}$$

where $U_\mu = H_{\mu,e} - H_{\mu,g}$ represents the energy gap contribution due to interaction with the μ component of the chromophore environment.

Using the linear response approximation and the classical limit derivation of eq 3 (see Appendix A for details), one obtains the expression for the component contributions[19,32] in terms of the component classical energy gap correlation function

$$C_\mu(t) = \langle\Delta E_\mu(t)\Delta E(0)\rangle - \langle\Delta E_\mu\rangle\langle\Delta E\rangle \tag{6}$$

$$\Delta\nu_\mu(t) - \Delta\nu_\mu(\infty) = \frac{C_\mu(t)}{hk_B T} \tag{7}$$

where ΔE_μ is the contribution to the energy gap between chromophore excited and ground states from interaction with the μ component (such as protein, solvent, a specific residue, etc.) and the thermal average can be taken over either the ground or the excited state, as above.

Defining the normalized component contribution as

$$S_\mu(t) = \frac{\Delta\nu_\mu(t) - \Delta\nu_\mu(\infty)}{\Delta\nu(0) - \Delta\nu(\infty)} \tag{8}$$

it follows that $S_\mu(t) = C_\mu(t)/C(0)$. Equations 3 and 7 allow one to determine the nonequilibrium solvation functions $S(t)$ and $S_\mu(t)$, using the equilibrium correlation functions $C(t)$ and

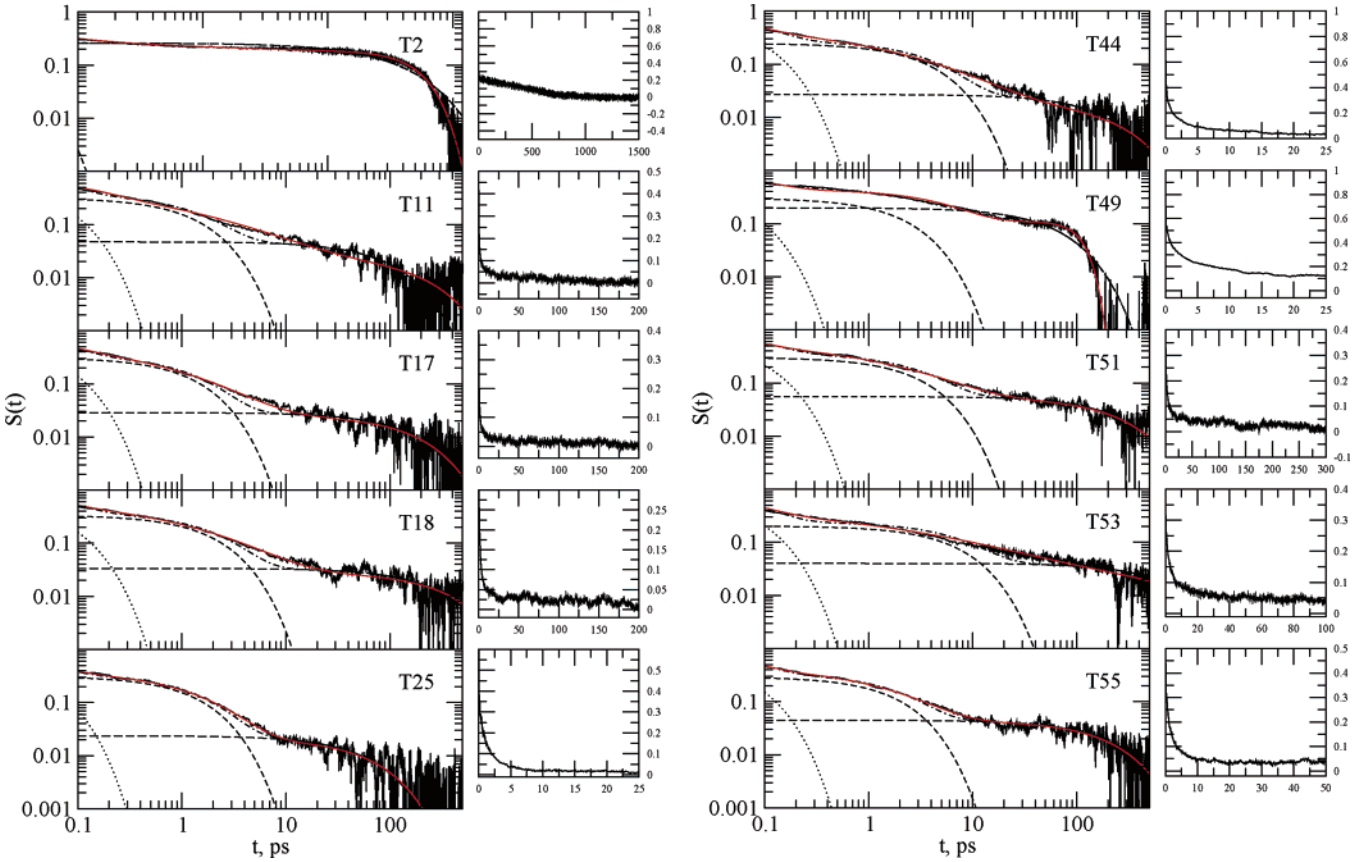

Figure 3. (Left panels in each column) $S(t)$ for indicated residues on a log−log scale with multiexponential fits (dot−dashed line) (Table 2) with individual components given by dotted, dashed, and long dashed lines as well as stretched/compressed multiexponential fits given by red lines. (Right panels in each column) $S(t)$ shown on a linear scale.

$C_\mu(t)$, which can be computed from an equilibrium molecular dynamics trajectory.

