Spatio-temporal hierarchy in the dynamics of a minimalist protein model

Yasuhiro Matsunaga, Akinori Baba, Chun-Biu Li, John E. Straub, Mikito Toda, Tamiki Komatsuzaki, and R. Stephen Berry

Citation: The Journal of Chemical Physics **139**, 215101 (2013); doi: 10.1063/1.4834415
View online: http://dx.doi.org/10.1063/1.4834415
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/139/21?ver=pdfcov
Published by the AIP Publishing

Spatio-temporal hierarchy in the dynamics of a minimalist protein model

Yasuhiro Matsunaga,[1] Akinori Baba,[2] Chun-Biu Li,[3] John E. Straub,[4] Mikito Toda,[5] Tamiki Komatsuzaki,[3,a)] and R. Stephen Berry[6,b)]
[1]*RIKEN, Advanced Institute for Computational Science, 7-1-26 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan*
[2]*RIKEN, The Center for Developmental Biology, 2-2-3 Minatojima-minamimachi, Chuo-ku, Kobe 650-0047, Japan*
[3]*Research Institute for Electronic Science, Hokkaido University, Kita 20 Nishi 10, Kita-ku, Sapporo 001-0020, Japan*
[4]*Department of Chemistry, Boston University, Boston, Massachusetts 02215, USA*
[5]*Physics Department, Nara Women's University, Kita-Uoya-Nishimachi, Nara 630-8506, Japan*
[6]*Department of Chemistry, University of Chicago, Chicago, Illinois 60637, USA*

(Received 1 August 2013; accepted 13 November 2013; published online 3 December 2013)

A method for time series analysis of molecular dynamics simulation of a protein is presented. In this approach, wavelet analysis and principal component analysis are combined to decompose the spatio-temporal protein dynamics into contributions from a hierarchy of different time and space scales. Unlike the conventional Fourier-based approaches, the time-localized wavelet basis captures the vibrational energy transfers among the collective motions of proteins. As an illustrative vehicle, we have applied our method to a coarse-grained minimalist protein model. During the folding and unfolding transitions of the protein, vibrational energy transfers between the fast and slow time scales were observed among the large-amplitude collective coordinates while the other small-amplitude motions are regarded as thermal noise. Analysis employing a Gaussian-based measure revealed that the time scales of the energy redistribution in the subspace spanned by such large-amplitude collective coordinates are slow compared to the other small-amplitude coordinates. Future prospects of the method are discussed in detail. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4834415]

I. INTRODUCTION

Energy flow,[1,2] such as anisotropic heat diffusion through protein residues, or vibrational energy transfer through the vibrational states of a protein, has long motivated experimental and computational studies toward its role in photosensing,[3–6] photosynthesis,[7] ligand-binding and dissociation,[8–13] and allostery.[14–16] For example, the existence of particular pathways of anisotropic heat diffusion was proposed to explain the allostery of the proteins belonging to the PDZ domain family.[15,16] In terms of the vibrational energy transfer,[17–19] time-resolved absorption spectroscopies have revealed the anisotropic energy transfers among peptide's or protein's collective modes.[20,21]

In general, energy flow pathways and the time scale of the equipartitioning of energy influence the kinetics of chemical reactions, such as the rate of conformational change of a protein.[22,23] In this regard, developing a method to identify and classify the energy flows in protein molecules has an impact in elucidating the mechanism of protein functions. To investigate such *nonstationary*, *anharmonic* dynamics of proteins, molecular dynamics (MD) simulation has been a powerful tool. Thus far, studies characterizing MD trajectories have focused primarily on static aspects or equilibrium distribution properties such as the free energy profile. For nonstationary and anharmonic dynamics, such approaches are not sufficient to characterize the process because the amplitudes and frequencies of the dynamics vary along the time evolution. Methods based on the time correlation functions or their frequency-domain representations (e.g., power spectra) suffer from the same problem since the assumption of the stationarity for the underlying process is not valid, particularly when relaxation by the excess energy transfer is the subject of investigation. To overcome this issue, various approaches, based on time-localized representations for nonstationary processes, have been proposed, including instantaneous normal mode analysis,[24,25] moving normal mode coordinates,[26] and windowed-Fourier approaches.[27,28]

