Solvent-induced backbone fluctuations and the collective librational dynamics of lysozyme studied by terahertz spectroscopy

K. N. Woods*
Physics Department, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA
(Received 24 November 2009; published 23 March 2010)

THz spectroscopy is used to investigate the dynamics of the globular protein hen egg white lysozyme under varying hydration and temperature conditions. An analysis of the experimental spectra has revealed that the amount of solvent in the hydration shell has a strong influence on the low-frequency protein conformational dynamics and also the arrangement of hydrogen bonds in the protein secondary structure. Furthermore at a hydration level >0.2 we identify collective backbone fluctuations in the protein secondary structure that are not present at low hydration. It is possible that these solvent induced modes are important for the biological function of the protein.

DOI: 10.1103/PhysRevE.81.031915 PACS number(s): 87.15.H−, 87.15.kr, 78.30.−j

I. INTRODUCTION

Proteins perform many of the biological functions in living cells. A long history of both experimental and theoretical investigations on protein dynamics has established that conformational changes are essential for protein function [1–5]. To this end, protein functionality is dependent on conformational rearrangement and flexibility. The internal dynamics in proteins have been found to fall into two main categories [6]. There are the fast motions that fluctuate within the minima of the energy landscape that defines the protein configuration space possibilities. And second, there are the slower large-scale thermally activated motions that represent the transitions between the valleys separating distinct protein conformations [7]. It has been suggested that fast thermal fluctuations are important for protein activity because they act as the "lubricant" that enables conformational changes in physiological time scales [8]. The results from both experimental and theoretical studies have suggested that picosecond time scale fluctuations taking place in the secondary structure in proteins have a strong influence in determining the reaction rates of those proteins [9–11]. While more recent numerical and computational studies examining fast protein fluctuations have illustrated that proteins are capable of localizing, storing, and transporting energy in the presence of spatial disorder on a subnanosecond time scale [12]. In light of these findings, the aim of the work presented in this article is to explore the picosecond time scale fluctuations taking place in the well-studied protein lysozyme, as a function of hydration and temperature. The expectation is that a detailed analysis of the THz detected vibrational modes in lysozyme may aid in the characterization and interpretation of the protein internal motions, occurring in the picosecond time scale regime, that have been conjectured to play a fundamental role in determining enzymatic activity [13,14].

II. MATERIALS AND METHODS

A. Sample preparation

Hen egg white lysozyme was purchased from Sigma-Aldrich (St. Louis, MO). To remove excess salt and other particles from the sample prior to the experiment, the lysozyme sample was dissolved in water and run through a desalting spin column with a buffer consisting of 10 mM NaH_2PO_4 and 0.01 mM ethylenediaminetetraacetic acid at pH 7.0. The concentration of the sample protein was determined by UV absorbance using a standard curve derived from a series of dilutions at 280 nm. The lysozyme solution sample used in the infrared experiments was prepared by diluting the stock protein sample to a concentration of 1 mg/ml. The unoriented film sample was prepared by placing 20 μl of the dilute protein solution sample (1 μg/μl) onto a high-resistivity silicon substrate for the THz spectroscopy experiment. The protein sample on the substrate was subsequently placed in a sealed sample chamber for approximately seven days while in the presence of a saturated salt solution to achieve the desired sample relative humidity (RH). From these film samples, the hydration content of the protein was determined by weighing. N-methylacetamide (NMA) was also purchased from Sigma and the experimental infrared sample was prepared in an analogous manner to the lysozyme sample.

B. THz spectroscopy experiments

The THz spectroscopy experiments were carried out on a Jasco FTIR—6000 series spectrometer. The lysozyme film samples were collected with a liquid helium cooled bolometer in the 15–250 cm^{-1} spectral range. The sample cell used in the experiments contained a 0.006 mm thickness polytetrafluoroethylene spacer (Specac Ltd., U.K.) and for each transmission measurement a 25-mm-diameter region of the protein film sample was illuminated with the THz beam to determine the absorbance.

To maintain the hydration level of the protein film during the experiment, the sample was placed in a sealed transmission cell consisting of two silicon windows. Reversibility of the temperature response of the protein sample, in terms of absorption features and intensity, was one criterion used to verify that the seal was maintained throughout the experiment. Comparison of the weight of the protein sample before and after the experimental measurement was also used to confirm that there was no significant water loss during the

*Corresponding author; knwoods@cmu.edu

duration of the experiment. In the spectral measurements presented each scan consists of 16 averaged scans and the infrared data was collected with a spectral resolution of 4 cm^{-1}. The 15–120 cm^{-1} THz spectra were collected with a 25 micron beam splitter while the data in the 100–250 cm^{-1} spectral region was collected with a 12 micron beam splitter. The temperature of the samples was varied using a SPECAC variable temperature cell. Using a combination of refrigerant and the control from the built-in temperature cell block heaters, the temperature of the sample could be adjusted from −190 °C to 30 °C.

C. Molecular dynamics simulation

1. Equilibrium molecular dynamics simulations

The molecular dynamics (MD) simulations were carried with the Gromacs package [15] version 3.2 using the Gromacs-96 force field. A starting structure of the hen egg white lysozyme configuration was initially downloaded from the protein databank. For the low hydration sample the protein was initially hydrated with solvent molecules in a 3 Å shell surrounding the protein surface, which corresponds to 515 water molecules and 8 Cl ions. After an initial 1 ns equilibration period, the number of water molecules was reduced to the closest 145 water molecules to the protein surface for the full simulation period. This is equivalent to a hydration level of approximately 0.18 g water per g protein (0.18 g/g). The sample with higher hydration content consists of 650 water molecules and 8 Cl ions and is equivalent to a hydration level of 0.81 g/g. In the simulations the SPC model of water was used. Energy minimization of the hydrated protein system was carried out by using a steepest descent method to a convergence tolerance of 0.001 kJ mol^{-1}. The energy minimization was followed by a MD run with constraints for 200 ps in which an isotropic force constant of 100 kJ mol^{-1} nm^{-1} was used on the protein atoms. During the restrained dynamics simulation the temperature and pressure of the system were kept constant by weak coupling to Berendsen temperature and pressure baths [16] and in all cases the protein, water, and ions have been coupled to the temperature and pressure baths separately. The final output configuration from the MD simulation with constraints was used as the starting configuration for the sample with high hydration content for the full 10 ns MD simulation. The final simulation was carried out with a 1 fs time step where the bonds between the hydrogen and the other heavier atoms were restrained to their equilibrium values with the linear constraints (LINCS) algorithm [17]. Particle mesh Ewald (PME) method [18,19] was used to calculate the electrostatic interactions in the simulation and was used with a real-space cutoff of 1.0 nm, a fourth order *B*-spline interpolation and a Fourier spacing of 0.12 nm. The molecular modeling software program VMD [20] was used to visualize the simulation final configurations.

We have used a rigid water model in our MD simulations of hydrated lysozyme. Previous computational studies have shown that simple three-site models of water (such as SPC) do not adequately reproduce some of the higher-frequency components of the experimental IR spectrum of liquid water [21,22]. To accurately reproduce these high frequency vibrational peaks in the spectrum, particularly the OH stretching vibration and the vibrational mode, the classical approximation of water used in the MD simulation is generally expanded with quantum corrections [23,24]. For instance, flexible water models constructed from a combination of rigid water models augmented by anharmonic terms for intramolecular interactions have successfully yielded the correct OH vibrational frequencies of liquid water [25,26]. Despite these shortcomings, recent experimental investigations have demonstrated that the empirical potentials of rigid water models utilized in MD simulations for observing water-protein interactions are capable of realistically capturing the fundamental IR transitions in the THz region ($\leq$200 cm^{-1}) of the spectrum [27,28]. The close correlation between experiment and simulation suggests that a rigid water model provides a reasonable depiction of the fast, cooperative hydrogen bonding dynamics of water molecules that form the protein hydration shell. It is unlikely that the use of a flexible water model in our MD simulations of lysozyme would dramatically alter the observed protein dynamics that have been used to aid in our interpretation of the experimental THz data on the protein.

The velocity autocorrelation function (VACF) of both atoms and molecules from the MD simulations were computed with the extended analysis tools that are included as part of the Gromacs software package. The VACF is defined by

$$\boldsymbol{C}_{\boldsymbol{v}}(\tau) = \frac{\langle \boldsymbol{v}_i(\tau) \cdot \boldsymbol{v}_i(0) \rangle}{\langle \boldsymbol{v}_i(0)^2 \rangle}, \tag{1}$$

where $\boldsymbol{v}$ refers to velocity and $\boldsymbol{i}$ denotes an atom or molecule in the simulation system. Fourier transform of the VACF is used to project out the underlying frequencies of the molecular processes associated with the correlated motions detected in the simulation. The dipole autocorrelation function from the MD simulations has also been calculated in an effort to interpret the infrared active dipole dynamics in the experimental investigation. The dipole autocorrelation function, given by

$$\boldsymbol{C}_{\mu}(\tau) = \frac{\langle \mu_i(\tau) \cdot \mu_i(0) \rangle}{\langle \mu_i(0)^2 \rangle}, \tag{2}$$

where $\boldsymbol{\mu}$ refers to the electron dipole moment, has been computed for the dipole moments of the protein. The corresponding THz dipole spectrum of the protein simulation trajectories were calculated by computing the Fourier Transform. Midinfrared spectra from the MD trajectories were computed using the VMD infrared density calculator plug-in.

Similarly, dynamical fluctuations within the hydrogen bond network within the protein or between the protein and the solvent molecules have been characterized by use of a correlation function [$C_{hb}(\tau)$], which averages over hydrogen bond pairs,

$$\boldsymbol{C}_{\boldsymbol{hb}}(\tau) = \frac{\langle \boldsymbol{hb}_i(\tau) \cdot \boldsymbol{hb}_i(0) \rangle}{\langle \boldsymbol{hb}_i(0)^2 \rangle}, \tag{3}$$

and has an existence function of either 0 or 1 ($hb(\tau) = [0,1]$) for a particular hydrogen bond i at time t. In this

analysis, a hydrogen bond is defined by using a geometrical criterion, where the center of mass distance is less than 3.5 Å, the (O···H) distance is smaller than 2.6 Å, and the ∠HO···O angle is smaller than 30°. Other weak interactions (Van der waal, electrostatic, etc…) in the protein system are identified as contacts within an appropriate cut-off distance.

