Spatial Dependence of Protein-Water Collective Hydrogen-Bond Dynamics

Matthias Heyden* and Douglas J. Tobias†

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

(Received 16 December 2012; revised manuscript received 10 May 2013; published 22 November 2013)

Using molecular dynamics simulations, we analyze collective vibrations in the hydration water of a small globular protein. We develop tools that allow spatial resolution of correlated protein and water motion, and use them to reveal correlated vibrations that extend up to 10 Å from the protein surface at far-infrared/THz frequencies that are sensitive to the chemical properties of the protein surface. Our results provide the first detailed description of long-range effects on protein hydration water dynamics and highlight the differences between single particle and collective dynamics, which are relevant in interpreting experimental observations.

DOI: 10.1103/PhysRevLett.111.218101 PACS numbers: 87.15.kr, 83.10.Rs, 87.15.A−

In aqueous solution, protein-induced effects on hydration water properties are typically described as short-range, i.e., restricted to the first one or two hydration layers, when structural properties and single particle dynamics are considered. The latter can be analyzed conveniently in molecular dynamics (MD) simulations [1–4] by following the trajectories of individual water molecules, and the results can be related to experimental observables, e.g., incoherent neutron scattering and NMR relaxation [5–7]. Single particle dynamics are often characterized in terms of single molecule diffusion rates, rotational correlation times, and the vibrational density of states (VDOS).

Data from recent experiments based on the far-infrared/THz absorption of aqueous solutions of biomolecular solutes suggest that single particle dynamics might not be sufficient to describe completely the coupling of protein and solvent dynamics, which is recognized to play a crucial role for protein folding and function [8–11]. THz experiments [12–14] probe the intermolecular vibrations of the hydrogen bond network in the hydration water of globular proteins. Recent THz studies revealed a change of the hydration water absorption coefficient within 10 Å around fully dissolved proteins, exceeding the first and second hydration layers. Initially, this experimental observation was interpreted primarily in terms of the retardation of single particle picosecond dynamics, despite the restriction of the latter to the first hydration layers. A recent *ab initio* MD simulation study on bulk water [15,16] explicitly analyzed vibrational motion at various frequencies in water and described the collective character of low-frequency intermolecular vibrational modes in the far-infrared/THz spectrum in terms of correlated motions extending over many molecules. This result suggests that the observation of long-range effects on the THz absorption of intermolecular vibrations in the hydration shell of a protein may be related to altered collective dynamics due to solute-solvent interactions. In this Letter, we extend previous methodology and apply it to an MD simulation of a protein to spatially resolve the spectrum of collective vibrations in the hydration water and to determine the length scale on which correlations of vibrational motion between protein atoms and their hydration water persist. We chose the N-terminal DNA-binding domain of the λ-repressor protein, which was employed in a previous experimental THz spectroscopy study [12], in order to compare our results to the experimental interpretation in terms of an extended hydration shell with a nonbulk absorption at far-infrared frequencies. In addition, we characterize the observed collective modes and analyze their sensitivity to specific protein-water interactions.

To analyze correlated vibrational motion between atoms of the protein surface and hydration water as a function of separation distance, we formulate a generalized VDOS that describes correlated vibrational motion. The standard definition of the VDOS is based on the Fourier transform of the autocorrelation function of atomic velocities

$$I(\omega) = \int e^{i\omega t}\langle \tilde{\mathbf{v}}(0)\cdot\tilde{\mathbf{v}}(t)\rangle dt, \quad (1)$$

where $\tilde{\mathbf{v}} = \sqrt{m}\mathbf{v}$ are mass-weighted atomic velocities, and the angular brackets denote an average over atoms and time origins. In the generalized version employed herein, we extend the definition to include velocity cross correlations as a function of separation distance between two atoms

$$I(\omega, r = |\mathbf{r}_j - \mathbf{r}_i|) = \int e^{i\omega t}\langle \tilde{\mathbf{v}}(0,\mathbf{r}_i)\cdot\tilde{\mathbf{v}}(t,\mathbf{r}_j)\rangle dt, \quad (2)$$

where $\mathbf{r}_i$ is the position of atom i. In practice, such a formulation of a spatially resolved cross-correlation spectrum is problematic for liquids at room temperature, and especially for the relatively low frequency motions of interest here, because molecular diffusion leads to ill-defined intermolecular separations on the time scale needed to sample the correlation function. To circumvent this problem, following previous work on bulk water [15,16], we utilize a localized, smooth density of the mass-weighted atomic velocities, in which the infinitely sharp delta function representing instantaneous atomic positions is

replaced by a three-dimensional Gaussian distribution with an appropriate width (e.g., $\sigma = 0.4$ Å [15,16])

$$\boldsymbol{\rho}_{\tilde{\mathbf{v}}}(t,\mathbf{r}) = \sum_i \tilde{\mathbf{v}}_i \delta(|\mathbf{r}_i(t) - \mathbf{r}|) \approx \sum_i \tilde{\mathbf{v}}_i \frac{1}{(2\pi\sigma^2)^{3/2}} e^{-(|\mathbf{r}_i(t)-\mathbf{r}|^2/2\sigma^2)}. \quad (3)$$