Protein dynamics are often investigated in the framework of the (generalized) Langevin equation where one can reduce the complex Newtonian dynamics of the atoms of the protein/solvent to the simple stochastic dynamics of a few degrees of freedom. In this approach, the energy transfer is interpreted as a dissipation due to a friction from the predefined "system" to the surrounding environment ("bath"). The rate of energy transfer can be related to the friction which can be estimated via the autocorrelation function of velocities (or the mean square displacements of coordinates). This approach, however, is not suitable for the detection of the energy flow *pathways* of interest in this study, because it is assumed that the excess energies always dissipate from the *predefined* system to the bath on average.

In this article, we present a wavelet-based approach,[29] focusing on a characterization of the vibrational energy transfer through the vibrational states of a protein. Unlike the

a)Electronic mail: tamiki@es.hokudai.ac.jp
b)Electronic mail: berry@uchicago.edu

conventional Fourier-based approaches, the wavelet transform provides a set of orthogonal basis functions in the time-frequency domain, and thus is suitable to detect the vibrational energy transfers. We demonstrate the method by applying it to time series derived from MD simulations of a coarse-grained minimalist protein model which is a 3-color, 46-bead model protein whose potential and free energy landscapes, kinetics, and dynamics have been well studied.[30–34]

While earlier studies that used the wavelet transform to analyze MD trajectories of proteins focused on identifying or clustering conformational states,[35–39] we use the approach to identify vibrational states and measure time scales for equipartitioning vibrational energy. To this end, we first apply the principal component analysis (PCA)[40–43] to decompose the various space scale contributions from the original multivariate MD trajectory. Then we calculate the momenta conjugate to the principal components (PCs) and analyze those using multiresolution decomposition via the discrete wavelet transform. By combining the methods in this way, the vibrational states and energy redistribution of the spatio-temporal dynamics of a protein are naturally characterized in a hierarchical manner. In the analysis, we postulate the low viscosity regime in condensed phase in which the system is weakly coupled to a heat bath by Berendsen's thermostat with a relatively small coupling time.

We find the vibrational energy transfers between the fast and slow time scales are clearly observed in a set of large-amplitude PCs during the folding/unfolding transitions in the minimalist protein model, while the rest of the small amplitude PCs behave as thermal noise. After identifying the states, we evaluate the time scale for equipartitioning of vibrational energy using a measure of the extent to which the distribution is Gaussian ("Gaussianity") based on the skewness of the time-localized distribution function.

II. METHODS

A. Multiresolution decomposition of collective momenta

How do biomolecules having several time and space scales evolve through the state space? The state space here involves not only the conformational degrees of freedom, on which the overdamped Langevin formulation is based, but also on their changes in time, that is, *momenta*. There is no general answer of what kind(s) of coordinate(s) or projection(s) is(are) best to represent collective motions of multidimensional systems. Among several possible candidates, we chose PCs which have been often used to reduce the dimensionality of multivariate time series because of their simplicity.[40–43]

The technical idea behind the PCA is to find the orthogonal eigenvectors that capture the large variances in the multivariate time series with the highest amount of information. Consider the variance-covariance matrix $\mathbf{R}$ of the Cartesian coordinates $\mathbf{q}$ (or mass-weighted coordinates $\mathbf{M}^{1/2}\mathbf{q}$) of the particle positions. For simplicity, we assume that each coordinate has the mean position as its origin. The variance-covariance matrix is a semi-definite positive matrix, and one can find the eigenvectors U that diagonalize $\mathbf{R}$

$$\mathbf{U}^{-1}\mathbf{R}\mathbf{U} = \mathbf{r}, \ (\mathbf{U}^T\mathbf{U} = \mathbf{I}).$$

The eigenvalue r_i, the ith element of the diagonal matrix $\mathbf{r}$, represents the variance of the ith PC, Q_i, defined as a linear combination of the original coordinates,

$$\mathbf{Q} = \mathbf{U}^T\mathbf{q}. \tag{1}$$

The larger the value of r_i, the better the ith PC, Q_i, represents the variance of the distribution of the system traversing through the high-dimensional conformational space. The $\{Q_i\}$ are sorted in the decreasing order of the variance, $r_1 \geq r_2 \geq, \ldots, r_{3N} \geq 0$ (where N is the number of particles).