Principal component analyses (PCAs) were carried out by diagonalzing the covariance matrix $C_{ij}=\langle(x_i-\langle x_i\rangle)(x_j-\langle x_j\rangle)\rangle$, where x denotes protein atomic positions in the $3N$-dimensional configurational space and the angular brackets represent the averages over the MD trajectory. Translational and rotational motions were removed by a least-squares fitting to a reference structure. The eigenvectors of $\boldsymbol{C}$ were determined by diagonalization with an orthonormal transformation matrix. The resulting eigenvectors from the transformation were used to determine the PCA modes with eigenvalues (λ) equivalent to the variance in the direction of the corresponding eigenvector. The MD trajectory was projected onto the principal modes to determine the principal components. The eigenvalues λ_i of the principal components denote the mean square fluctuation of the principal component i and are arranged so that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{3N}$. Using this arrangement, the trajectories were filtered along the first principal component to analyze the collective dynamics taking place within the protein. For the analysis, the first 2 ns of the simulation have been used for equilibration and the remaining 8 ns of the simulation of have been divided into two parts (each 4 ns in duration) for comparison. It should be noted that the simulations used for the correlation analyses in this work are rather short when compared with the MD trajectories that are generally used (hundreds of nanoseconds) to elucidate the large scale correlated motions in proteins. In this sense, it is also important for us to outline our experimental aim for this investigation. Our goal in this investigation is to provide further insight into the role that fast, picosecond time scale fluctuations in lysozyme may have in influencing the longer time structural changes in the protein. It has been suggested that the internal dynamics of proteins can be modeled as a mixture of fast and slow modes. In this model, the faster regime has a characteristic relaxation time of a few picoseconds and corresponds to diffusion within a single well of the high-dimensional configuration landscape of the protein. The longer diffusion regime, on the other hand, is characteristic of motion between wells. Consequently, it is unclear if the protein is in more than one configuration during the time duration of the MD simulations utilized for this investigation. But based on the cosine content of PC's analyzed, there is a convergence of sampling around at least one conformation and in this view is useful for our aims for this investigation.

Full correlation analyses (FCA) on the MD trajectories were carried out by using the algorithm developed by Lange and Grubmüller [29]. In short, FCA relies on mutual information (MI) to identify correlations of residues from the MD trajectory. Fluctuations in the coordinates of the protein are considered to be uncorrelated if

$$\boldsymbol{p}(\boldsymbol{x}) = \prod_{i=1}^{3N} \boldsymbol{p}_i(\boldsymbol{x}_i), \tag{4}$$

where $\boldsymbol{p}(\boldsymbol{x})$ signifies the canonical ensemble given by $\boldsymbol{p}(\boldsymbol{x}) = \boldsymbol{e}^{-E_i\beta}/\boldsymbol{Z}$, in which Z refers to the partition function, E the energy, β the inverse temperature, and $\boldsymbol{p}_i(\boldsymbol{x}_i)=\int \boldsymbol{p}(\boldsymbol{x})\boldsymbol{dx}_{i\neq j}$ the marginal density. If there are correlations in the system then they can be identified by MI,

$$\boldsymbol{I}(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3,\ldots\ldots\boldsymbol{x}_{3N}) = \sum_{i=1}^{3N} \boldsymbol{H}(\boldsymbol{x}_i) - \boldsymbol{H}(\boldsymbol{x}), \tag{5}$$

where $\boldsymbol{H}(\boldsymbol{x})=-\int \boldsymbol{p}(\boldsymbol{x})\log \boldsymbol{p}(\boldsymbol{x})\boldsymbol{dx}$ designates the information entropy. MI measures any correlation in the system, both linear and higher order. FCA minimizes the mutual information between the collective coordinates in the protein system, of which the coordinates from the PCA are used as an initial guess. The generalized correlation coefficient, defined by

$$\boldsymbol{r}_{\boldsymbol{MI}}(\boldsymbol{x}_i,\boldsymbol{x}_j) = \left(1-\boldsymbol{e}^{-2\boldsymbol{I}(\boldsymbol{x}_i,\boldsymbol{x}_j)/\boldsymbol{d}}\right)^{1/2}, \tag{6}$$

is used to visualize and interpret the correlations detected with FCA.

2. Nonequlibrium molecular dynamics simulations

Nonequilibrium molecular dynamics simulations on both the low and high hydration protein systems were carried out with the simulated annealing algorithm as implemented in Gromacs. The initial configuration and velocities for the nonequilibrium simulations were taken from the final configuration of the 10 ns equilibrium MD simulations. The reference temperature of both the protein and the solvent molecules in the protein system were dynamically changed after the initial 2 ps of the simulation from 310 to 298 K. The initial temperature drop was followed by a 1.8 ns simulation at 298 K. The temperature change employed in the nonequilibrium MD simulation is equivalent to an energy of approximately 300 kJ/mol.

III. RESULTS AND DISCUSSION

A. THz spectrum of hen egg white lysozyme

In the experimental measurements carried out in this work, hen egg white lysozyme (HEWL) has been measured at three different hydration levels. The samples have been prepared as unoriented, solid films maintained at different relative humidity. As a film, the protein samples have been measured at low hydration with a RH of 75%, which is equivalent to a hydration level (h) of 0.18 g/g; at an intermediate hydration level at 85% RH (h=0.32 g/g); and also at full hydration at 97% RH (h=0.81 g/g) [30]. In the discussion of the experimental data, we will refer to the hydration levels of the protein sample as low hydration (h=0.18 g/g), intermediate hydration (h=0.32 g/g), and high hydration (h=0.81 g/g). Previous experimental measurements and computational investigations have determined that there is a correlation between hydration level and protein enzymatic activity [31–34]. At low hydration the enzyme in not catalytically active; while at the intermediate hydration level the enzyme has been found to be functional but has very low activity [35]. The sample at the highest hydration level studied in this experimental investigation has been shown to possess all of the functionality of the active protein [36]. Although, the smaller number of free water molecules

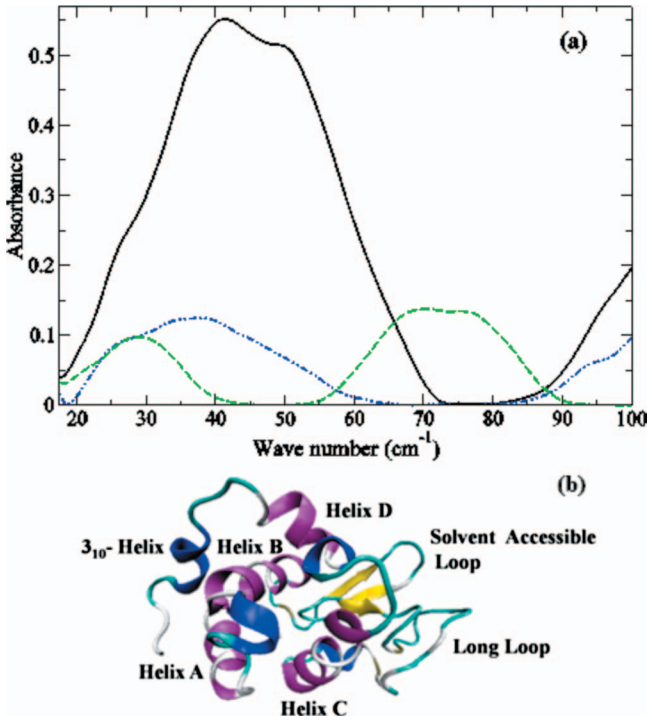

FIG. 1. (Color) (a) Experimental THz spectrum of room temperature lysozyme in the 20–100 cm^{-1} spectral region at 97% RH (black solid line), 85% RH (blue dotted line), and 75% RH (green dashed line). (b) Secondary structure of HEWL with regions labeled as they are referred to in the text of the manuscript. The α—helices in the protein are purple; 3_{10} helices are blue; turns and loops are cyan; and β—sheets are yellow.

in the hydration shell when compared to a solution sample is believed to affect the rate of the motions taking place in the protein.

In our experimental measurements we find that the THz spectrum of lysozyme displays a markedly different spectrum at low hydration when compared with that at high hydration [Fig. 1(a)]. The low hydration room-temperature sample of lysozyme exhibits prominent peaks at approximately 28, 68, and 75 cm^{-1}. As the hydration level is increased from low to intermediate hydration, the peaks at 68 and 75 cm^{-1} disappear and there is a peak that emerges at approximately 36 cm^{-1}. At the highest hydration level studied, the room temperature sample of lysozyme possesses a broad peak centered at around 45 cm^{-1}. One topic that we would like to try to address in the discussion regarding the experimental results detected in the THz spectra of lysozyme is the relationship between the collective vibrational modes detected in the protein and their possible relationship with enzymatic activity.

B. Equilibrium moleculard dynamics simulation results

The lowest-frequency large amplitude mode at approximately 30 cm^{-1} in the lysozyme sample at both low and intermediate hydration in Fig. 1 appear at a similar frequency as the peak observed in previous inelastic neutron scattering (INS) experiments on lysozyme [37,38]. This low-frequency band has also been observed in subsequent Raman [39] and OHD-RIKES experiments [40] on lysozyme. In these earlier experimental studies, the presence of the ~30 cm^{-1} peak in the low-frequency spectrum of the protein has been ascribed to collective librational dynamics of polar exposed side chains. In an effort to understand the origin of the peak at ~30 cm^{-1}, as well as the other infrared bands detected in our experiments, we have carried out MD simulations on lysozyme at hydration levels equivalent to the lowest and highest hydrated samples measured in our experimental investigation. An analysis of the Fourier transform of the VACF of side chain residues in the MD simulation on lysozyme at both hydration levels has uncovered bands at approximately 18, 33, 66, 110 (shoulder), and 135 cm^{-1} in the <200 cm^{-1} spectral region. Hence, the identification of the ~30 cm^{-1} peak as a side chain fluctuation from earlier investigations is consistent with the analysis from the MD simulations carried out in this work. A peak centered at 75 cm^{-1} in the low hydration experimental spectrum in Fig. 1 has only been detected in low-frequency Raman experiments [39,41] as well as other simulation [42] and vibrational spectroscopy measurements [40] on globular proteins, including lysozyme. Its absence from INS spectra suggests that it is not highly influenced by hydrogen bonding dynamics. As a consequence, it has been hypothesized that this mode arises from torsional motions occurring in the protein backbone.