The correlation functions that we calculate are $\langle E_{\mu,\mathrm{int}}(t)E_{\mathrm{int}}(0)\rangle$, where E_{int} is the electrostatic interaction energy of the whole residue (main chain and side chain) with the rest of the system, both protein and explicit water with ions, and $E_{\mu,\mathrm{int}}$ is the electrostatic interaction energy of the whole residue with μth component of surrounding environment, such as the entire solvent or protein. This means that in eq 6 ΔE_μ and ΔE are replaced by $E_{\mu,\mathrm{int}}$ and E_{int}, respectively. The correlation function $\langle E_{\mu,\mathrm{int}}(t)E_{\mathrm{int}}(0)\rangle$ and the energy gap correlation function (eq 6) give the same type of information concerning fluctuations in the electrostatic field, acting on the residue due to the surroundings, the quantity of interest here; more explicitly, $\Delta E = \Sigma(q_{i,\mathrm{excited}} - q_{i,\mathrm{ground}})\phi_i$, where ϕ_i is the electrostatic potential that the ith charge experiences, while we use $\Delta E = \Sigma q_{i,\mathrm{ground}}\phi_i$. The *normalized* correlation functions $S(t)$ and $S_\mu(t)$ obtained for the present definition correspond exactly to the energy gap result for an the excited state that differs from the ground state by merely scaling all charges of the chromophore by the same factor. Moreover, if the transition dipole moment for the chromophore is collinear with the dipole moment of the ground state (in fact, a good approximation for the Aladan probe, for example[6]) and the dipole moment contribution is dominant, then the calculated solvation functions $S(t)$ and $S_\mu(t)$ obtained with the electrostatic interaction energy are the same as the ones obtained by using the usual definition of the energy gap function. (This is because the magnitude of the dipole moment does not appear in $S(t)$ or $S_\mu(t)$, due to normalization.) Most importantly, in experiments, the TDSS is used to infer the polar solvation dynamics in the *ground* state through a linear response approximation. In the present simulations, we calculate the polar solvation dynamics in the ground state directly, so that the choice of the probe is arbitrary. Using eq 7, the component solvation function $S_\mu(t)$ is determined from $C_\mu(t)$; see also the Appendix.

III. Results and Discussions

A. Relaxation Behavior for Different Positions in the Protein. Figure 1 shows the position of the probe sites. As can be seen, they are distributed throughout the protein with different burial (ranging from 5% to 45%) in different types of secondary structures (α-helix, β-sheet, or loop); see Table 1. Figure 3 shows log−log plots of $S(t)$ versus t.[33] In agreement with the experiments of Cohen et al.[6] the calculated relaxation dynamics are heterogeneous for different positions in the protein. The log−log plots make it possible to present all the results on comparable scales and, equally important, to be able to show clearly their multiexponential form. The plots also demonstrate that at long times the relative uncertainties are large, although they are small on an absolute scale. Most of the $S(t)$ functions have a simple form with a major decrease (of 85−99%) within a few picoseconds followed by much slower relaxation of the remaining tail; T2 and T16 are exceptions. To show the short time relaxation more clearly, we included linear plots next to the log−log plots. Within the accuracy of the simulation results, at least two exponentials are required. T16 is an exception, in that it shows a decay best described as a slowly decaying underdamped oscillator.[34] Fits of the relaxation behavior of $S(t)$ are given in Table 2. The $S(t)$ function decays on time scales varying from picoseconds to nanoseconds. The relaxation for 8 out of 11 sites can be described reasonably well with a sum of three exponentials corresponding to short, intermediate, and long time

TABLE 2: Parameters for Pure and Stretched/Compressed Multiexponential Fits: $\sum a_i \exp(-t/\tau_i)$ and $\sum a_i \exp[-(t/\tau_i)^{b_i}]$, respectively.

residue	a_1	τ_1 (ps)	a_2	τ_2 (ps)	a_3	τ_3 (ps)	b^a
T2	0.74	0.18	0.26	380			
T2	0.39	0.073	0.14	0.64	0.29	0.93	$b_3 = 0.30^b$
T11	0.63	0.065	0.32	1.39	0.05	100	
T11	0.26	0.14	0.72	0.31	0.02	230	$b_2 = 0.32$
T16						>1000	
T17	0.66	0.064	0.31	1.3	0.03	170	
T17	0.34	0.061	0.63	0.41	0.03	190	$b_2 = 0.46$
T18	0.63	0.073	0.34	1.98	0.03	300	
T18	0.30	0.060	0.68	0.62	0.02	370	$b_2 = 0.44$
T25	0.66	0.045	0.31	1.4	0.03	62	
T25	0.60	0.040	0.38	1.1	0.02	68	$b_2 = 0.74$
T44	0.72	0.084	0.25	3.9	0.03	190	
T44	0.40	0.065	0.58	0.81	0.02	260	$b_2 = 0.38$
T49	0.50	0.061	0.30	2.2	0.20	65	
T49	0.56	0.086	0.33	5.4	0.11	120	$b_3 = 3.2$
T51	0.64	0.091	0.30	3.1	0.06	270	
T51	0.39	0.068	0.56	1.1	0.05	290	$b_2 = 0.48$
T53	0.76	0.076	0.20	7.3	0.04	810	
T53	0.41	0.077	0.57	0.67	0.02	1300	$b_2 = 0.26$
T55	0.66	0.067	0.29	1.9	0.05	210	
T55	0.54	0.056	0.41	1.2	0.05	220	$b_2 = 0.64$

[a] The other stretching/compression parameters b are set to 1. [b] $a_4 = 0.18$, $\tau_4 = 550$ ps, $b_4 = 2.1$.

relaxation behaviors. In these cases, a slight improvement to the fits, clearly visible on the log−log scale, can be achieved by using a stretched exponential for the intermediate time relaxation. Interestingly, the long time relaxation for residues T2 and T49 decays faster than exponential and can be fitted reasonably well with "compressed" exponential functions ($b > 1$) (see also below).[35] Due to the slow convergence of the relaxation for residues T2, T16, and T49 only estimates of relaxation rates could be made.