The conjugate momentum P_i to the Q_i is derived from a canonical transformation from the old conjugate variables $\mathbf{q}$ and $\mathbf{p}$ to the new conjugate variables $\mathbf{Q}$ and $\mathbf{P}$. Consider a generating function F defined by[44]

$$F(\mathbf{q}, \mathbf{P}) = \sum_{i=1}^{3N}\sum_{j=1}^{3N} U_{ji} q_j P_i.$$

From Eq. (1), we derive that

$$\frac{\partial F}{\partial P_i} = Q_i = \sum_{j=1}^{3N} U_{ji} q_j,$$

and

$$\frac{\partial F}{\partial q_j} = p_j = \sum_{i=1}^{3N} U_{ji} P_i.$$

The new momenta $\mathbf{P}$ conjugate to $\mathbf{Q}$ are then given by

$$\mathbf{P} = \mathbf{U}^T\mathbf{p}.$$

If the harmonic approximation holds, the motion of the Q_i of larger variance tends to be slower and the vibrational frequency of the corresponding momentum P_i is invariant during the course of time evolution. On the other hand, with anharmonic dynamics, because of the time dependence of the amplitudes and frequencies of the motions, the fast and slow time scale modes can mix even in a single Q_i.

In this article, we propose a multiresolution decomposition of the collective momentum $P_i(t)$ based on orthogonal wavelets.[29,45] This way, a given P_i is decomposed into contributions from a hierarchy of different time scales,

$$P_i(t) = \sum_{j=1}^{n} d_i^{(j)}(t) + a_i(t). \tag{2}$$

Each of the $d_i^{(j)}(t)$ is called the *detail* at the jth *level* of time scale $2^j\Delta t$, and is given by

$$d_i^{(j)}(t) = \sum_{k=0}^{(N_T\Delta t)/(2^j\Delta t)-1} d_{ik}^{(j)}\psi_k^{(j)}(t),$$

where N_T is the total time step of the trajectory $N_T = 2^n$, n is an arbitrary integer), and Δt is the sampling time. $\{d_{ik}^{(j)}\}$ are

the wavelet coefficients of P_i, defined by

$$d_{ik}^{(j)} = \int_0^{N_T \Delta t} P_i(t) \psi_k^{(j)}(t)\, dt.$$

Here, $\{\psi_k^{(j)}(t)\}$ are the wavelet bases, which form a set of orthogonal bases in $L^2(\mathbb{R})$, and are given by scaled and translated versions of a single mother wavelet $\Psi(t)$,

$$\psi_k^{(j)}(t) = \frac{1}{\sqrt{2^j \Delta t}} \Psi\left(\frac{t - k2^j \Delta t}{2^j \Delta t}\right).$$

Increasing j to $j+1$ enlarges the function $\psi_k^{(j)}$ by a factor of two, and thus the time scale of changes in the amplitude of the momentum $P_i(t)$ (making the most important contribution to the jth level of time scale) is $2^j \Delta t$. In practice, the multiresolution decomposition is truncated by a finite level n ($<\log_2 N_T$), although the maximum level is strictly determined by the total time steps N_T, with the relation $n \sim \log_2 N_T$. In this article, $n = 10$ is used, which will be later found to be sufficient to capture most of the vibrational motions in the minimalist protein model. $a_i(t)$ in Eq. (2) is called the *approximation*, which includes all the contributions to the time dependence of $P_i(t)$ on levels higher than the truncation level n. It is often expected that $a_i(t)$ corresponds to aperiodic diffusive motion (e.g., Ref. 45). As the mother wavelet, the eight coefficient wavelets of Daubechies[29,45] are employed (see Fig. 1). These functions are nonzero only in a finite interval (i.e., finite support) and behave well with respect to time-localization. The computational complexity of the discrete wavelet transform is $O(N)$, superior to that of the fast Fourier transform algorithm, $O(N \log N)$, used in the conventional analyses. Also, the data parallelization of the code for the set of momenta is straightforward.