The aim of our investigation on the low-frequency dynamics taking place in lysozyme is to gain further insight about the role that the observed differences in the experimentally detected THz fluctuations of the protein, at different hydration levels, may play in determining the protein's function. To this end, we have conducted a PCA on the lysozyme trajectories to aid in the identification of the functionally relevant correlated motions that take place in the protein at the two different hydration levels. In these analyses, the PCA modes with the largest fluctuation amplitude are selected to both identify and quantify the correlated motions taking place in the lysozyme secondary structure. The expectation is that identification of both the positive (*in phase*) and negative (*out of phase*) atomic correlations from the simulation may provide valuable insight into the fast fluctuations taking place in the protein that are important for establishing enzymatic activity on a much longer time scale. In the low hydration simulation of lysozyme, PCA of the largest amplitude mode reveals that the *in phase* correlated motion with the largest magnitude takes place between valine 2 (or VAL 2, where the numeral 2 denotes that it is the second amino acide in the protein sequence) and phenylalanine 3 (PHE 3) with phenylalanine 38 (PHE 38). A structural analysis of the protein trajectory at low hydration has uncovered that PHE 38 is in direct contact with both VAL 2 and PHE 3. The largest magnitude *out of phase* correlation in the low hydration protein occurs between arginine 5 (ARG 5) and glycine 117 (GLY 117). ARG 5 is the first residue of Helix A, one of the well defined α helices in the α lobe of the lysozyme secondary structure [Fig. 1(b)]. GLY 117, also in the α lobe of the protein, resides in a solvent exposed turn region. An NMR solution structure of hen egg white lysozyme has established that GLY 117 has the highest main chain solvent accessibility

in the protein [43]. As a consequence, it is likely that interaction with the solvent plays an important role in modulating the *out of phase* correlation between ARG 5 and GLY 117. Perhaps for this reason it is not entirely surprising that the addition of a greater number of water molecules into the protein hydration shell has a strong influence on the types and number of correlated motions taking place in the protein. The *in phase* correlation of residues 2 and 3 with that of residue 38 is still present in the high hydration simulation of lysozyme. But the magnitude of this correlation is greatly reduced when compared with the low hydration system. Analysis of the largest amplitude PCA mode in the high hydration simulation of lysozyme reveals a sharp increase in the number of large amplitude *out of phase* correlations. We find evidence of strong correlations of residues in Helix A with residues in the 3_{10} helix in the same (α–) lobe. Large magnitude *out of phase* correlations are also evident in residues in the solvent accessible loop region (residues 101–104) of the protein with residues in the same (α lobe) 3_{10} helix. Specifically, the *out of phase* residue correlations occur between ALA 10 and ALA 11 in Helix A with ALA 122 in the 3_{10} helix. In the solvent accessible loop region GLY 102 and ASN 103 are correlated with TRP 123 in the 3_{10} helix.

In our analyses of the concerted motions taking place in the lysozyme secondary structure, we have also carried out a full correlation analysis (FCA) to assess the nonlinear contributions to the correlations taking place within the secondary structure of the protein. Our motivation is to identify anharmonic fluctuations that may be constructive in formulating an appreciation about the relationship between energy transport pathways and conformational fluctuations taking place within the protein secondary structure. FCA optimizes MI rather than fluctuation amplitude from the trajectories obtained by MD simulation. In contrast to the Pearson correlation coefficients used in the PCA analysis, the canonical correlations used in the FCA do not differentiate between correlated and anticorrelated motions. However, FCA is capable of quantifying both linear and higher-order correlations from the simulation trajectories. In our analysis we have selected FCA modes with the highest anharmonicity and collectivity. Modes that fit these criteria are characteristic of the motions that are most likely to be detected in our experimental spectrum [44]. They are also the motions that have a greater likelihood of revealing connections in the protein structure that are essential for biological activity [45]. The FCA modes with the highest anharmonicity in the low hydration simulation of lysozyme reveal highly correlated motion within the turn and loop regions of the molecule [Fig. 2(a)]. Specifically there are large magnitude correlations among the residues in the long loop region in the β lobe of the protein (residues 61–78). Residues in the long loop are also strongly linked with SER 86, a residue that is located in a turn region in the α lobe of the protein. Additionally, the residues in the long loop of the protein are correlated with residues 100–102 in the solvent accessible loop region and the dynamical motion of GLY 102 is also linked with residues 122–125 located in the 3_{10} helix of the α lobe. These same residues in the 3_{10} helix (122–125) are also correlated with SER 86. The strongest *nonlinear* contributions of the correlations from the FCA modes analyzed arise specifically

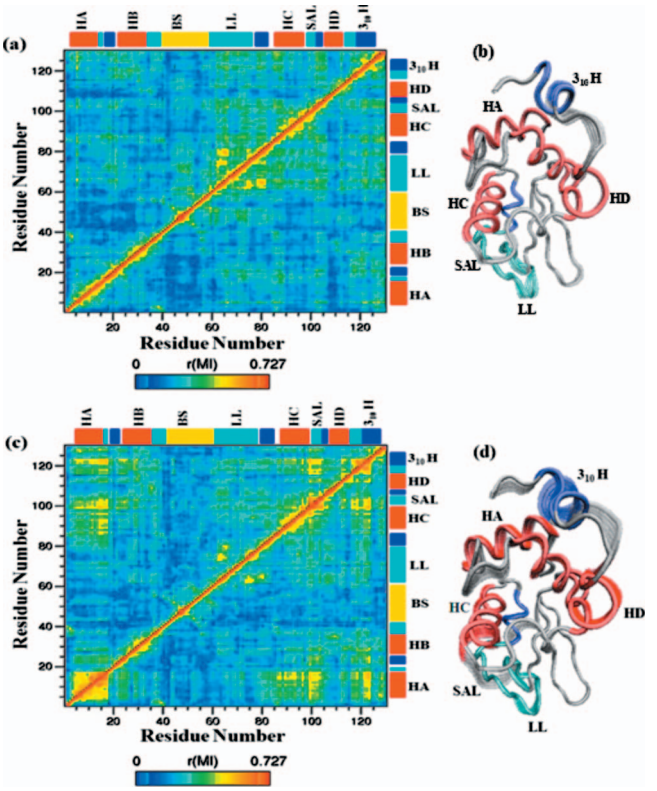

FIG. 2. (Color) Generalized correlation coefficient (r_{MI}) for lysozyme C_α atoms at (a) low hydration and (c) high hydration. The strength of the correlation between residues is represented by a color mapping scheme. Secondary structural regions of the protein (b) and (d) are superimposed over 10 frames of the respective trajectories in an effort to indicate the amplitude of the motion.

from GLY 102 with SER 86 and SER 86 with ALA 122 in the low hydration simulation of lysozyme. Both the nonlinear and linear correlations taking place in the molecule culminate into a motion that can be described as an out of plane twisting motion mostly affecting the lower (β–) lobe.

In the high hydration simulation of lysozyme, the FCA of the modes with the appropriate characteristics reveals strong intrahelical correlations between the residues in helix A [Fig. 2(b)]. The residues in helix A are also correlated with residues in the turn regions in the α lobe along with residues that are part of the solvent accessible loop region (residues 86–102) and with residues that form helix C. Motion of the residues in helix A are also linked dynamically with the residues in the turn region preceding the 3_{10} helix as well as with residues within the 3_{10} helix itself. Additionally, from these analyses we have detected prominent correlations of residues 102 and 103 in the solvent accessible loop region with residues in the 3_{10} helix (118–122). *Nonlinear* associations in the highly hydrated protein system take place between GLY 102 with ALA 122. There are also large magnitude *nonlinear* contributions arising from interactions of VAL 120 with the residues in helix A. In addition, VAL 120 has nonlinear associations with both GLY 102 and ALA 122. The dynamics of the highest anharmonic mode in the high hydration protein is dominated by fluctuations within the helical regions of the protein. Particularly, the 3_{10} helix in the α lobe and helices A

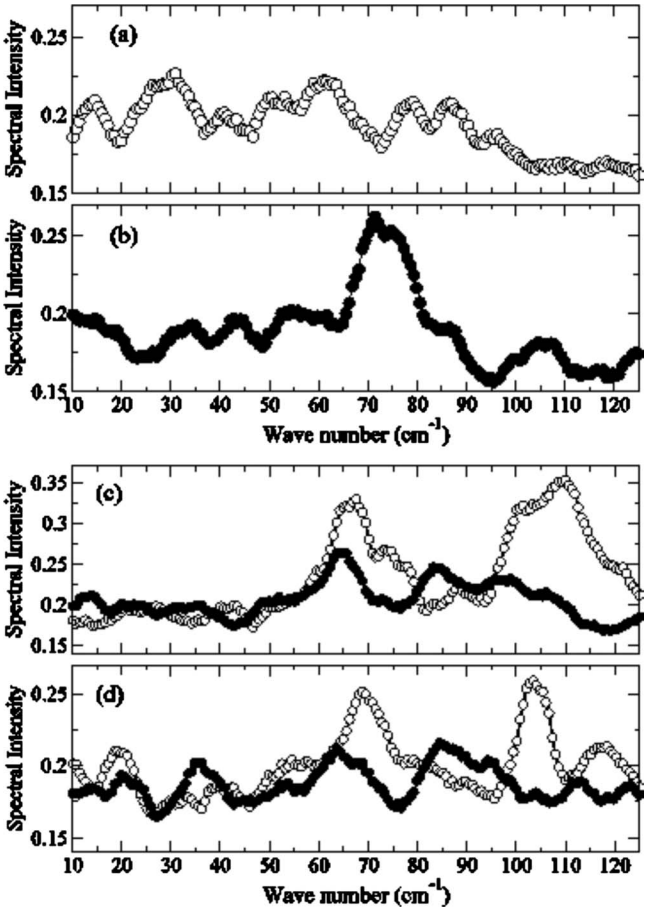

FIG. 3. VACF spectrum of protein carbonyl O···N atom correlations from (a) the high hydration and (b) the low hydration MD simulation of lysozyme. VACF spectrum highlighting (c) $C_\beta\cdots O$ atom interactions between VAL 2 and PHE 38 and (d) the contact spectrum resulting from interaction of VAL 2 methyl groups with the PHE 38 O atom from the low hydration (filled circles) and the high hydration (open circles) MD simulation of lysozyme.

and C have the largest amplitude motion. Interestingly, we have detected that there is also a noticeable reduction in the amplitude of the oscillations taking place within the turn and loop regions of the protein secondary structure when compared with the drier protein.