While the localized density does not necessarily contain contributions from a single atom or molecule (however, at each instant of time the contribution of the closest atom will dominate for a properly chosen σ), its location in space is well defined. Using smooth velocity densities, we formulate a localized VDOS via

$$I(\omega,\mathbf{r}) = \int e^{i\omega t}\langle \boldsymbol{\rho}_{\tilde{\mathbf{v}}}(0,\mathbf{r})\boldsymbol{\rho}_{\tilde{\mathbf{v}}}(t,\mathbf{r})\rangle dt. \quad (4)$$

A comparison between the VDOS of bulk water oxygen atoms (> 10 Å from the closest nonhydrogen protein atom) obtained via Eq. (4) and the standard definition [Eq. (1)] in Fig. 1 shows a modest broadening of the hydrogen bond (HB) bending [17] signal around 50 cm^{-1} due to the use of the smooth velocity density. However, the main features of the low frequency spectrum, particularly the HB bending and HB stretching peak or shoulder at 200 cm^{-1} are well reproduced, demonstrating the validity of the approach. To analyze velocity correlations between atoms of the protein and hydrating water molecules, we first choose protein atoms from the protein–water interface and further subdivide them into surface patches formed by hydrophilic and hydrophobic amino acid side chains. The standard VDOS

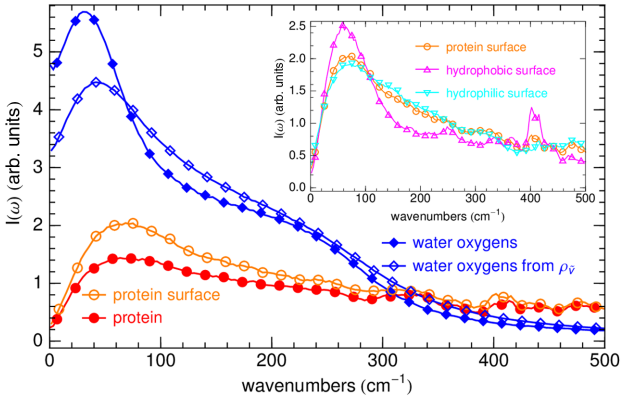

FIG. 1 (color online). Comparison of the VDOS obtained for bulk water oxygen atoms (> 10 Å distance to the protein) via the standard velocity autocorrelation formalism [Eq. (1), blue line, full diamonds] and the localized velocity density [Eq. (4), blue line, open diamonds]. In addition, the VDOS obtained for nonhydrogen protein atoms (red, full circles) and the protein-water interface (orange, open circles) are shown. Inset: Comparison of the VDOS for the full protein surface (orange, open circles) and hydrophobic (magenta, open triangles, up) and hydrophilic (cyan, open triangles, down) patches. All VDOS shown are normalized by the area enclosed with the frequency axis.

[Eq. (1)] for nonhydrogen atoms (which dominate the low frequency spectrum as shown for water in Fig. S1 [18]) obtained for the total protein and its surface is shown in Fig. 1. For the hydrophobic surface of the protein, the inset of Fig. 1 demonstrates a shift of intensities from the 200 cm^{-1} region toward lower frequencies around 50 cm^{-1}, which can be attributed to the absence of protein-water HBs and their corresponding stretch vibrations. Otherwise, the computed spectra are qualitatively similar at frequencies below 300 cm^{-1} ($\approx$10 THz) and exhibit a significant overlap with the vibrational spectrum of the water HB network.