B. Kinetic temperature of collective coordinates over scales

The multiresolution decomposition of collective momenta is expected to shed light on the time scale of the equipartitioning of vibrational energy, or the redistribution of the energies between different scales of the collective motion. Before going into the characterization of the temporal behavior of multiple levels, we first define the time-averaged property of the kinetic energies of collective motions. The averaged kinetic energy, or kinetic temperature over scales,[45] from the output of the wavelet analysis, in analogy with conventional power spectra, provides us with a decomposition of the temperature of the system over the underlying time scales of its dynamics.

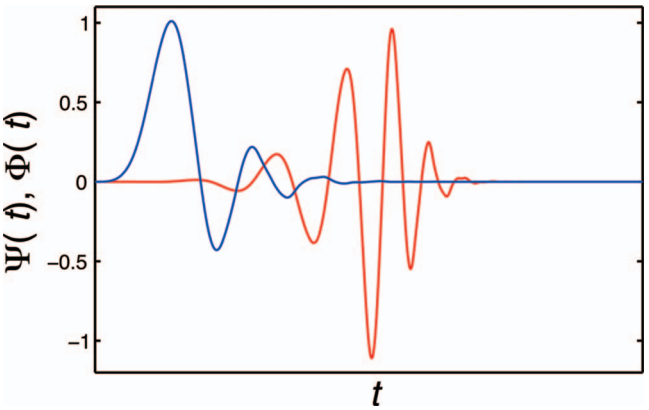

FIG. 1. Mother wavelet function of the eight coefficient Daubechies wavelet $\Psi(t)$ (red) and the corresponding scaling function $\Phi(t)$ (blue). Since the Daubechies wavelet has no explicit expression, the functions were approximated by the iterations of inverse discrete wavelet transform.

The kinetic energy of the system is rewritten in terms of the momenta $\mathbf{P}$ conjugate to the PCs $\mathbf{Q}$,

$$K(t) = \frac{1}{2m} \sum_{i=1}^{3N} p_i^2(t) = \frac{1}{2m} \sum_{i=1}^{3N} P_i^2(t).$$

Here, we used the orthogonal property of the PCs and assumed that the masses are uniform. See Appendix A for the derivation of the kinetic energy under a general coordinate transformation. If the equipartitioning holds for $\{p_i\}$ and $\{P_i\}$ under the ergodicity condition, the temperature of the system is proportional to the time-averaged kinetic energy,

$$\begin{aligned} \langle K \rangle &= \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^{\tau} K(t)\, dt \\ &= \frac{1}{2m} \sum_{i=1}^{3N} \langle p_i^2 \rangle \\ &= \frac{1}{2m} \sum_{i=1}^{3N} \langle P_i^2 \rangle \\ &= \frac{3}{2} N k_{\mathrm{B}} T. \end{aligned} \tag{3}$$

We substitute the multiresolution decomposition Eq. (2) into Eq. (3) and make use of the orthogonality relation of the wavelet basis. Then, the averaged kinetic energy can be decomposed as

$$\langle K \rangle = \frac{1}{2m} \sum_{i=1}^{3N} \left(\sum_{j=1}^{n} \langle (d_i^{(j)})^2 \rangle + \langle a_i^2 \rangle \right).$$

This expression provides a resolution of the temperature of the system over the levels of time scales $\{2^j \Delta t\}$ for each collective momentum P_i,

$$T = \sum_{i=1}^{3N} T_i = \sum_{i=1}^{3N} \left(\sum_{j=1}^{n} T_i^{(j)} + T_i^a \right), \tag{4}$$

where

$$T_i^{(j)} = \langle (d_i^{(j)})^2 \rangle / (3N k_{\mathrm{B}} m),$$

and

$$T_i^a = \langle a_i^2 \rangle / (3N k_{\mathrm{B}} m).$$

Hereafter, these temperatures, $T_i^{(j)}$, and T_i^a are referred to as *kinetic temperatures* over scales.