C. Nonlinear associations and collective vibrational modes detected in the THz spectra of lysozyme

The analysis of the motions contributing to the nonlinear associations from the simulation may provide some insight into how we may interpret the experimental measurements carried out in this work. From the analyses of the simulation trajectories, it has been determined that the drier protein has a larger number of intramolecular associations, particularly in the turn and loop regions. The Fourier transform of the VACF of GLY 102, a residue associated with nonlinear dynamics in the low hydration protein, reveals a prominent peak at approximately 75 cm^{-1} in the $<$200 cm^{-1} spectral region [Figs. 3(a) and 3(b)]. Further analysis has revealed that this peak is largely attributed to carbonyl O···backbone *N* correlations arising from protein intramolecular associations between residues within the turn and loop regions in the protein secondary structure. On a similar note, we have also observed that the number of protein intramolecular hydrogen bonds is for the most part, greater in the low hydration system when compared with the higher hydration simulation of the protein. In addition to the carbonyl O···backbone *N* correlations, we have also determined that the low-frequency oscillations in the low hydration system are also strongly affected by the direct contact of the residues that actively form the β bride in the protein structure. Particularly, an analysis of the simulation trajectory has revealed that the VAL 2 Cβ atom is in contact with the PHE 38 carbonyl oxygen (O). In response to the dynamical connection between VAL 2 and PHE 38, there is a bending mode at about 63 cm-1 in the VACF spectrum highlighting the main chain C$\beta\cdots$carbonyl O association [Fig. 3(c)] that develops in response to the protein intramolecular interactions in the β bridge. Bands at this frequency have also been detected from other C$\beta\cdots$O atom associations in the low hydration system; although, the magnitude of the peak from these other contributions are less prominent in the vibrational spectrum. In our analysis we have uncovered that the bending motion that results from the main chain C$\beta\cdots$carbonyl O interactions taking place between the turn and loop regions of the protein has a direct affect on the other main chain and side chain carbon fast atomic fluctuations, particularly the methyl side chain groups. As an example, a plot of the methyl group motion of VAL 2 from the VACF spectrum in response to the contact between PHE 38 carbonyl O with VAL 2 Cβ [Fig. 3(d)] displays a small peak in the spectrum located at 63 cm^{-1} (attributed to the C$\beta\cdots$O interaction). In addition, there are two other noticeable bands in the spectrum located at approximately 33 and 88 cm^{-1}. The band at 33 cm^{-1} is associated with methyl group torsional fluctuations that arise as a result of the O—Cβ interactions, while the band at 88 cm^{-1} is due to a Cα—carbonyl O atom deformation in the protein backbone that also forms in response to the protein intramolecular interaction. The mode at 88 cm^{-1} has been detected inVAL 2 as well as other residues involved in turn and loop region protein intramolecular correlations. It is interesting to note that the VAL 2—PHE 38 interaction spectra in the highly hydrated protein simulation of lysozyme [Figs. 3(c) and 3(d)], has a large amplitude band at a slightly higher frequency (65 cm^{-1}) than the band observed in the low hydration spectrum. The shift in frequency and increase in intensity reflects the stronger interaction of the protein with the solvent. Although we have found no evidence of direct contact between the VAL 2 nonpolar side chains with the solvent, we hypothesize that the stronger interaction of the solvent with other polar side chains in the protein secondary structure influence the fluctuations of residues remote from the protein surface where the solvent interaction with the protein is maximized. Experimentally, we have uncovered a band in the experimental THz spectrum of pure room temperature liquid water at approximately 70 cm^{-1} that we have attributed to a water bending mode of water molecules not arranged in a tetrahedral structural configuration [46]. An analysis of the VACF of the oxygen atom of the water molecules from the simulations of lysozyme carried out in this

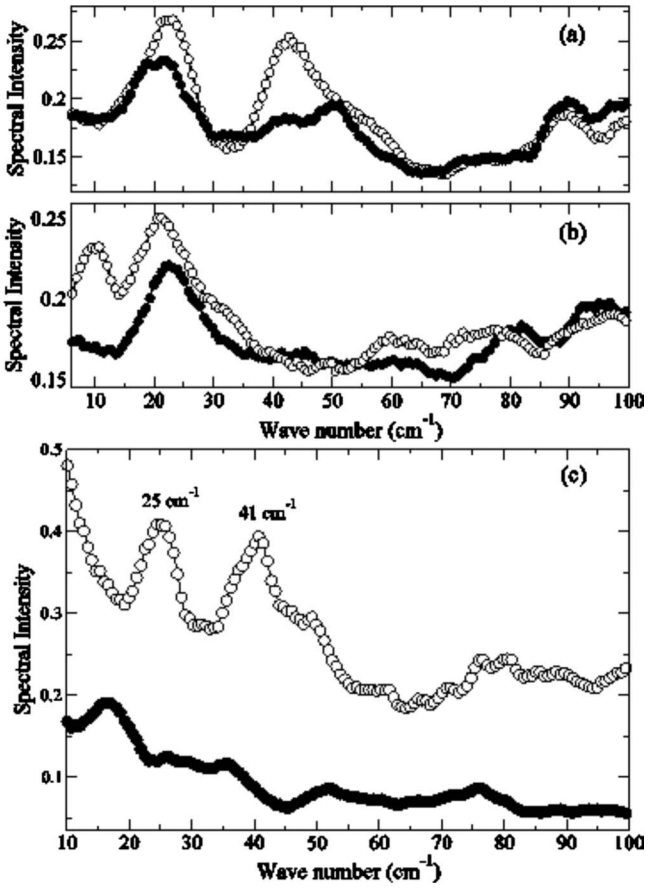

FIG. 4. VACF spectrum of (a) ALA 11 and (b) ALA 122 from the low hydration (filled circles) and the high hydration (open circles) MD simulation of lysozyme. (c) Backbone dipole correlation spectrum from the MD simulation of residues in helix A at low hydration (filled circles) and at high hydration (open circles).

investigation have uncovered a band at 65 cm^{-1} that does not strongly involve hydrogen atoms. Therefore, we accredit this band to an intermolecular bending mode taking place within the hydrogen bonded network of the solvating water molecules. The band at 33 cm^{-1}, prominent in the spectrum accentuating the localized methyl group torsional motion in the low hydration simulation [Fig. 3(d)], is not present in the high hydration simulation of lysozyme. In its place, the methyl group torsional oscillation of VAL 2, along with the methyl side chain groups of other nonpolar residues in the protein structure, generate a peak at $\sim$110 cm^{-1}. The peak at 110 cm^{-1} reflects the shift toward a more active solvent side chain interaction that forms in response to an increased number of water molecules in the protein hydration shell.

The VACF spectrum of ALA 122 [Fig. 4(b)] in the low hydration protein simulation has a peak at approximately 22 cm^{-1} that is attributed to the carbonyl C main chain Cβ atom motion of correlated residues that are strongly modulated by both the solvent and methyl group fluctuations. In the highly hydrated simulation of lysozyme, the 22 cm^{-1} mode has a stronger intensity reflecting the large magnitude out of phase correlation that takes place between ALA 11 (helix A) and ALA 122 (3_{10} helix). Interestingly, the analysis of the VACF of SER 86 did not uncover any prominent bands from the simulation trajectory in the accessible THz region.

Using the results from the MD simulation, it is now possible to re-evaluate the experimental THz spectrum of the low hydration sample of lysozyme with greater insight. In Fig. 5(c), the band at 75 cm^{-1} is prominent at 93 K and decreases in amplitude as the temperature of the protein sample is increased. The temperature behavior of the 75 cm^{-1} band is consistent with a backbone mode [47]. Hence, it is possible that the 75 cm^{-1} is related to the correlated backbone motion identified at a similar frequency in the simulation spectrum in the low hydration protein. Furthermore, the band at 75 cm^{-1} is close in frequency to the backbone mode detected in previous Raman experimental investigations [40,41]. Figure 5(c) reveals that the low hydration sample of lysozyme also includes a peak at 68 cm^{-1} that is only observed at temperatures $\geq$250 K. Based on our analysis of the equivalent low hydration simulation system, we hypothesize that this band reflects a protein intramolecular bending mode. It is likely that the temperature dependence of this band stems from the connection that exists between protein flexibility and temperature. The protein transition temperature (T_g) for a number of different proteins has been found to lie in the range of 200–230 K [48–50]. For this reason, it is probable that the 68 cm^{-1} is only visible in the spectrum only after the transition temperature has been reached and the protein flexibility is increased. The peak at 28 cm^{-1} in the experimental spectrum [Figs. 5(c)–5(f)] initially increases in amplitude as the temperature is raised, but begins to decrease in amplitude above 250 K. Based on our earlier analysis of the correlated motions taking place within the secondary structure in the low hydration simulation of lysozyme, we deduce that the infrared peak at 28 cm^{-1} arises as a result of backbone main chain correlations between residues in the turn and loop regions of the protein secondary structure. Curiously, we have also identified peaks from the simulation, that arise from protein dynamical motions ascribed to nonlinear associations, that we do not detect experimentally. Namely, we find prominent peaks at 85 and 88 cm^{-1} in the simulation spectrum that *are not* visible in the experimental spectrum of lysozyme. These peaks have been linked with backbone fluctuations involving deformation of $C_\alpha \cdots O$ and $C_\beta \cdots O$ atoms that form as a result of intramolecular correlations in the protein secondary structure. In an effort to develop a better understanding about the nature and characteristics of the backbone motions taking place in lysozyme, we have also carried out THz spectroscopy experiments on NMA, a model backbone peptide. In Fig. 6 the THz spectrum of NMA in the $\leq$100 cm^{-1} spectral region uncovers peaks at 20, 57, 63, 73, 82, and 88 cm^{-1}. The band at 82 cm^{-1} encounters some temperature dependence, as reflected in the experimental spectrum, while the rest of the peaks in the spectrum appear to be somewhat temperature independent. It is interesting to note that, excluding the lowest-frequency band at 20 cm^{-1}, there is also no change in the frequency of the bands when the solvent is exchanged from H_2O to D_2O implying that the modes in the spectrum have justly been labeled as backbone modes. The temperature insensitivity of the modes at 63, 73, and 88 cm^{-1} is consistent with backbone modes that we have