Unlike Bagchi and co-workers,[18] we have not found correlation between the relaxation time scale and the exposure to the solvent. No such correlation is apparent even for residues whose dominant relaxation is due to solvent (see section IIIB). The results also show that the relaxation rates vary substantially within the same secondary structure elements; e.g., T16, T17, and T18 are all in the same β-strand, yet the relaxation dynamics differ by an order of magnitude. This is in accord with the fact that for probed residues the relaxation is converged (to 10−20%) within an interaction radius of 5 Å (Figure 5). Hence, the probe senses interactions primarily with the local environment. This observation is consistent with the estimate of the range of the chromophore environment interactions obtained from pressure-tuning hole-burning experiments.[36]

Figure 6 shows the fluctuations of electrostatic interaction energies of several residues (T2, T16, T25, and T49) with the environment. The fluctuations of T2, T16, and T49 are much slower than those of T25; the protein switches between long-lived states with different interaction energies. The origin of this behavior is analyzed in the next section. Infrequent occurrences of such switches lead to broken ergodicity on the accessible time scales, which results in slow convergence and anomalous relaxation of $S(t)$ at long times. Compressed exponential behavior has been observed in light-scattering experiments on jammed soft materials such as colloidal gels, concentrated emulsions, and micellar polycrystals.[35]

B. Contributions to the Polar Solvation. To understand the origin of the polar solvation for Thr residues, we decompose $S(t)$ based on eq 7. For each of the Thr residues we determined the contributions due to interactions with the protein (P), the entire solvent (water + ions) (W), those from water molecules hydrogen-bonded to the probe residue (B_1), and those hydrogen-bonded to any part of the protein (B_2). A hydrogen bond is assumed to be formed if the distance between hydrogen bond donor and hydrogen is less than 2.4 Å and hydrogen bond donor−hydrogen−hydrogen bond acceptor angle is more than 120°. We focus on four residues (T16, T17, T18, and T25) that illustrate the types of behavior observed for all of the residues. We use linear scales here with the time adjusted to correspond to the relaxation rate; for log−log plots, see Figure S1 of the Supporting Information. The corresponding decomposition of the relaxation for the remaining residues is also given in the Supporting Information (Table S1 and Figures S2−S5).

The results for T17 (Figure 7) are very similar to those obtained by Nilsson and Halle.[19] The slow relaxation is

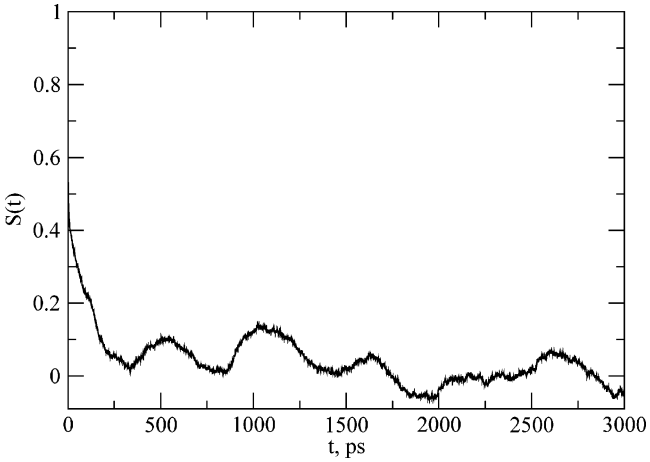

Figure 4. $S(t)$ for T16 on a linear scale. We plot T16 separately because it shows underdamped behavior that is very different from that of other residues (see text and Figure 8).

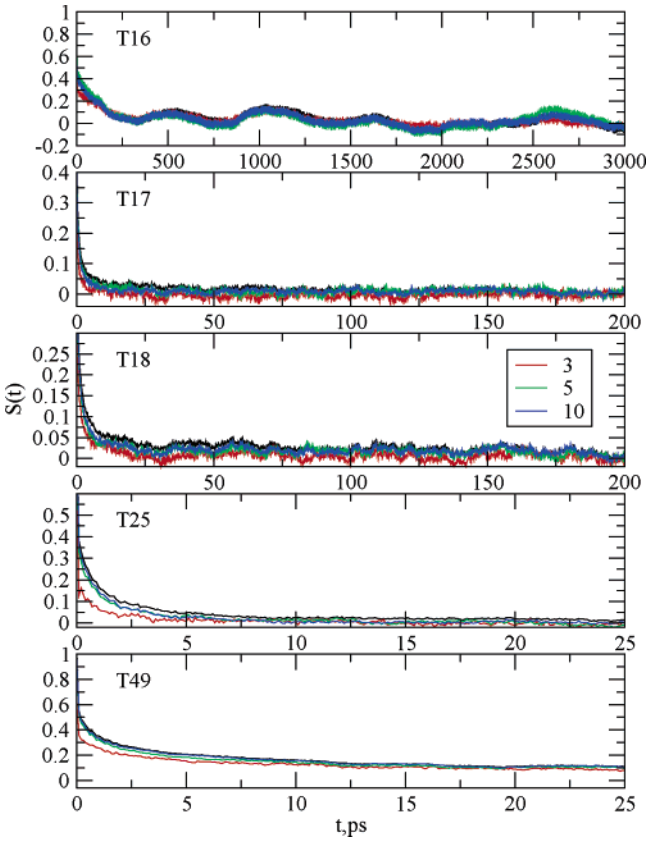

Figure 5. Convergence of $S(t)$ with cutoff (in Å).