Atoms of the protein surface are fairly well localized over time scales needed to sample the required correlation functions for the Fourier transformation (tens of ps). Furthermore, any diffusional motion and rotation of the protein can be removed from the trajectory prior to analysis by applying a rotational-translational transformation that minimizes the root mean squared deviation of the protein backbone. Hence, replacement of atomic velocities by a localized velocity density is only required for hydration water oxygens to resolve spatial correlations. This density is sampled at selected points in space, which are chosen by their distance to the closest heavy atom of the respective protein surface (Fig. S2 [18]). Correlated vibrational motion between atoms of the protein surface and water are conveniently described as

$$I(\omega, r = |\mathbf{r}_P - \mathbf{r}_{OW}|) = \int e^{i\omega t}\langle \tilde{\mathbf{v}}_P(0,\mathbf{r}_P)\boldsymbol{\rho}_{\tilde{\mathbf{v}},OW}(t,\mathbf{r}_{OW})\rangle dt, \quad (5)$$

where the subscripts P and OW indicate protein atoms and water oxygens, respectively.

To facilitate analysis of the collective properties of the various vibrational motions, we display the computed correlation spectra as a function of the reciprocal distance to the protein surface, $k = 2\pi/r$, in Fig. 2 (the actual distance r is shown on the alternative x axis and an alternate representation of the same data as a function of the distance r is shown in Fig. S3 [18]). This representation allows us to observe dispersive behavior of collective vibrational modes via the dispersion, namely, a k dependence of the mode frequency ω at the maximum of the absolute intensity, which indicates a propagating wave with velocity $v_p = d\omega/dk$ [19]. Our interpretation is analogous to the analysis of the related density current spectra $I_J(k,\omega)$, which describe spatially correlated density fluctuations and can be determined from the experimentally accessible dynamic structure factor $S(k,\omega)$ via $I_J(k,\omega) = (\omega^2/k^2)S(k,\omega)$ [16,19]. The approach described here allows us to resolve spatial correlations of atomic velocities directly in real-space, as well as to select and analyze site–specific collective dynamics of the protein surface and its hydration water. Negative intensities indicate antiparallel velocity vectors characteristic of stretch vibrations, while positive intensities describe correlated in–phase vibrational motion in the same direction. Separate analysis