Li and co-workers[45] previously applied this decomposition of kinetic temperature to the time series of the total kinetic energy of solid and liquid argon. They found the shift of the peak maximum in the kinetic temperatures toward fast

time scales as temperature was increased, resulting from both the higher mean particle velocities, leading to a shorter time between collisions, and from the steeper portion of the repulsive potential. On the other hand, in this article, we further decomposed the temperature into contributions from collective momenta conjugate to the PCs. This decomposition is a key step to characterize different space scale dynamics of proteins. In fact, it will be shown that collective momenta have different temperature "spectra" depending on their space scales in the latter analysis.

C. Gaussianity of collective momentum distribution

In order to evaluate the time scale of the equipartitioning of vibrational energy, or the redistribution of the energies between different scales of the collective motion, we measure the time-dependence of the Gaussianity of the momentum P_i by using the third order central moment of P_i,

$$\gamma_i = \langle (P_i - \langle P_i \rangle)^3 \rangle,$$

where $\langle \cdot \rangle$ is the expectation. This measure, called the (unnormalized) skewness, evaluates the asymmetry of the shape of the distribution function. Negative values for the skewness indicate that the distribution has a long tail in the left (relative to the right) and positive values for the skewness indicate the distribution has a long tail in the right (relative to the left). For a Gaussian distribution, the skewness becomes zero since the Gaussian has a completely symmetric shape.

In general, the estimation of the skewness requires large numbers of samples and should be carefully evaluated for samples with limited length. Thus, instead of determining the absolute value of the skewness, we diagnose the convergence of the estimates of the skewness as a function of time. By evaluating the time scale of the convergence, we estimate the time scale for which the distribution of P_i relaxes to a Gaussian distribution.

If the estimation of the skewness, $\hat{\gamma}_i$, is calculated over the subset of the trajectory with step length n, which is sufficiently shorter than the total step length N_T, one can make a number of estimations $\{\hat{\gamma}_i(n)\}$ at different regions of the time window with length $n\Delta t$ (taken not to overlap with each other). If the snapshots of trajectories in different time windows are uncorrelated, the variance of the estimation $\hat{\gamma}_i(n)$ is expected to decrease monotonically, as n increases. Specifically, the following scaling relation holds:

$$\mathrm{Var}(\hat{\gamma}_i(n)) = \frac{\mu_6}{n}, \tag{5}$$

where μ_6 is the sixth order moment of P_i.

On the other hand, if the snapshots are statistically correlated during a certain time step τ_{dep}, the above relation is expected to be deformed. A simple stochastic process with low-order correlations yields (see Appendix B for details),

$$\mathrm{Var}(\hat{\gamma}_i(n)) = \frac{\mu_6 \tau_{\mathrm{dep}}(\tau, n)}{n}, \tag{6}$$

where τ is the decay time of P_i. τ_{dep} is regarded as the typical time scale for which the distribution of P_i relaxes to a Gaussian distribution. Existence of multiple decay times and higher-order correlations may also deform the simple scaling relation (Eq. (5)) in a similar manner to Eq. (6) but with more complicated dependence of τ_{dep} on the decay times. We here focus on the time scale of τ_{dep} in terms of Gaussianity and alleviate the explicit calculation of the decay times and higher order correlations.

III. MODEL

We demonstrate the method by applying it to a frequently studied variant of a 3-color, 46-bead protein model[46] whose potential and free energy landscapes, kinetics, and dynamics have been studied extensively.[30–34] The model is composed of hydrophobic (B), hydrophilic (L), and neutral (N) beads. The potential energy function of the model (called the BLN model) is described by

$$\begin{aligned} V = {} & \frac{1}{2}\sum_{i}^{\mathrm{bond}} K_R(R_{i,i+1} - R_0)^2 + \frac{1}{2}\sum_{i}^{\mathrm{angle}} K_\theta(\theta_i - \theta_0)^2 \\ & + \epsilon \sum_{i}^{\mathrm{dihedral}} [A_i(1 + \cos\phi_i) + B_i(1 + \cos 3\phi_i)] \\ & + 4\epsilon \sum_{ij}^{\mathrm{nonbonded}} C_{ij}\left[\left(\frac{\sigma}{R_{ij}}\right)^{12} - D_{ij}\left(\frac{\sigma}{R_{ij}}\right)^{6}\right]. \end{aligned}$$