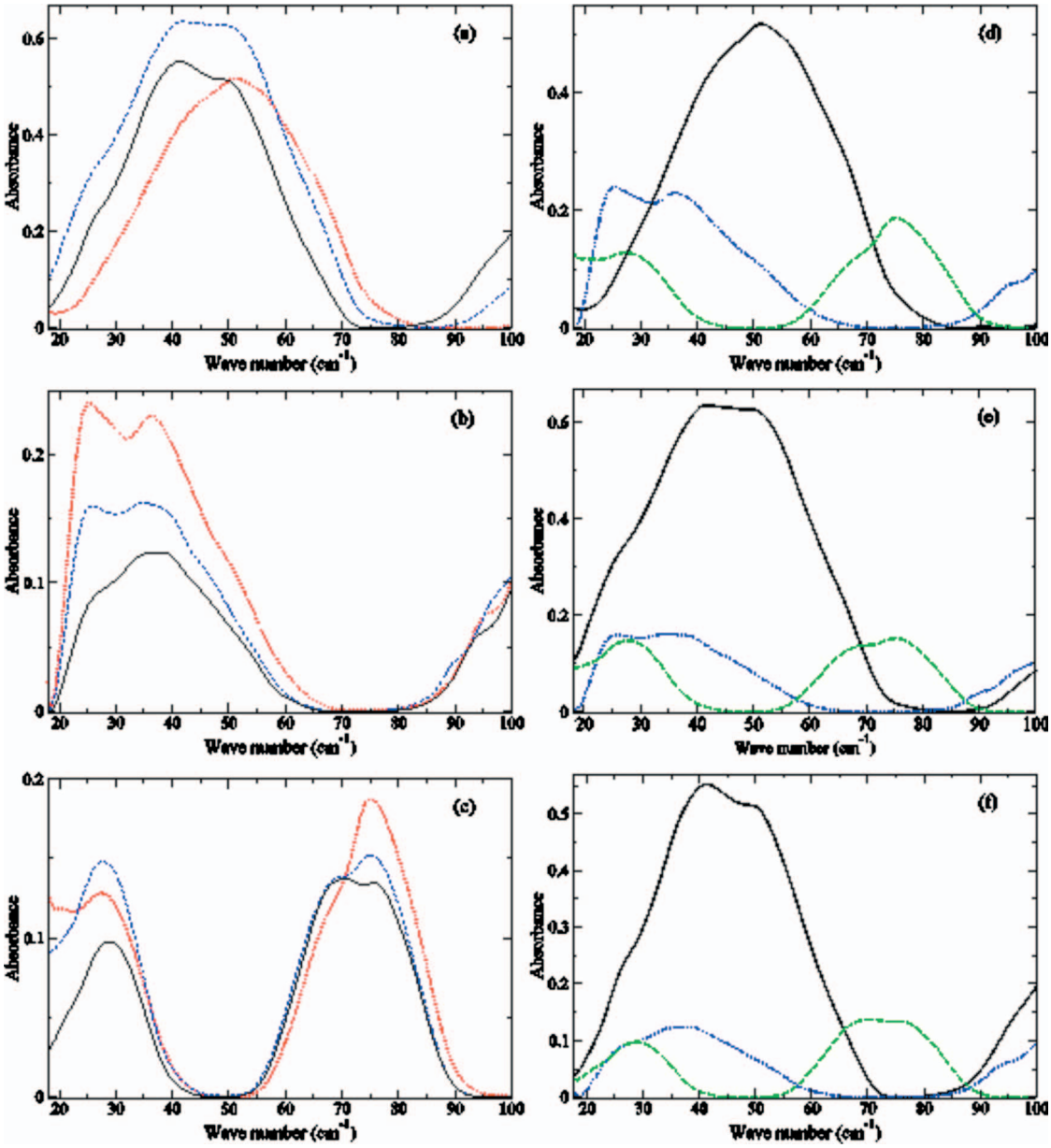

FIG. 5. (Color) Experimental THz spectrum of lysozyme in the 20–100 cm^{-1} spectral region at (a) 97% RH, (b) 85% RH, and (c) 75% RH and at temperatures 93 K (red dotted line), 250 K (blue dashed line), and at 303 K (black solid line). Overlay of the experimental THz spectra of lysozyme in the 20–100 cm^{-1} spectral region at 97% RH (black solid line), 85% RH (blue dot—dot-dashed line), and at 75% RH (green dashed line) at (d) 93, (e) 250, and (f) 303 K.

uncovered in our simulation analyses of lysozyme. The temperature dependence of the 82 cm^{-1} mode may be explained by its possible connection to an intramolecular protein mode uncovered at a similar frequency (85 cm^{-1}) in the simulation of the low hydration protein. From our analysis of the low hydration simulation trajectory of lyoszyme, we have determined that the 85 cm^{-1} mode is attributed to a protein main chain fluctuation strongly coupled to protein intramolecular hydrogen bonds. Further evidence to support this hypothesis has been uncovered in the analysis of the protein main chain dipole correlations taking place in the low hydration simulation sample. Peaks in the dipole autocorrelation analysis (not shown) appear at 12, 32, 62, and 85 cm^{-1} in the spectrum with a small shoulder located at 75 cm^{-1}. As an aside, it is interesting to note that the backbone mode at 32 cm^{-1} overlaps with the motion associated with side chain fluctuations detected in the velocity correlation analysis (Sec. III B). The band at 32 cm^{-1} is not present in the NMA spectrum suggesting that it arises from a side chain backbone correlation that is not present in the small model peptide.

Similar to what has been observed in the spectrum of ALA 122, the VACF spectrum of ALA 11 in the higher hydration simulation of lysozyme in Fig. 4(a) has a more intense band at 22 cm^{-1} when compared with the lower hydration spectrum. The amplitude difference in the residue correlation spectrum in the two protein systems reflects the stronger (*out of phase)* correlation between residues in helix A with residues in the 3_{10} helix. The *out of phase* correlations

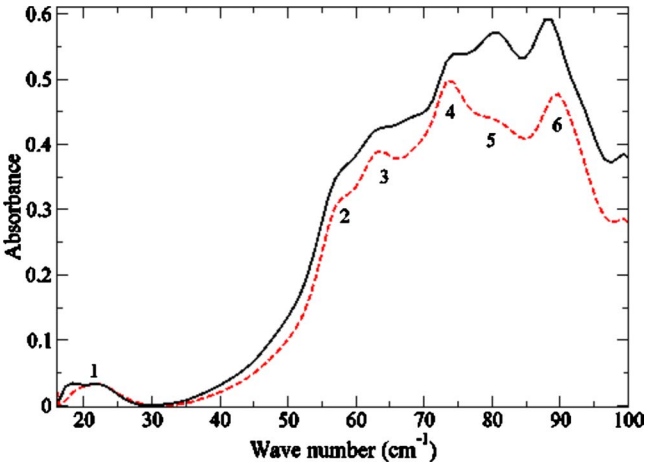

FIG. 6. (Color online) Experimental THz spectrum of NMA at 97% RH in the 20–100 cm^{-1} spectral region at 93 K (red dashed line) and at 303 K (black solid line). Prominent absorption bands are found at approximately (1) 20, (2) 57, (3) 63, (4) 73, (5) 82, and (6) 88 cm^{-1}.

taking place in the protein grow in number when additional water molecules are added to the protein hydration shell. Consequently, the analysis has also uncovered that in the high hydration sample of lysozyme the number of solvent—protein side chain hydrogen bonds are significantly enhanced when contrasted with the low hydration system. From the analysis of the protein dynamical correlations from the simulation, we have also established that the high hydration protein possesses fast backbone fluctuations that are not present at low hydration. For instance, we have identified a backbone fluctuation in helix A that is initiated as a result of the intrahelical correlations taking place *within* the α helix when the hydration level of the protein is increased. In Fig. 4(c), a plot of the backbone dipole correlations that arise in response to the intrahelical interactions taking place in helix A confirm the presence of bands at 25 and ~40 cm^{-1} in the spectrum that are not prominent in the low hydration simulation of lysozyme. We believe the band at 40 cm^{-1} stems from a backbone torsional fluctuation involving the backbone *N* atoms in helix A that are triggered as a result of the large magnitude out of phase correlation that take place between residues in helix A with residues in the 3_{10} helix as the hydration level in the protein is increased. An out of plane backbone deformation has also been identified in helix A at approximately 50 cm^{-1}. The solvent induced backbone fluctuation involves torsion of the C_α, and carbonyl C atoms with a small contribution from the *N* backbone atoms. Using the knowledge gained from analyses of the dynamical motions taking place in the protein from the simulations, it is now possible to reconsider the origin of the modes uncovered in our experimental THz spectra of lysozyme at intermediate and high hydration. In the intermediate hydration experimental spectrum [Figs. 5(b) and 5(d)–5(f)] there are peaks at 25 and ~40 cm^{-1}, that we surmise reflect backbone correlations taking place between residues in helix A with those in the (α lobe) 3_{10} helix. We believe that the peak close to 40 cm^{-1} is associated with the backbone fluctuation in helix A that has been observed in the simulation spectrum of lysozyme. At intermediate hydration, both the 25 cm^{-1} and ~40 cm^{-1} mode are present even at the lowest temperature studied [Figs. 5(b) and 5(d)] implying that they are both associated with backbone fluctuations taking place in the protein secondary structure. As the hydration level is increased from intermediate to high hydration, there is a peak that appears in the spectra at approximately 50 cm^{-1} [Fig. 5(a)]. In the experimental high hydration sample of lysozyme [Figs. 5(a) and 5(d)–5(f)], the peaks at 25 and 41 cm^{-1} are only prominent at temperatures $\geq$250 K. The temperature dependence of these modes in the high hydration protein sample might reflect the proteins greater interaction with the solvent and possibly the larger contribution of side chain motion to the modes. All of the <100 cm^{-1} modes that we have identified in our investigation of lysozyme appear to originate from backbone motions that are modulated by side chain fluctuations. This becomes even more apparent in the experimental spectrum when the solvent (H_2O) is replaced by D_2O. With the exception of the lowest frequency mode, there is a general blue shift of all of the detected modes in lysozyme with the solvent isotope exchange. We do not observe this same shift in the <100 cm^{-1} modes uncovered in NMA, presumably because in the small peptide the backbone motion is not as closely intertwined with the low frequency side chain fluctuations. The 50 cm^{-1} mode, on other hand, is visible at all of the temperatures investigated [Figs. 5(a) and 5(d)–5(f)] in the experimental high hydration lysozyme sample. Its absence from the lower hydrated samples suggests that its presence in the experimental spectrum is related to a solvent influenced conformational change that occurs within the protein structure. Our interpretation is consistent with the results of a recent NMR investigation examining the structure of lysozyme as a function of hydration [51]. The NMR results have revealed that the lysozyme secondary structure has a more open conformation when the hydration level in the protein is increased. It is likely that the 50 cm^{-1} mode detected in our experimental spectrum reflects the solvent induced backbone deformation uncovered in our analysis on the equivalently hydrated simulation sample of lysozyme at a similar frequency (~50 cm^{-1}).

It is important to point out that the peak observed at 25 cm^{-1} in both the intermediate and high experimental spectrum [Figs. 5(a) and 5(b)] likely has a different origin than the 28 cm^{-1} mode uncovered in the low hydration protein sample [Fig. 5(c)]. The fast fluctuations taking place in the driest sample studied have been found to be dominated by protein intramolecular associations. In the low hydration experimental spectrum [Figs. 5(c)–5(f)] we believe that the ~30 cm^{-1} mode is associated with methyl group oscillations that are amplified as a result of protein backbone—main chain (CN$\cdots$O) correlations. Previously in this section, we uncovered a mode at 63 cm^{-1} in the simulation of the low hydration protein, reflecting the protein intramolecular interaction [Fig. 3(c)]; which coincidentally, also had a strong influence on the methyl group fluctuation at about 30 cm^{-1} [Fig. 3(d)]. We have also observed that as the hydration level in the system is increased there is a noticeable shift in the frequency of this backbone mode from 63 to 65 cm^{-1}. We conjecture that this shift reflects the stronger influence of the solvent dynamics on the protein backbone motion. Accord-

TABLE I. Vibrational modes[a] (in cm^{-1}) detected in the nonequilibrium MD simulation of lysozyme within 800 ps after the initial temperature change.