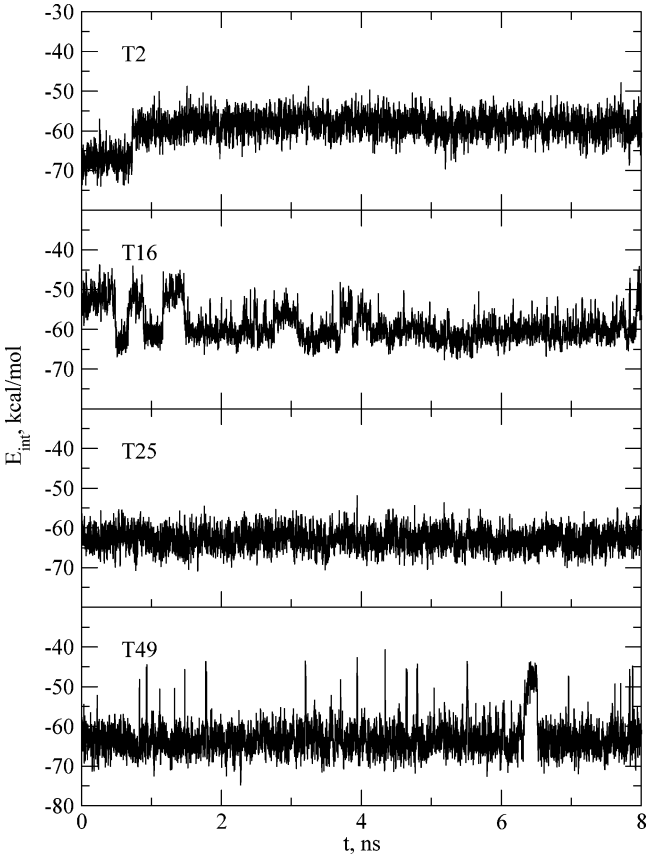

Figure 6. Fluctuation of electrostatic interaction energies of residues T2, T16, T25, and T49 with the environment. The fluctuations are smoothed by using running averages with 2 ps windows.

dominated by protein contributions (>200 ps for T17), while the fast relaxation (within a picosecond) is due to the contributions of the solvent. Clearly, the origin of the long time relaxation of these residues is due to not hydration dynamics but rather to protein motion. However, such a simple separation of time scales does not always hold. For example, for T25 (Figure 7), the contributions due to the protein (3 ps) and dominant contributions due to the solvent (2.2 ps) are on approximately the same time scale.

For T16 and T18 (Figure 7), the dominant contributions to $S(t)$ are from interactions with the solvent and mainly with waters that are hydrogen-bonded to the probed residues (B_1); waters hydrogen-bonded to other residues do not make a significant contribution. For these residues we examined whether there is a relationship between the slow relaxation time and the "escape time" of long resident ("bound") water molecules into "bulk" states. To estimate the escape times, we follow the trajectory of each water molecule from the moment it has formed hydrogen bonds with a given residue until its does not have any protein contacts within a hydration radius of 4.75 Å[37] ("bulk" state). The escape time along that trajectory is the time that it takes the hydrogen-bonded molecule to enter the bulk. If the molecule returns from the "bulk" state and forms a hydrogen bond again with the same residue, then it is treated as a new molecule from the moment it has formed this hydrogen bond and the counter is restored. Histograms of escape times for residues T16 and T18 are given in the Supporting Information (Figures S6 and S7). We determine that these "bound" water molecules are relatively tightly held by the protein with mean escape times of 101 and 24 ps for T16 and T18 residues, respectively. This is shorter than the observed relaxation rates given above, particularly for T16; the relaxation times are greater

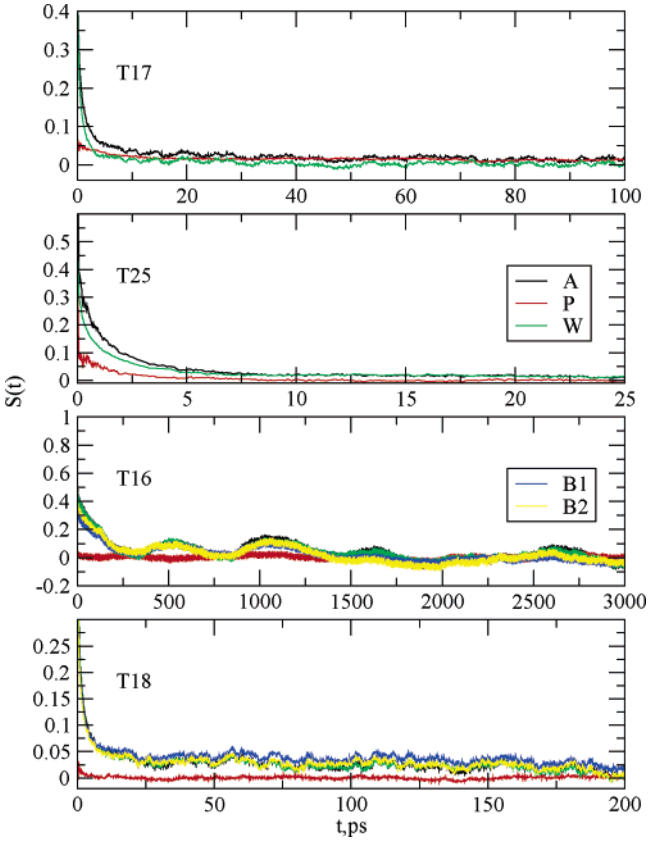

Figure 7. Contributions to $S(t)$ for indicated residues due to interactions with protein (P), the entire solvent (water + ions) (W), and both protein and solvent (A). For T16 and T18 the contributions of B_1 and B_2 (see text) are given as well.