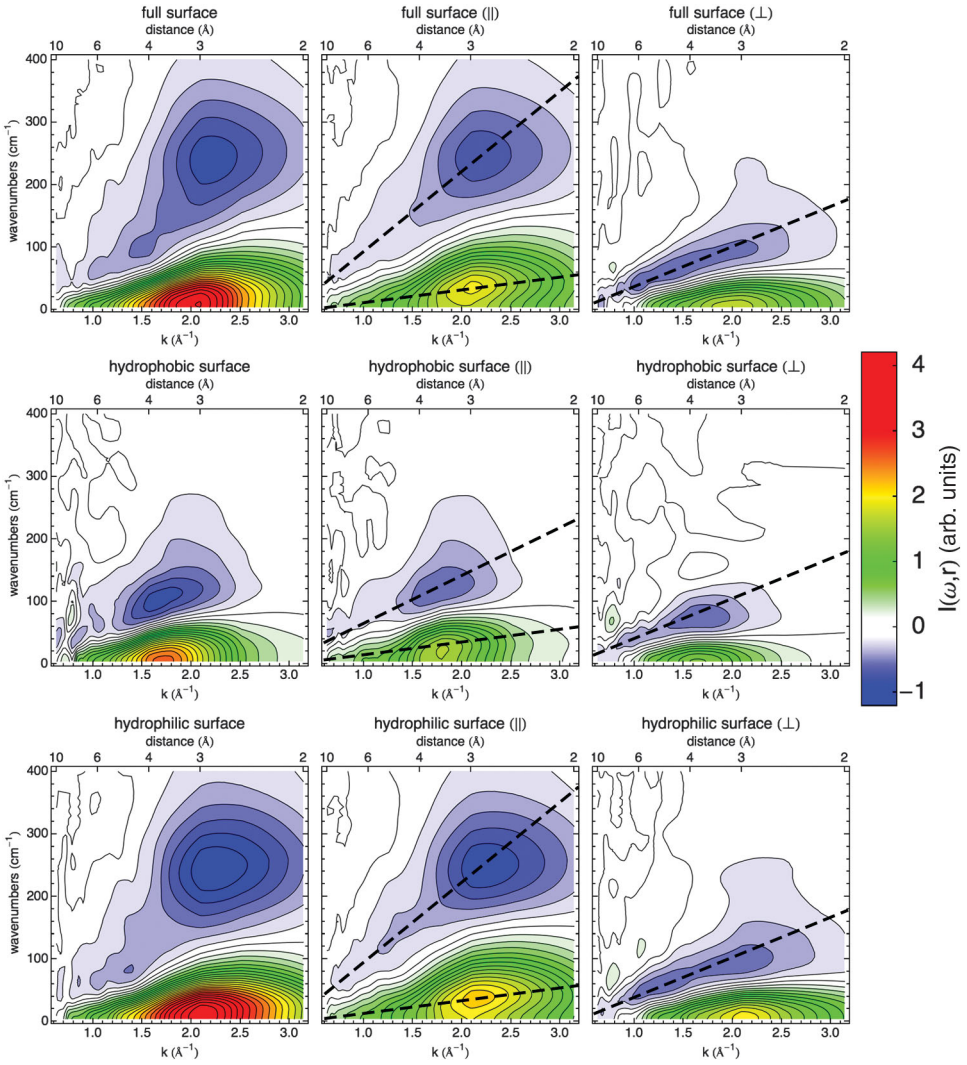

FIG. 2 (color). Cross-correlation spectra [Eq. (6)] showing correlated vibrational motion of nonhydrogen atoms of the protein surface with hydration water oxygens as a function of $k = 2\pi/r$, where r is the distance from the protein surface. Top, center, and bottom panels show the results for hydration water of the full protein, hydrophobic patches, and hydrophilic patches of the protein surface, respectively. The left column panels contain the results from overall correlations of atomic velocities, while the center and right columns display contributions from velocity components ∥ and ⊥ to the protein surface normal, respectively. Dashed lines denote corresponding dispersion curves for each dispersive mode, the slopes of which yield propagation velocities.

of velocity components parallel (∥) and perpendicular (⊥) to the protein surface normal allows us to distinguish longitudinal and transverse collective motion in the hydration water of the protein.

In the top row of panels in Fig. 2, the spectra of velocity cross correlations between protein surface atoms and hydration water, and separately, the ∥ and ⊥ components, are shown as a function of k. Two distinct ∥ modes that dominate the total spectrum can be identified: (1) a high frequency mode with negative intensities (HFN) originating in protein-water HB stretch vibrations, and (2) a low frequency mode with positive intensities (LFP) in a frequency range that resembles the regime of HB bending motions in water (Fig. 1). Both modes exhibit a maximum absolute intensity in the first hydration layer ($k = 2.1$–2.3 Å^{-1}; $r = 2.8$–3.0 Å) describing the initial intermolecular vibration between the protein and hydration water molecules. The propagation of these vibrations into higher hydration shells can then be followed by tracing the absolute intensity maximum towards larger separation distances from the protein surface, i.e., lower values of k, which are accompanied by a linear decrease in frequency in Fig. 2 in accord with dispersive behavior. The key observation is that the intensity of velocity cross correlations between protein surface atoms and hydration water oxygens remains non-negligible up to k values below 1 Å^{-1}, corresponding to distances of 8–10 Å from the protein surface (see also Fig. S3 [18]). At this distance, the frequencies of these correlated motions have decreased to below 100 cm^{-1} (3 THz), with intensities of the LFP and HFN modes

dominating frequencies below 1 THz and from 1 to 3 THz, respectively. We note that this observation coincides with apparently opposing experimental observations for the protein hydration shell absorptions at frequencies below 1 THz [20] (no apparent change of hydration shell THz absorption vs bulk water) and frequencies from 2.25 to 2.55 THz [12] (increased THz absorption in the hydration shell). Independent of the impact of the correlated vibrations associated with the LFP and HFN modes on apparent THz absorption coefficients in the hydration shell, our result demonstrates that correlated, collective motion involving protein atoms and hydration shell water molecules persists on a length scale that exceeds the first and second hydration layers. The k dependence of the protein-water collective vibrational mode frequencies, indicated by dashed lines in Fig. 2, allows us to determine wave propagation velocities of $\approx$2500 and 400 m/s, respectively, the former being slower than the fast sound propagation velocity of 3800 m/s obtained from longitudinal current spectra in protein hydration water alone in a comparable k range [19].

The $\perp$ components exhibit a dispersive mode with negative intensities, which originates at HB bending frequencies and also features non-negligible atomic velocity correlations at $k < 1$ Å^{-1}. Its dispersive behavior indicates a shear wave due to the vibrations perpendicular to the propagation direction (away from the protein surface), with a propagation velocity of $\approx$ 1250 m/s. At zero frequency, nondispersive diffusive motion gives rise to positive intensities in the cross-correlation spectrum that are restricted to $k > 1$ Å^{-1}.