≤100 ps	100–200 ps	200–400 ps	400–600 ps	600–800 ps
1684	1659	1493	1493(*l*)	1327(*l*)
1010	1327	1161(*h*)	1161(*h*)	1161(*l*)
336	1010	1010(*l*)	1010(*l*)	662(*l*)
130	336	167	995(*l*)	167
110	40(*h*), 50(*h*)		167	
	63(*l*), 75(*l*)			
60(*HW*)	65(*HW*),			
	116(*HW*)			

[a]The symbols *h*, *l*, and *HW* represent high hydration protein, low hydration protein, and solvent hydrogen atoms, respectively, and refer to the vibrational modes with appreciable amplitude only under the described conditions (*h* or *l*) or from specific molecules (*HW*).

ingly, this shift in backbone motion is accompanied by a change in the consequent methyl group oscillation from ~30 to 110 cm^{-1} in response to the increased protein main chain solvent interactions.

In our analyses of lysozyme dynamics as a function of hydration we have established that the protein intramolecular interactions are more significant at lower hydration levels. We have also determined that as the hydration level is increased there are a greater number of correlated motions taking place between residues in distant parts of the protein secondary structure. Specifically, we have observed that additional water molecules added to the lysozyme hydration shell trigger large amplitude correlated motion between residues in helix A with those in the 3_{10} helix in the lysozyme α lobe. The correlated motion in the protein at the high hydration level is accompanied by an enhancement of solvent side chain hydrogen bonds. Perhaps it is the combination of these two characteristics that is significant in activating fast backbone fluctuations that are not visible in the lower hydration protein systems. It is probable that at lower hydration, strong intramolecular protein associations inhibit the formation of anharmonic fluctuations in the protein secondary structure.

D. Energy dissipation and anharmonic fluctuations within the protein secondary structure

To comprehend the relationship between picosecond time scale conformational fluctuations taking place in the protein secondary structure and energy dissipation, a nonlinear molecular dynamics (NMD) simulation has been conducted on both the low and high hydration protein to determine the dominate motions that are activated as thermal energy moves through the molecule. Our interest in the nonlinear MD experiment is to identify the fast relaxation processes in the protein that may be set in motion from an excitation equivalent to energy released during macromolecular substrate binding or perhaps when a chemical bond in the protein is broken. From this analysis we have mapped the sequence of motions in the secondary structure that are triggered as the excess thermal energy, introduced in the simulation, moves through the molecule [52]. From this analysis we have determined that greater interaction with the solvent is in an important factor in determining both the rate of the thermal energy dissipation and the pathways formed for the dissipation process. From the MD simulation we have discerned that the general pathway of energy flow through the molecule is as follows: the initial response to the excitation involves the C=O stretching of the Amide I vibration of residues in the turn and loop regions of the protein secondary structure (1684 cm^{-1}) along with low-frequency helical deformations, which include methyl group torsional fluctuations in the 3_{10} helix (in the protein α lobe) at 110 cm^{-1} and a collective main chain bending vibration at 130 cm^{-1} created by intrahelical dipole correlations in helix A (Table I). In the 100–200 ps time frame after excitation, the energy moves from the Amide I vibration of residues in the turn and loop regions of the protein secondary structure to the Amide I vibration of α helices. During this time frame, the bending mode, arising from collective intramolecular correlations taking place within helix A (130 cm^{-1}), continue to grow in amplitude and are accompanied by a localized C_α—H bending vibration at 1327 cm^{-1} that also takes place in the α helix. In the time range spanning 200–400 ps energy moves from Amide I vibration into the Amide II region of the protein secondary structure. The Amide II vibrational response at 1493 cm^{-1} is greater in the lower hydration simulation protein and it is at this point that the pathway for energy dissipation in the two systems begins to diverge. In the low hydration sample, the vibrational energy becomes trapped in the Amide II vibrational mode while in the highly hydrated protein the energy moves efficiently from the backbone modes into the side chain methyl groups of the protein. Perhaps the underlying reason for the deviation can be attributed to the strong intramolecular association taking place in the protein at low hydration; and as a consequence, a weaker interaction with the solvent when compared with the highly hydrated protein. In the early phase of the energy dissipation scheme we find that collective modes associated with side chain methyl group oscillations in the 3_{10} helix (110 cm^{-1}) and the intramolecular main chain hydrogen bonding taking place within helix A (130 cm^{-1}) are important for moving energy out of the protein. Both mechanisms are driven by the proteins increased interaction with the solvent. In the preced-

TABLE II. A list of several of the prominent vibrational frequencies and their assignments uncovered in the MD simulations of HEWL.

Frequency (cm^{-1})	Assignment	Frequency (cm^{-1})	Assignment
1684	Amide I turns and loops	167	α helix main chain deformation
1659	Amide I α helix	130	Intramolecular main chain backbone deformation
1493	Amide II	120	Main chain solvent deformation
1327	α helix Cα—H main chain deformation	110	Methyl group torsion
1161	Side chain methyl group rocking	75	Backbone torsion
1010	Phenyl group ring breathing	63	Intramolecular backbone bending
995	Out of phase phenyl group vibration	50	α helix out of plane backbone deformation
662	In phase phenyl group vibration	41	α helix backbone fluctuation
336	α helix main chain bending	25	Out of plane backbone deformation coupled with side chain torsion

ing sections we determined that these associations were characteristic of dynamics only found in the highly hydrated protein. Therefore, we postulate that these anharmonic fluctuations taking place in the protein are triggered by side chain solvent interactions. The solvent side chain interactions greatly increase the amplitude of the methyl group oscillations in the 3_{10} helix. And through nonlinear coupling among residues in the secondary structure, they are also responsible for the intrahelical associations in helix A that are activated when additional water molecules are included in the protein hydration shell. 200 ps after the initial quench, low-frequency backbone torsional fluctuations in helix A, at approximately 40 cm^{-1} and 50 cm^{-1}, in the well hydrated protein, are initiated. The fast backbone fluctuations couple with the Amide I vibration at 1659 cm^{-1}. It is possible that these low frequency collective modes are vital for moving the excess energy out of the protein. The collective backbone modes do not occur in the low hydration protein sample dynamics and as a result, beginning at $t>200$ ps the drier protein begins to experience a "back up" of vibrational energy flow. Stronger intramolecular interactions in the low hydration protein cause energy to dissipate along a different pathway. For instance, in the low hydration protein, we have determined that one of the main avenues for excess energy transfer is through intramolecular ring deformations of the protein nonpolar side chains (995 and 662 cm^{-1}). The greater intensity and lifetime of the Amide II vibration in the low hydration protein also attests to the stronger intramolecular backbone interactions that ultimately inhibit efficient transfer of the excess energy out of the protein. During the same time frame, in the fully hydrated protein, collective methyl group fluctuations of the protein side chains are effective in moving energy out of the protein. This process is initiated with methyl group oscillations in the 3_{10} helix, where solvent side chain interactions in the protein are highest. The energy eventually moves through the protein via torsional motion of side chain methyl groups of the other residues in the protein, as illustrated by the band that appears in the IR spectrum at 1161 cm^{-1} during this time period.

From the FCA of the equilibrium MD simulations described in an earlier section, it has been determined that VAL 120 is engaged in anharmonic interatomic interactions in the high hydration system (Table II). The nonlinear contribution to the protein dynamics from the residue is not present in the low hydration system. At high hydration, VAL 120 has nonlinear associations with residues in helix A as well as with residues in the (α lobe) 3_{10} helix. For this reason we would like to use the response of this residue as a window to monitor the important collective dynamics that occur in response to the excitation and to determine the role of nonlinear processes in the energy dissipation mechanism. In VAL 120 the initial response to the thermal energy excitation includes modes at 120, 50, and 40 cm^{-1} [Fig. 7(b)]. The collective vibrational modes at 40 and 50 cm^{-1} are attributed to fast backbone fluctuations taking place in helix A; while the mode at 120 cm^{-1} is ascribed to a protein backbone mode that is highly influenced by solvent hydrogen bonding dynamics. We find a large peak at this frequency in the VACF spectrum highlighting the librational motion that results from the protein backbone-solvent hydrogen bonds [Fig. 8(a)]. We have also uncovered a mode at a similar frequency in our experimental THz spectrum of lysozyme [Figs. 8(b)–8(e)]. In the experimental spectrum, the band at $\sim$120 cm^{-1} has a higher intensity in the high hydration sample when compared to the lowest hydrated sample studied, suggesting a much stronger backbone interaction with the solvent. Interestingly, in a structural analysis of the protein simulation at both low and high hydration we have found that in the drier protein VAL 120 exists as a residue at the end of a turn region in the protein secondary structure. In the high hydration simulation of lysozyme, VAL 120 is the first residue in the 3_{10} helix. The transformation of VAL120 from a residue forming part of a turn to a residue located in a helix implies that there is at least a slight modification of the hydrogen bonding within the protein secondary structure that is initiated when more water molecules are introduced into the protein hydration shell. Using VAL 120 as a guide to follow the role of collective fluctuations in the energy dissipation process in the drier