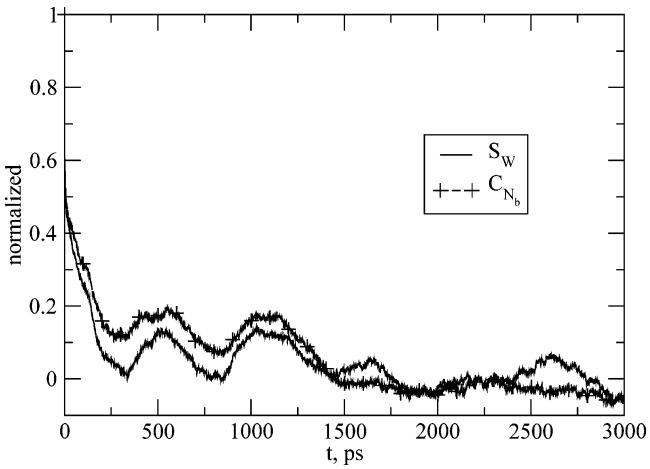

Figure 8. Normalized correlation function of the number of water molecules hydrogen-bonded to T16 and the solvent contribution $S_W(t)$ to the correlation function $S(t)$.

than 1 ns and 300 ps, respectively. Thus, the $S(t)$ relaxation times do not reflect the rate of exchange of a given water molecule between "bound" and "bulk" water states, as assumed in some analyses.[2,12]

From the time series in Figure 9, it is evident that the positive excursion in the electric field, relative to the average behavior, is due to the binding of a reduced number of water molecules. Consequently, the correlation function, C_{N_b}, of the fluctuations of the number of water molecules hydrogen-bonded to the residue T16 evolves on the same time scale as $S(t)$ (Figure 8). The slow N_b fluctuations could be due, in principle, to either slow exchange of "bound" with "bulk" water molecules or

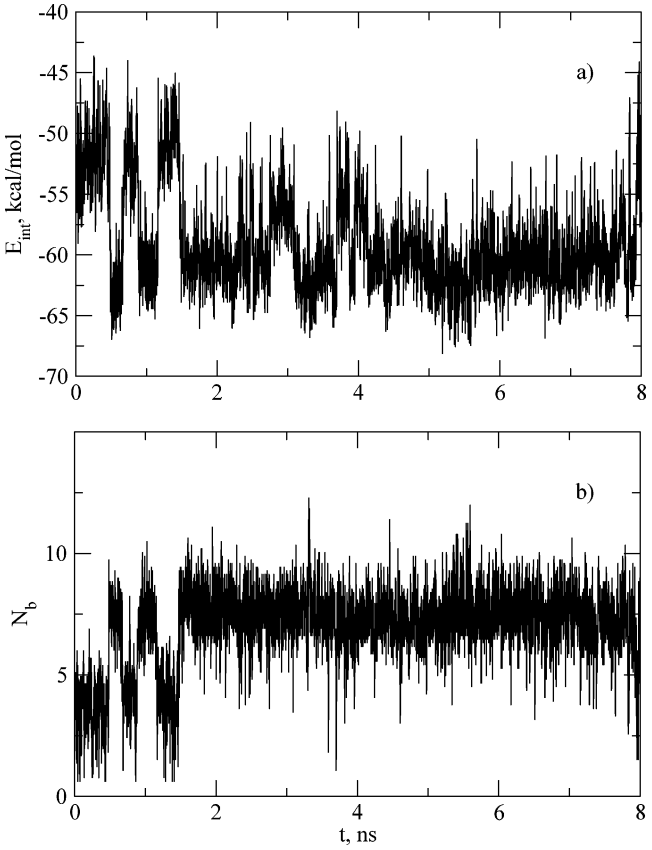

Figure 9. (a) Fluctuations of electrostatic interaction energy of the residue T16 with the environment. (b) The number of water molecules hydrogen-bonded to T16. The fluctuations are smoothed by using running averages with 2 ps windows.

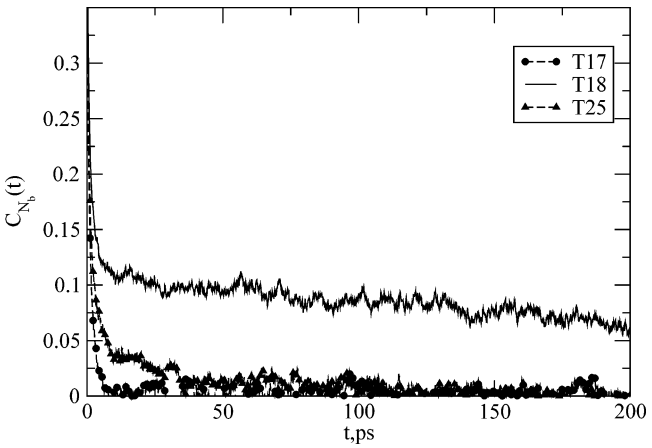

Figure 10. Normalized correlation functions of the fluctuations of the number of water molecules hydrogen-bonded to residues T17, T18, and T25 (N_b). Means (standard deviation) of N_b are 11.5 (2.7), 5.4 (1.9), and 4.7 (2.0), respectively.

exchange among protein conformations with different numbers of "bound" water molecules. Since the exchange between "bulk" and "bound" waters occurs on a much shorter time scales (see above), only the latter cause remains. This is consistent with the fact that for T17 and T25, where the solvent component $S(t)$ exhibits very short relaxation (Figure 10), C_{N_b} relaxes very quickly (within 5 and 50 ps, respectively), while for T18, where the solvent component $S(t)$ exhibits very long relaxation, C_{N_b} relaxes on a slower time scale.