The details of the parameters are given in Refs. 30–33. Throughout this study, the units of energy, temperature, bead mass, and time are ϵ, ϵ/k_{B}, m, and $t^* = \sigma\sqrt{m/\epsilon}$, unless otherwise noted explicitly. For small single-domain proteins, an ideal funnel-type energy landscape[47] has been postulated as one of the most fundamental properties, which manifests a two-state like transition. In this article, in order to shed light on the nature of coarse-grained dynamics of folding inherent to small proteins, we impose the Gō-type bias[31,48,49] toward the global minimum structure of the BLN model and apply our analyses to this biased variant of the BLN model. For the constant temperature MD simulation, Berendsen's thermostat was used[50] with an integration time $0.0025t^*$ in which the system is coupled to a heat bath with a coupling time of $0.50t^*$. In the scheme of Berendsen's thermostat, the system is weakly coupled to a heat-bath, and thus it is suitable to investigate the dynamical properties of the original Hamiltonian system under an isothermal condition compared with other thermostats.[51] In most studies of low-frequency, collective motions in proteins, it is assumed that the dynamics is overdamped by referring to the viscosity of real water. In contrast, our choice of the coupling time of $0.50t^*$ may underdamp the low-frequency motions compared to those in real water although the damping time of velocity cannot be directly compared with the coupling time since Berendsen's thermostat rescales not velocities but the total kinetic energy. The viscosity dependence of the folding rates for the BLN models was first investigated by Klimov and Thirumalai.[52] In a later study by Best and Hummer,[53] the origin of viscosity dependence of the (macroscopic) folding rate was explained by the "internal friction" arising from nonlinear couplings of intra-protein interactions and non-Markovian effects in microscopic tran-

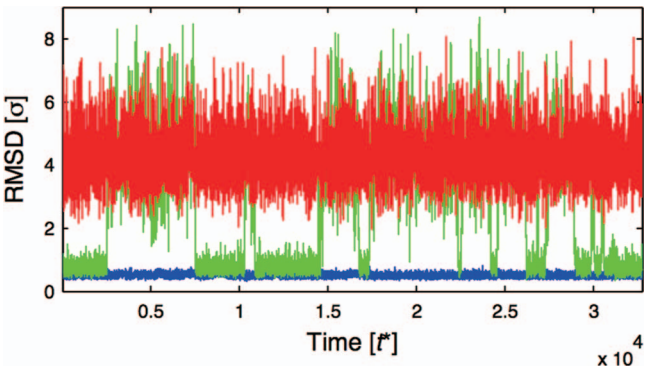

FIG. 2. Time courses of root mean square deviations with respect to the global minimum structure at $T = 0.2$ (blue), 0.6 (green), and 2.0 (red).

sitions on the rough energy landscape. Thus, even when the viscosity is so large that the dynamics looks "overdamped" at macroscopic scale, energy flow pathways and the time scale of the equipartitioning of energy at microscopic scale could affect the kinetics at macroscopic scale, such as the folding rates. In this regard, developing a method to identify the hierarchy of energy flows should be of crucial importance in bridging microscopic dynamics and macroscopic behavior in complex environments.

The model exhibits a sharp two-state like transition around the folding temperature $T_f \sim 0.60$. Simulations were performed over a wide range of temperatures that is below the folding temperature T_f ($T = 0.2$), around T_f ($T = 0.6$), and above T_f ($T = 2.0$). In Fig. 2, time courses of the root mean square deviations (RMSD) with respect to the global minimum structure at $T = 0.2$, 0.6, and 2.0 are shown. Trajectories were recorded every 100 integration time steps, i.e., $\Delta t = 0.0025t^* \times 100 = 0.25t^*$, and $2^{17}(= N_T)$-step snapshots were analyzed at each temperature (thus, the full trajectory length is $2^{17}\Delta t = 2^{17} \times 0.25t^* = 32\,768t^*$).