In light of these results, the question arises whether the observed long-range correlations of vibrational motion and the collective dynamics are influenced by the protein, or if they are simply a manifestation of the average collective dynamics in the liquid medium [19]. To answer this question, we analyze separately the collective protein-hydration water dynamics in the vicinity of hydrophobic and hydrophilic parts of the protein surface (center and bottom rows of panels of Fig. 2).

The surface of the DNA-binding domain of the λ-repressor is primarily hydrophilic ($\approx$70% for a water–molecule sized probe, $\approx$ 85% for a probe size >6 Å, indicating a preference of solvent exposed hydrophobic residues for concave crevices on the protein surface; see Fig. S2). Hence, collective dynamics in the vicinity of hydrophilic side chains resemble the collective dynamics of the entire protein. However, in the vicinity of hydrophobic side chains, a distinct behavior is observed for the collective dynamics parallel to the protein surface normal. The absence of direct protein-water HBs at the hydrophobic surface results in a lower frequency HFN mode, which gives rise to a slower longitudinal wave propagating with 1500 m/s (vs 2500 m/s observed at hydrophilic parts of the protein surface). The modified characteristics of this mode indicate, therefore, that correlated vibrational and collective motions of the protein surface and its hydration water are affected by the chemical properties of the protein and, hence, are distinct from the average collective properties of the medium.

While the characteristics of the $\parallel$ LFP mode and the correlations of the $\perp$ components are not significantly different in the vicinity of the hydrophobic surface, their maximum absolute intensities are decreased relative to the hydrophilic case (on the order of 10%–25%), indicating weaker correlations. Additionally, the peak maxima for all modes are shifted to lower $k \approx 1.8$ Å^{-1}, corresponding to distances of roughly 3.5 Å for the hydrophobic surface (compared to $k = 2.1$–2.3 Å^{-1} and $r = 2.8$–3.0 Å for the hydrophilic surface) due to a larger preferred separation distance between water molecules in the first hydration shell and hydrophobic protein surfaces.

We conclude that vibrational motion in the hydration water HB network of a biomolecular solute cannot be understood solely in terms of single particle dynamics. The results in Fig. 2 and Fig. S3 [18] show clearly that the vibrational modes are collective and that correlated vibrational motion exists for atom separations up to 10 Å, in particular in the frequency window below 100 cm^{-1}. Thus, single particle dynamical parameters provide only a partial description of solute-induced effects on solvent dynamics in the vicinity of a protein.

The inability of single particle dynamics to capture the long-range correlations observed in Fig. 2 and Fig. S3 [18] is demonstrated in Fig. 3, where the localized VDOS from velocity autocorrelations are shown for water molecules in the hydration water of the protein as a function of the distance to the protein surface. Distortions from the bulk water spectrum are found only within the first hydration

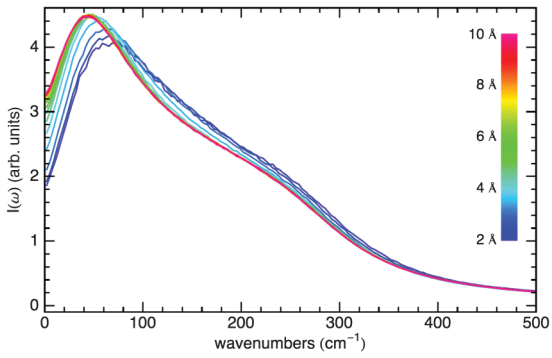

FIG. 3 (color). Single particle VDOS computed from autocorrelations of localized velocity densities [Eq. (4)] of hydration water oxygens as a function of distance to the protein surface. The data shown corresponds to distances from 2 to 10 Å in 0.5 Å steps according to the color code. Figure S3 [18] shows the equivalent result for hydration water of hydrophobic and hydrophilic parts of the protein surface, separated into the VDOS from velocity components $\parallel$ and $\perp$ to the protein surface normal. All VDOS shown are normalized by the area enclosed with the frequency axis.

layer, i.e., within 4 Å of the protein surface. This result is in accord with previous findings [21], as well as other experimental and theoretical data on single particle dynamics [2,4–7], and clearly indicates that the potential energy surface on which single particle dynamics occur (primarily determined by water structure) is indistinguishable from bulk beyond the first hydration layer.