IV. Conclusions

By using 11 different Thr residues in the immunoglobulin binding domain of protein G as probes for molecular dynamics simulations of polar solvation, we are able to obtain an understanding of the variation in the relaxation behavior and the relative contributions of the protein and the aqueous solvent. The polar solvation dynamics have been found to be heterogeneous and dependent on the location in the protein. There was no apparent dependence on accessible surface area. The probes have been found to report on interactions within 5 Å vicinity (at least for the probes studied in detail here).

We have found that there are cases when polar solvation relaxation is very slow and is dominated by interactions with water molecules hydrogen-bonded to a given probe. Although in these cases water molecules have been found to have long residence times (in the range of 20−100 ps depending on the site), the slow polar solvation relaxation does not represent the lifetime of "bound'' water molecules. Instead, it corresponds to interconversion among protein conformations with different numbers of "bound'' water molecules; i.e., there is a coupling between hydration and conformational dynamics. As a result, the interpretation of TDSS measurements is rather complex, making difficult experimental decomposition of the contributions to TDSS relaxation such as is available from the simulations. Variation of the solvent (as done in ref 4) would be of a great value in this respect.

Appendix A: Derivation of the Component Time-Dependent Stokes Shift

In this section, we derive the relationship between the component time-dependent Stokes shift and the classical component energy gap fluctuation correlation function given by eq 7. The analogous relationship between the full time-dependent Stokes shift and the classical energy gap fluctuation correlation functions (eq 3) has been used extensively in the past. The mere difference in index of μ between eq 3 and eq 7 is deceptive. There are many ways that $C(t)$ can be split into its components. For example, $C(t)$ was usually split into protein−protein, water−water, and water−protein components, i.e., $C(t) = C_{p,p} + C_{w,w} + 2C_{p,w}$, where $C_{\mu,\mu'}(t) = \langle \Delta E_\mu(t) \Delta E_{\mu'}(0) \rangle - \langle \Delta E_\mu \rangle \langle \Delta E_{\mu'} \rangle$. However, only $C_\mu(t)$ (eq 6) directly corresponds to $S_\mu(t)$. In the original work by Nilsson and Halle[19] eq 7 was used for the first time in the context of the TDSS. However, they did not derive eq 7 but rather referred to a lengthy review of Bernard and Callen.[32] We present here a much simpler derivation of eq 7 using the modern language of time-ordered exponentials and the notion of Kubo transformed correlation functions. As an intermediate result, we obtain a quantum-mechanical analog of eq 7, which is applicable to quantum-dynamics simulations. Our derivation is based on an elegant and concise derivation of eq 3 by Chernyak and Mukamel.[30]

We begin with the positive and negative time-ordered exponentials (W_+ and $W\_$, respectively) defined as in chapter 2 of ref 31. For perturbative approaches, it is useful to split the Hamiltonian into the zeroth-order Hamiltonian H_0 and the perturbation V. The forward propagator can then be written as

$$e^{-i/\hbar(H_0+V)t} = e^{-i/\hbar H_0 t} W_+(t) \tag{A1}$$

By differentiating both sides of this equation, we obtain

$$W_+(t) = -\frac{i}{\hbar} V(t) W_+(t) \tag{A2}$$

with

$$V(t) = e^{i/\hbar H_0 t}\, V\, e^{-i/\hbar H_0 t} \tag{A3}$$

Iterative solution of eq A2 gives

$$W_+(t) = 1 - \frac{i}{\hbar}\int_0^t dt_1\, V(t_1) + \left(\frac{i}{\hbar}\right)^2 \int_0^t dt_2 \int_0^{t_2} dt_1\, V(t_2)V(t_1) \ldots . \tag{A4}$$

This series is commonly called a positive time-ordered exponential and is denoted by

$$W_+(t) = \exp_+\left(-\frac{i}{\hbar}\int_0^t d\tau\, V(\tau)\right) \tag{A5}$$

Similarly, the backward propagator can be represented as

$$e^{i/\hbar(H_0+V)t} = W_-(t)e^{i/\hbar H_0 t} \tag{A6}$$

with

$$\dot{W}_-(t) = \frac{i}{\hbar}\, W_-(t)V(t) \tag{A7}$$

The iterative solution of this differential equation gives

$$W_-(t) = 1 + \frac{i}{\hbar}\int_0^t dt_1\, V(t_1) + \left(\frac{i}{\hbar}\right)^2 \int_0^t dt_2 \int_0^{t_2} dt_1\, V(t_1)V(t_2) \ldots . \tag{A8}$$

This series is commonly called a negative time-ordered exponential and is denoted by

$$W_-(t) = \exp_-\left(\frac{i}{\hbar}\int_0^t d\tau\, V(\tau)\right) \tag{A9}$$

The expression for the component time-dependent Stokes shift (eq 5)

$$\Delta\nu_\mu = \frac{1}{h}\, Tr[U_\mu\, e^{-iH_e t/\hbar}\, \rho_g\, e^{iH_e t/\hbar}] \tag{A10}$$

(it is assumed that $Tr\rho_g = 1$) can be written in terms of the energy gap fluctuation operators

$$\Delta U = U - \langle U\rangle \tag{A11}$$

$$\Delta U_\mu = U_\mu - \langle U_\mu\rangle \tag{A12}$$

as

$$h\Delta\nu_\mu(t) - h\Delta\nu_\mu(0) = Tr[e^{i/\hbar(H_g+\Delta U)t}\, \Delta U_\mu\, e^{-i/\hbar(H_g+\Delta U)t}\, \rho_g] \tag{A13}$$

where $\langle A\rangle = TrA\rho_g$.