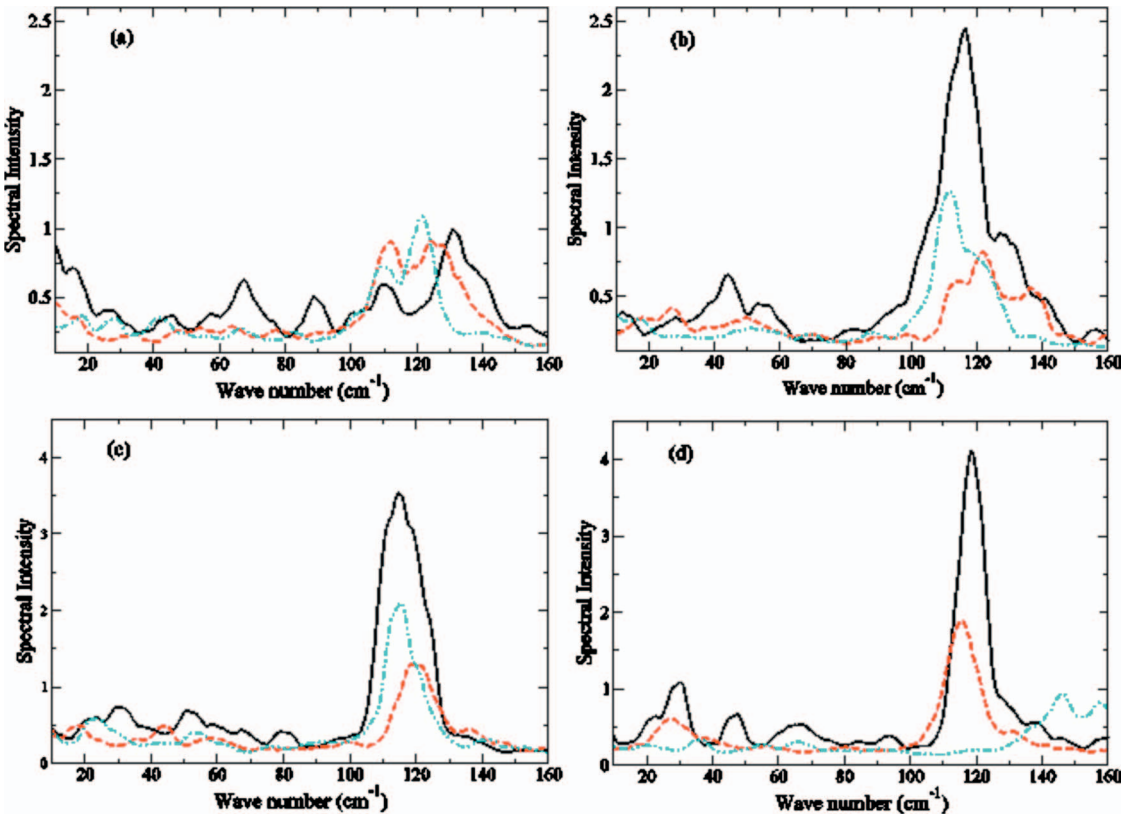

FIG. 7. (Color) VAL 120 VACF spectrum from the nonequilibrium MD simulation (a) at low hydration and (b) at high hydration 200 ps after excitation (black solid line), 400 ps after excitation (cyan dot-dot-dashed line), and 600 ps after excitation (red dashed line). GLY 102 VACF spectrum from the nonequilibrium MD simulation (c) at low hydration and (d) at high hydration 200 ps after excitation (black solid line), 400 ps after excitation (cyan dot-dot-dashed line), and 600 ps after excitation (red dashed line).

protein [Fig. 7(a)], we find that the early response to the excitation is dominated by intramolecular modes. Specifically, we observe modes at ~63, 85, and 165 cm^{-1}; all of which are indicative of collective fluctuations that are dominated by protein *intramolecular* main chain backbone interactions. We find further confirmation of the strong protein intramolecular correlations taking place in the low hydration sample in the THz experimental spectrum in Figs. 8(b) and 8(e). In the experimental spectrum only the driest protein has a prominent band at 165 cm^{-1}, reflecting intramolecular backbone main chain interactions. In contrast, the high hydration sample has strong solvent interactions, which are particularly evident through the side chain solvent hydrogen bonds. It is possible that this close association with the solvent dynamics is essential for choreographing other anharmonic correlations within the protein structure. For instance, we have deduced that solvent interaction with the 3_{10} helix is the trigger for the anharmonic methyl group torsional fluctuations (~110 cm^{-1}) that are important for initiating the energy transfer process in the molecule. Similarly, the intrahelical associations detected in helix A (130 cm^{-1}) and the fast backbone fluctuations within the helix (40 and 50 cm^{-1}) develop as a result of the localized intrapeptide associations that are modulated by the solvent. In the experimental spectrum there is a band close to 110 cm^{-1} in the intermediate and highly hydrated protein that we construe is related to methyl group fluctuations [Figs. 8(b)–8(d)]. Its absence in the low hydration sample of the protein and its temperature dependence [Figs. 8(c)–8(e)] are characteristic of a mode that is highly modulated by the solvent. We also observe a shoulder in the THz spectrum at 140 cm^{-1} at the coldest temperature studied that is also only present in the high and intermediate hydrated protein sample. We believe that the mode at 140 cm^{-1} is associated with the intrahelical dipole correlations that take place in helix A when additional solvent molecules are introduced into the protein hydration shell. The presence of both the intrapeptide mode at 140 cm^{-1} and backbone mode at 50 cm^{-1} at 93 K in Fig. 8 suggests that the structure of the protein is slightly altered in a high solvent environment when compared with the dehydrated protein.

From the simulation it has become evident that in the drier protein, solvent interaction with the protein is most prominent in the long loop region rather than with the solvent accessible side chains as has been observed in the fully hydrated protein system. From our previous analyses, we have established that hydration is important for forming correlated motion among residues in distant parts of the protein. In both the low and highly hydrated protein, motion of GLY 102, a residue in the long loop region of the protein secondary structure, adds a nonlinear component to the protein dynamics. Nonetheless, from our analysis of the energy dissipation process, we have determined that although the solvent accessible surface of the residue is similar in both protein

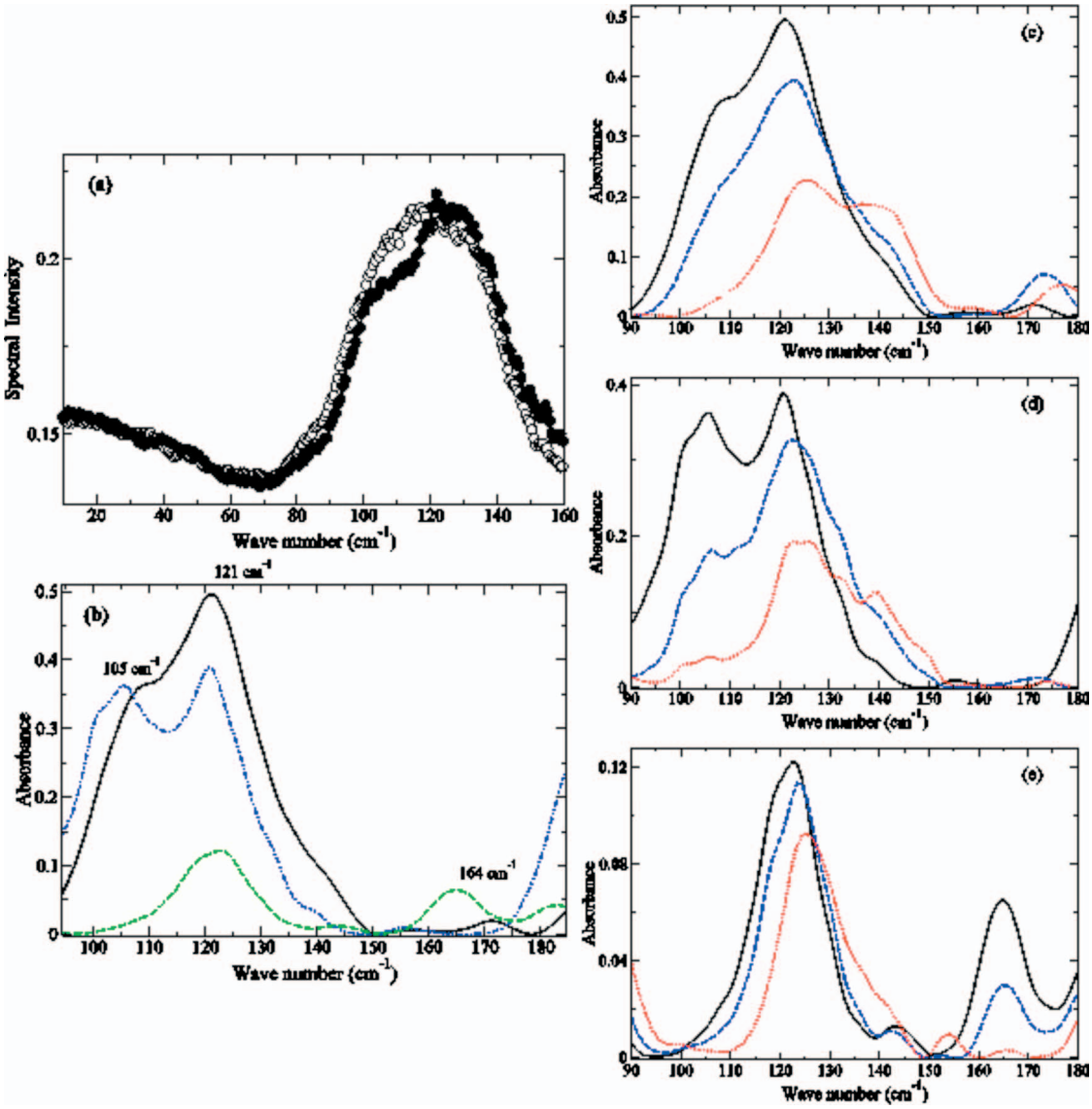

FIG. 8. (Color) (a) VACF spectrum of protein backbone—solvent hydrogen bonds from the MD simulation of lysozyme at low hydration (solid circles) and at high hydration (open circles). (b) Experimental THz spectrum of room-temperature lysozyme in the 90–180 cm^{-1} spectral region at 75% RH (green dashed line), 85% RH (blue dot-dot-dashed line) and 97% RH (black solid line). Experimental THz spectrum of lysozyme in the 90–180 cm^{-1} spectral region at (c) 97% RH, (d) 85% RH, and (e) 75% RH and at temperatures 93 K (red dotted line), 250 K (blue dashed line), and at 303 K (black solid line).

systems, the rate of energy dissipation is vastly dissimilar. In Figs. 7(c) and 7(d), we find that within 600 ps the excess energy has completely dissipated from the highly hydrated protein backbone mode at 120 cm^{-1}, but within that same time frame energy lingers in the drier protein backbone mode. Perhaps the progression of the energy transfer process from the backbone mode makes more sense if one considers the residues to which it is anharmonically coupled. In the high hydration protein, GLY 102 is anharmonically coupled to residues in the 3_{10} helix in which side chain interactions with solvent are maximized. In the drier sample, GLY 102 has the strongest correlation with a residue in a partially buried turn region in the protein secondary structure. Hence, in the latter case, GLY 102 does not possess strong coupling with the solvent. This particular aspect of the protein dynamics seems to be essential in promoting the other nonlinear phenomena observed in the highly hydrated protein sample. From our investigation we have determined that nonlinearity in the protein is promoted by interaction of intrapeptide helical vibrations, Amide I vibrations, and local displacements such as those identified at 110, 50, and 40 cm^{-1}. The lack of interaction with the solvent also provides an explanation for the "build up" of vibrational energy observed in IR spectrum of the low hydration protein in response to the excess thermal energy introduced in the protein in the nonlinear MD simulation.