The question remains how the observed protein-induced long-range collective dynamics affect hydration water vibrations if their vibrational spectrum, i.e., the VDOS, remain unchanged. We suggest that the observed correlated vibrations reflect a nonrandom phase relation between vibrational motions on the protein surface and the hydration water. While the intermolecular HB network vibrations of water are intrinsically collective, even in the bulk [15], the protein solute can act as a synchronizing element for the vibrations of water molecules in its hydration shell, with consequences, for example, for the frequency dependent dipole fluctuations in the protein hydration water compared to independent collective oscillators in the bulk. The latter provides the connection to the experimentally observed changes of THz absorption coefficients in hydration shells of several globular proteins up to 10 Å thickness [12–14], which cannot be explained satisfactorily by single particle dynamics. It was observed, furthermore, that the hydration water absorption at THz frequencies is sensitive to destabilizations of the native folded state [13]. Our analysis suggests that changes in the exposure of hydrophilic or hydrophobic groups on the protein surface due to a shift in the native $\rightleftharpoons$ partially unfolded equilibrium, will affect collective vibrations in the hydration water, in accord with this experimental observation, although other factors, such as protein flexibility or stiffness, are likely to play a role as well.

We expect that consideration of long-range correlated motion between proteins and their hydration water will be crucial for a complete understanding of protein dynamics in crowded environments such as the interior of the cell. In addition, a detailed description of collective dynamics may be relevant to predicting the optimal functioning of proteins in a specific environment.

M. H. acknowledges financial support from the German Academy of Sciences Leopoldina. This work is supported by the Cluster of Excellence RESOLV (EXC 1069) funded by the Deutsche Forschungsgemeinschaft.

*Present address: Max-Planck-Insitut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, D-45470 Mülheim an der Ruhr, Germany.
matthias.heyden@uci.edu
†dtobias@uci.edu

[1] B. Bagchi, Chem. Rev. **105**, 3197 (2005).
[2] R. Abseher, H. Schreiber, and O. Steinhauser, Proteins **25**, 366 (1996).
[3] M. Tarek and D. J. Tobias, Biophys. J. **79**, 3244 (2000).
[4] N. Sengupta, S. Jaud, and D. J. Tobias, Biophys. J. **95**, 5257 (2008).
[5] M. Settles and W. Doster, Faraday Discuss. **103**, 269 (1996).
[6] J. Qvist, E. Persson, C. Mattea, and B. Halle, Faraday Discuss. **141**, 131 (2009).
[7] A. Froelich, F. Gabel, M. Jasnin, U. Lehnert, D. Oeserhelt, A. M. Stadler, M. Tehei, M. Weik, K. Wood, and G. Zaccai, Faraday Discuss. **141**, 117 (2009).
[8] P. W. Fenimore, H. Frauenfelder, B. H. McMahon, and F. G. Parak, Proc. Natl. Acad. Sci. U.S.A. **99**, 16047 (2002).
[9] M. Chaplin, Nat. Rev. Mol. Cell Biol. **7**, 861 (2006).
[10] Y. Levy and J. N. Onuchic, Annu. Rev. Biophys. Biomol. Struct. **35**, 389 (2006).
[11] P. Ball, Chem. Rev. **108**, 74 (2008).
[12] S. Ebbinghaus, S. J. Kim, M. Heyden, X. Yu, U. Heugen, M. Gruebele, D. M. Leitner, and M. Havenith, Proc. Natl. Acad. Sci. U.S.A. **104**, 20749 (2007).
[13] S. Ebbinghaus, S. J. Kim, M. Heyden, X. Yu, M. Gruebele, D. M. Leitner, and M. Havenith, J. Am. Chem. Soc. **130**, 2374 (2008).
[14] B. Born, S. J. Kim, S. Ebbinghaus, M. Gruebele, and M. Havenith, Faraday Discuss. **141**, 161 (2009).
[15] M. Heyden, J. Sun, S. Funkner, G. Mathias, H. Forbert, M. Havenith, and D. Marx, Proc. Natl. Acad. Sci. U.S.A. **107**, 12068 (2010).
[16] M. Heyden, J. Sun, H. Forbert, G. Mathias, M. Havenith, and D. Marx, J. Phys. Chem. Lett. **3**, 2135 (2012).
[17] G.-E. Walrafen, J. Phys. Chem. **94**, 2237 (1990).
[18] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.111.218101 for a description of the protocol used for simulations of the solvated λ^*_{6-85}-repressor fragment and technical details of the analysis.
[19] M. Tarek and D. J. Tobias, Phys. Rev. Lett. **89**, 275501 (2002).
[20] J. Xu, K. W. Plaxco, and S. J. Allen, Protein Sci. **15**, 1175 (2006).
[21] S. Chakraborty, S. K. Sinha, and S. Bandyopadhyay, J. Phys. Chem. B **111**, 13626 (2007).