Through the use of eqs A1 and A6, it is convenient for perturbative treatment to express this equation in terms of time-ordered exponentials in eqs A5 and A9, analogous to eqs A5 and A9

$$\delta E_\mu = \left\langle \exp_-\left(\frac{i}{\hbar}\int_0^t d\tau\, \Delta U(\tau)\right)\Delta U_\mu(t)\, \exp_+\left(-\frac{i}{\hbar}\int_0^t d\tau\, \Delta U(\tau)\right)\right\rangle \tag{A14}$$

where $\delta E_\mu = h\Delta\nu_\mu(t) - h\Delta\nu_\mu(0)$ and

$$\Delta U_\mu(t) = e^{i/\hbar H_g t}\, \Delta U_\mu\, e^{-i/\hbar H_g t}$$

$$\Delta U(t) = e^{i/\hbar H_g t}\, \Delta U\, e^{-i/\hbar H_g t} \tag{A15}$$

By using the expansions of the time-ordered exponentials given in eqs A4 and A8 and retaining only the lowest-order terms in ΔU, the relationship between component TDSS and component response function χ_μ is obtained as

$$\delta E_\mu = \int_0^t d\tau\, [\chi_\mu(t-\tau)] \tag{A16}$$

where

$$\chi_\mu(t-\tau) = -\frac{i}{\hbar}\langle[\Delta U_\mu(\tau), \Delta U(\tau)]\rangle$$

corresponding to the analogous relationship for the full time-dependent Stokes shift, eq A16 is exact when U is a Gaussian variable.[30]

We now define the quantum component energy gap operator correlation function as

$$C_\mu(t) = \langle\Delta U_\mu(t)\Delta U(0)\rangle \tag{A17}$$

Since ΔU_μ and ΔU are Hermitian, the component response function can be written as

$$\chi_\mu(t) = -\frac{i}{\hbar}\,(C_\mu(t) - C_\mu^*(t)) \tag{A18}$$

A symmetrized quantum analog of the classical correlation function can be introduced as

$$C_\mu^S(t) = \frac{C_\mu(t) + C_\mu^*(t)}{2} \tag{A19}$$

Using an energy representation one can show that

$$\tilde{C}_\mu^*(-\omega) = e^{-\beta\hbar\omega}\tilde{C}_\mu(\omega) \tag{A20}$$

where $\beta = (1/k_B T)$ and the Fourier transform is defined as

$$\tilde{A}(\omega) = \int_{-\infty}^{\infty} dt\, e^{i\omega t}\, A(t) \tag{A21}$$

Therefore

$$\tilde{\chi}_\mu(\omega) = -\frac{i}{\hbar}\,(1 - e^{-\beta\hbar\omega})\tilde{C}_\mu(\omega)$$

$$\tilde{C}_\mu^S(\omega) = \frac{1}{2}\,(1 + e^{-\beta\hbar\omega})\tilde{C}_\mu(\omega) \tag{A22}$$

and the fluctuation−dissipation relation follows

$$\tilde{\chi}_\mu(\omega) = -\frac{2i}{\hbar}\tanh\frac{\beta\hbar\omega}{2}\,\tilde{C}_\mu^S(\omega) \tag{A23}$$

By substituting this expression into eq A16, one obtains

$$\delta E_\mu = \beta\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\,(e^{-i\omega t} - 1)\tilde{C}_\mu^K(\omega) \tag{A24}$$

with

$$\tilde{C}_\mu^K(\omega) = \frac{\tanh\dfrac{\beta\hbar\omega}{2}}{\dfrac{\beta\hbar\omega}{2}}\,\tilde{C}_\mu^S(\omega)$$

where $\tilde{C}^{K}_{\mu}(\omega)$ is a Fourier transform of the Kubo transformed[38] correlation function given by

$$C^{K}_{\mu,g}(t) = \frac{1}{\beta}\int_0^{\beta} d\lambda \, \langle \Delta U_{\mu}(t)\Delta U(i\hbar\lambda)\rangle \qquad \text{(A25)}$$

This shows that

$$h\Delta\nu_{\mu}(t) - h\Delta\nu_{\mu}(\infty) = \beta C^{K}_{\mu,g}(t) \qquad \text{(A26)}$$

where g indicates that the thermal average is taken over the ground state.

Similarly, by writing

$$e^{-\beta H_g} = e^{-\beta(H_e - U)} \qquad \text{(A27)}$$

and expanding in powers of $\Delta U' = U - \langle U\rangle_e$, one can show that a corresponding expression holds for the thermal average over excited state, i.e,

$$h\Delta\nu_{\mu}(t) - h\Delta\nu_{\mu}(\infty) = \beta C^{K}_{\mu,e}(t)$$

where e indicates that the thermal average is taken over the excited state. By taking the high-temperature limit, eq 7 follows from eqs A24, A26, and A28

Acknowledgment. We thank Steven Boxer, Bruce Cohen, Lennart Nilsson, Wei Yang, Paul Maragakis, Martin Spichty, Isaac Sztainbuch, and Ivana Adamovic for stimulating discussions. A seminar by Steven Boxer at Harvard University initiated our interest in this problem. NAMD was developed by the Theoretical and Computational Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign. The computations were performed in part on the Crimson grid and the Bauer Computer Center. This work was supported in part by an National Institutes of Health (NIH) grant to M. K. and a NIH National Research Service Award fellowship F32 GM073324 to AAG.