E. Protein localized rigidity and energy dissipation pathways

In the work carried out during this investigation to explore the picosecond time scale dynamics in lysozyme, we have uncovered an unexpected connection between the root–

mean-square deviation (RMSD) of residues in the protein secondary structure and energy dissipation. For instance, although the main chain solvent accessible surface (SAS) of VAL 120 is much greater in the highly hydrated protein, the RMSD of the residue is much lower in the fully hydrated sample when compared with that of the drier protein. In the highly hydrated protein we find that residues 118–119 and 121–123, the amino acids preceding and succeeding VAL 120 have much higher RMSD suggesting that VAL 120 is a rigid site in an otherwise flexible region of the protein. In the low hydration protein sample VAL 120 remains as a flexible residue in a highly flexible turn region of the protein. We have also found that overall, the loop and turn regions in the protein secondary structure are much more flexible at low hydration when compared to the highly hydrated protein. Furthermore, the nonlinear associations that we have uncovered in the highly hydrated protein are almost entirely absent in low hydration protein dynamics. For this reason, we speculate that rigidity localized within certain regions of the protein secondary structure is required for forming channels for energy transport within the protein structure. The nonlinear coupling of spatially overlapping modes in resonance may play an important role in promoting a delocalization mechanism that facilitates the energy transfer process across the protein. Our results are in line with a computational study [53] on the thermal transport coefficient of another globular protein, myoglobin. The results of this earlier investigation have revealed that anharmonicity enhances heat flow by transferring energy among protein localized normal modes. In the fully hydrated protein we have identified anharmonic fluctuations that arise in response to the nonlinear interactions between VAL 120 with residues in helix A and the 3_{10} helix in the protein α lobe. GLY 102, a residue in the solvent accessible loop region in the protein secondary structure, has also been linked with nonlinear associations with residues in the same 3_{10} helix in the highly hydrated protein. The formation and interaction of these nonlinear localized modes in the protein secondary structure may provide a vibrational energy channel that is capable of altering energy barriers on a short time scale (picoseconds) and may have import implications for chemical reactions and/or enzymatic catalysis taking place in the protein on a much longer time scale.

IV. CONCLUSIONS

In this work, we have used infrared spectroscopy to probe the collective vibrations in lysozyme under varying hydration and temperature conditions. Numerous experimental and theoretical studies have demonstrated that the solvent dynamics in the protein hydration shell have a strong influence on protein function [54–58]. From the experimental and MD simulation studies carried out during this investigation on lysozyme, we have determined that the low hydration protein is dominated by intramolecular interactions and in turn, the low-frequency dynamics are mostly governed by localized methyl group fluctuations. Our findings are consistent with previous neutron scattering experiments [32,59] on lysozyme and other model peptides; where under dry or low hydration conditions the protein dynamics are found to be confined to rotational motion of the methyl groups. At higher hydration we report the activation of collective backbone fluctuations in the protein secondary structure and the formation of anharmonic side chain fluctuations that are not present in the low hydration protein. Both types of collective excitations are a direct result of the increased interaction of the protein with the solvent in the highly hydrated lysozyme sample. Additionally, the solvent appears to play a central role in orchestrating communication throughout the entire protein by means of correlated fluctuations. Specifically, we have observed that fast, nonlinear dynamical fluctuations connect distant parts of the protein; that consequently, also provide a pathway for the transport of energy throughout the protein.

ACKNOWLEDGMENT

Carnegie Mellon University Berkman Faculty Development Fund.

[1] A. Bakan and I. Bahar, Proc. Natl. Acad. Sci. U.S.A. **106**, 14349 (2009).
[2] H. J. C. Berendsen and S. Hayward, Curr. Opin. Struct. Biol. **10**, 165 (2000).
[3] A. J. Wand, Nat. Struct. Mol. Biol. **8**, 926 (2001).
[4] P. W. Fenimore, H. Frauenfelder, B. H. McMahon, and F. G. Parak, Proc. Natl. Acad. Sci. U.S.A. **99**, 16047 (2002).
[5] A. Tomita, T. Sato, K. Ichiyanagi, S. Nozawa, H. Ichikawa, M. Chollet, F. Kawai, S.-Y. Park, T. Tsuduki, T. Yamato, S.-Y. Koshihara, and S.-I. Adachi, Proc. Natl. Acad. Sci. U.S.A. **106**, 2612 (2009).
[6] S. Hammes-Schiffer, Biochemistry **41**, 13335 (2002).
[7] H. Frauenfelder, S. Sligar, and P. Wolynes, Science **254**, 1598 (1991).
[8] J. Fitter, R. E. Lechner, and N. A. Dencher, Biophys. J. **73**, 2126 (1997).
[9] H.-X. Zhou, S. T. Wlodek, and J. A. McCammon, Proc. Natl. Acad. Sci. U.S.A. **95**, 9280 (1998).
[10] L. D. Barron, L. Hecht, and G. Wilson, Biochemistry **36**, 13143 (1997).
[11] S. Fischer, J. C. Smith, and C. S. Verma, J. Phys. Chem. B **105**, 8050 (2001).
[12] D. M. Leitner, Annu. Rev. Phys. Chem. **59**, 233 (2008).
[13] E. Cornicchi, G. Onori, and A. Paciaroni, Phys. Rev. Lett. **95**, 158104 (2005).
[14] L. Zhang, L. Wang, Y.-T. Kao, W. Qiu, Y. Yang, O. Okobiah, and D. Zhong, Proc. Natl. Acad. Sci. U.S.A. **104**, 18461 (2007).
[15] E. Lindahl, B. Hess, and D. der van Spoel, J. Mol. Model. **7**, 306 (2001).
[16] H. J. C. Berendsen, J. P. M. Postma, W. F. v. Gunsteren, A. DiNola, and J. R. Haak, J. Chem. Phys. **81**, 3684 (1984).

[17] H. B. Berk Hess, J. C. Herman Berendsen, and Johannes G. E. M. Fraaije, J. Comput. Chem. **18**, 1463 (1997).
[18] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, J. Chem. Phys. **103**, 8577 (1995).
[19] T. Darden, D. York, and L. Pedersen, J. Chem. Phys. **98**, 10089 (1993).
[20] W. Humphrey, A. Dalke, and K. Schulten, J. Mol. Graph. **14**, 33 (1996).
[21] H. J. Bakker and J. L. Skinner, Chem. Rev. **110** (3), 1498 (2010).
[22] K. Toukan and A. Rahman, Phys. Rev. B **31**, 2643 (1985).
[23] J. Marti, E. Guardia, and J. A. Padro, J. Chem. Phys. **101**, 10883 (1994).
[24] F. Paesani, W. Zhang, D. A. Case, T. E. Cheatham III, and G. A. Voth, J. Chem. Phys. **125**, 184507 (2006).
[25] S. Amira, D. Spångberg, and K. Hermansson, Chem. Phys. **303**, 327 (2004).
[26] J. Lobaugh and G. A. Voth, J. Chem. Phys. **106**, 2400 (1997).
[27] 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).
[28] U. Heugen, G. Schwaab, E. Bründermann, M. Heyden, X. Yu, D. M. Leitner, and M. Havenith, Proc. Natl. Acad. Sci. U.S.A. **103**, 12301 (2006).
[29] O. Lange and H. Grubmüller, Proteins: Struct., Funct., Bioinf. **70**, 1294 (2008).
[30] I. D. Kuntz, Jr. and W. Kauzmann, Adv. Protein Chem. **28**, 239 (1974).
[31] M. Tarek and D. J. Tobias, Biophys. J. **79**, 3244 (2000).
[32] J. H. Roh, V. N. Novikov, R. B. Gregory, J. E. Curtis, Z. Chowdhuri, and A. P. Sokolov, Phys. Rev. Lett. **95**, 038101 (2005).
[33] W. Doster, S. Cusack, and W. Petry, Nature (London) **337**, 754 (1989).
[34] M.-C. Bellissent-Funel, J. Mol. Liq. **84**, 39 (2000).
[35] J. A. Rupley and G. Careri, Adv. Protein Chem. **41**, 37 (1991).
[36] J. Perez, J.-M. Zanotti, and D. Durand, Biophys. J. **77**, 454 (1999).
[37] M. Diehl, W. Doster, W. Petry, and H. Schober, Biophys. J. **73**, 2726 (1997).
[38] K. D. Moeller, G. P. Williams, S. Steinhauser, C. Hirschmugl, and J. C. Smith, Biophys. J. **61**, 276 (1992).
[39] H. Urabe, Y. Sugawara, M. Ataka, and A. Rupprecht, Biophys. J. **74**, 1533 (1998).
[40] G. Giraud, J. Karolin, and K. Wynne, Biophys. J. **85**, 1903 (2003).
[41] L. Genzel, F. Keilmann, T. P. Martin, G. Winterling, Y. Yacoby, H. Frohlich, and M. Makinen, Biopolymers **15**, 219 (1976).
[42] T. Simonson and D. Perahia, Faraday Discuss. **103**, 71 (1996).
[43] H. Schwalbe, S. B. Grimshaw, A. Spencer, M. Buck, J. Boyd, C. M. Dobson, C. Redfield, and L. J. Smith, Protein Sci. **10**, 677 (2001).
[44] M. Walther, P. Plochocka, B. Fischer, H. Helm, and P. U. Jepsen, Biopolymers **67**, 310 (2002).
[45] D. Tobi and I. Bahar, Proc. Natl. Acad. Sci. U.S.A. **102**, 18908 (2005).
[46] K. N. Woods and H. Wiedemann, Chem. Phys. Lett. **393**, 159 (2004).
[47] W. Nusser and R. Kimmich, J. Phys. Chem. **94**, 5637 (1990).
[48] F. Parak and E. W. Knapp, Proc. Natl. Acad. Sci. U.S.A. **81**, 7088 (1984).
[49] M. Ferrand, A. J. Dianoux, W. Petry, and G. Zaccaï, Proc. Natl. Acad. Sci. U.S.A. **90**, 9668 (1993).
[50] A. Ostermann, R. Waschipky, F. G. Parak, and G. U. Nienhaus, Nature (London) **404**, 205 (2000).
[51] G. Diakova, Y. A. Goddard, J.-P. Korb, and R. G. Bryant, J. Magn. Reson. **189**, 166 (2007).
[52] See supplementary material at http://link.aps.org/supplemental/10.1103/PhysRevE.81.031915 for figures of the calculated midinfrared spectra of lysozyme from the nonequilibrium MD simulations.
[53] X. Yu and D. M. Leitner, J. Chem. Phys. **122**, 054902 (2005).
[54] G. Zaccai, Philos. Trans. R. Soc. Lond. B Biol. Sci. **359**, 1269 (2004).
[55] P. W. Fenimore, H. Frauenfelder, B. H. McMahon, and R. D. Young, Proc. Natl. Acad. Sci. U.S.A. **101**, 14408 (2004).
[56] J. Swenson, H. Jansson, J. Hedstrom and R. Bergman, J. Phys. Condens. Matter **19**, 205109 (2007).
[57] V. Kurkal, R. M. Daniel, J. L. Finney, M. Tehei, R. V. Dunn, and J. C. Smith, Biophys. J. **89**, 1282 (2005).
[58] H. Nakagawa, Y. Joti, A. Kitao, and M. Kataoka, Biophys. J. **95**, 2916 (2008).
[59] D. Russo, G. L. Hura, and J. R. D. Copley, Phys. Rev. E **75**, 040902 (2007).