Supporting Information Available: Figures and a table characterizing protein and solvent contributions to $S(t)$. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Fleming, G. R.; Cho, M. *Annu. Rev. Phys. Chem.* **1996**, *47*, 109−134.
(2) Pal, S.; Zewail, A. *Chem. Rev.* **2004**, *104*, 2099−2123.
(3) Karplus, M.; Kuriyan, J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6679−6685.
(4) Lampa-Pastirk, S.; Beck, W. *J. Phys. Chem. B* **2004**, *108*, 16288−16294.
(5) Lu, W.; Kim, J.; Qiu, W.; Zhong, D. *Chem. Phys. Lett.* **2004**, *388*, 120−126.
(6) Cohen, B. E.; McAnaney, T. B.; Park, E. S.; Jan, Y. N.; Boxer, S. G.; Jan, L. Y. *Science* **2002**, *296*, 1700−1703.
(7) Cohen, B.; Pralle, A.; Yao, X.; Swaminath, G.; Gandhi, C.; Jan, Y.; Kobilka, B.; Isacoff, E.; Jan, L. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 965−970.
(8) Changenet-Barret, P.; Choma, C.; Gooding, E.; DeGrado, W.; Hochstrasser, R. M. *J. Phys. Chem. B* **2000**, *104*, 9322−9329.
(9) Pal, S.; Zhao, L.; Zewail, A. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 8113−8118.
(10) Kamal, J. K. A.; Zhao, L.; Zewail, A. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 13411−13416.
(11) Pal, S.; Peon, J.; Zewail, A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 15297−15302.
(12) Pal, S.; Peon, J.; Bagchi, B.; Zewail, A. *J. Phys. Chem. B* **2002**, *106*, 12376−12395.
(13) Frenkel, D.; Smit, B. *Understanding Molecular Simulations*; Academic Press: San Diego, 1996.
(14) Allen, M.; Tildesley, D. *Computer Simulations of Liquids*; Oxford University Press: New York, 1987.
(15) Ladanyi, B. M.; Maroncelli, M. *J. Chem. Phys.* **1998**, *109*, 3204.
(16) Faeder, J.; Ladanyi, B. *J. Phys. Chem. B* **2001**, *105*, 11148.
(17) Vivian, J. T.; Callis, P. R. *Biophys. J.* **2001**, *80*, 2093−2109.
(18) Bandyopadhyay, S.; Chakraborty, S.; Balasubramanian, S.; Bagchi, B. *J. Am. Chem. Soc.* **2005**, *127*, 4071−4075.
(19) Nilsson, L.; Halle, B. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13867−13872.
(20) Pal, S.; Peon, J.; Zewail, A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 10964−10969.
(21) Halle, B. *Phil. Trans. R. Soc. London, Ser. B* **2004**, *359*, 1207−1224.
(22) MacKerell, A. D., Jr.; Bashford, D.; Bellott, R. L.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R. Straub, J.; Watanbe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(23) Brooks, C. L., III.; Karplus, M. *J. Chem. Phys.* **1983**, *79*, 6312−6325.
(24) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(25) Kale, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. *J. Comput. Phys.* **1999**, *151*, 283−312.
(26) Brooks, B.; Bruccoleri, R.; Olafson, B.; Olafson, B.; States, D.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(27) Essmann, U.; Perera, L.; Berkowitz, M.; Darden, T.; Lee, H.; Pedersen, L. *J. Chem. Phys.* **1995**, *103*, 8577−93.
(28) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(29) Feller, S.; Zhang, Y.; Pastor, R.; Brooks, B. *J. Chem. Phys.* **1995**, *103*, 4613−4621.
(30) Chernyak, V.; Mukamel, S. *J. Chem. Phys.* **2001**, *114*, 10430−10433.
(31) Mukamel, S. *Principles of Nonlinear Optical Spectroscopy*; Oxford University Press: New York, 1995.
(32) Bernard, W.; Callen, H. B. *Rev. Mod. Phys.* **1959**, *31*, 1017−1044.
(33) McMahon, B.; Frauenfelder, H. The Energy Landscape. In *Single Molecule Spectroscopy: Nobel Symposium Lectures*; Rigler, R., Orrit, M., Basche, T., Eds.; Springer: New York, 2001; pp 257−276.
(34) Swaminathan, S.; Ichiye, T.; Gunsteren, W. V.; Karplus, M. *Biochemistry* **1984**, *21*, 5230−5241.
(35) Cipelletti, L.; Ramos, L.; Manley, S.; Pitard, E.; Weitz, D. A.; Pashkovski, E. E.; Johansson, M. *Faraday Discuss.* **2003**, *123*, 237−251.
(36) Lesch, H.; Schlichter, J.; Friedrich, J.; Vanderkooiy, J. M. *Biophys. J.* **2004**, *86*, 467−472.
(37) Boresch, S.; Willensdorfer, M.; Steinhauser, O. *J. Chem. Phys.* **2003**, *120*, 3333−3347.
(38) Kubo, R. *J. Phys. Soc. Jpn.* **1957**, *12*, 570.