The PCA was performed for the Cartesian coordinates of the beads at each temperature. Before conducting the PCA, a fitting preprocess was applied for each trajectory in order to eliminate the global translations and rotations of the molecule. In the process, the least-squares fitting was recursively applied with respect to the average structure until the convergence of the average structure was achieved. It was confirmed that the last (smallest) six eigenvalues of the variance-covariance matrix were almost zero, indicating that the translations and rotations were properly eliminated. Also, the robustness of the eigenvectors was checked by comparing the eigenvectors calculated using the first and last half subsets of the full trajectory (given in Fig. S1 in the supplementary material[54]). The convergence of the eigenvectors was considered to be reasonable, suggesting that the PC modes reflect the (true) ensemble of our system.

It was found that the 1st PC (Q_1) makes a large contribution to the total variance of the system (30.9% for $T = 0.2$, 48.2% for $T = 0.6$, and 26.7% for $T = 2.0$) compared to the others (e.g., the 2nd PC (Q_2) has 11.9% for $T = 0.2$, 12.6% for $T = 0.6$, 11.2% for $T = 2.0$). The contributions rapidly decrease as a function of the PC index, and the contribution of the 50th PC (Q_{50}) is, for example, 0.2% for $T = 0.2$, 0.02% for $T = 0.6$, and 0.06% for $T = 2.0$ (given in Fig. S2 in the supplementary material[54]). After the PCA, the mode structures of the PCs were visualized on the average structure for each temperature (given in Figs. S3 and S4 in the supplementary material[54]). The visualization revealed that the 1st PC modes of $T = 0.6$ and $T = 2.0$ have collective opening/closing motions involving the terminal strands. Since the BLN model consists of four strands, and three flexible loop regions connecting them, the rigid-body like motions of the strands are reasonable considering the energetics. Also, the 1st PC mode at $T = 0.2$ is related to the fluctuations of the loop and terminal regions. On the other hand, the structures of the 50th PC mode show anti-correlated motions between adjacent beads in the sequence, involving local deformations of the rigid strands irrespective of temperatures.

IV. RESULTS

A. Multiresolution decomposition

Figure 3 shows the time propagation of the details, $d_1^{(j)}(t)$, and the approximation, $a_1(t)$, of $P_1(t)$ of the largest amplitude collective motion (the 1st PC or $Q_1(t)$) at three different temperatures, that is, below the folding temperature T_f ($T = 0.2$), around T_f ($T = 0.6$), and above T_f ($T = 2.0$). Here, the multiresolution decomposition was carried out up to 10 levels corresponding to time scales from $t = 2^1\Delta t$ to $t = 2^{10}\Delta t$, where Δt is the sampling time. Note that the time scale of the first level, $t \approx 2^1\Delta t$, corresponds to that of one oscillation of bond stretching between two beads. At $T = 0.2$ the fluctuations of $P_1(t)$ mostly originate from the details of $d_1^j(t)$, $j = 4, 5$, reflecting that the system is trapped within a potential basin in the folded state. In contrast, at $T = 2.0$, the details with slow time scales ($d_1^{(j)}(t)$, $j = 5, \ldots, 9$) become dominant due to the fact that the system explores a much wider region on the potential energy landscape requiring a longer time scale for the excursion than that in the folded state.

In turn, at $T = 0.6 \sim T_f$, the fluctuations of the details show mutual dependence across scales along the course of time evolution. Comparing Fig. 3 with the RMSD plot (Fig. 2) shows how the amplitudes of the fluctuation of the details reveal precisely when the protein exhibits a folding or unfolding transition. From the figure, it is clear that the vibrational energy redistribution occurs between the fast and slow time scales. For instance, when the unfolding transition takes place, the fluctuations at the fast time scales ($d_1^{(j)}(t)$, $j = 1, \ldots, 4$), i.e., vibrational energy with high-frequencies, are transferred to the slow time scales ($d_1^{(j)}(t)$, $j = 6, \ldots, 8$), i.e., vibrational energy with low-frequencies. At the folding transition, the inverse energy transfer can be observed from the low to high frequency components in the collective coordinate. The detailed correlations between the RMSD and slow and fast time scales were investigated in Fig. S10 in the supplementary material.[54] While it was found that a fast time scale ($d_1^{(3)}(t)$) responds simultaneously to the structural transitions, we could not determine the exact timings of the transitions for a slow time scale ($d_1^{(8)}(t)$) because of the nature of the multiresolution decomposition (as there is lower time resolution for slower components).

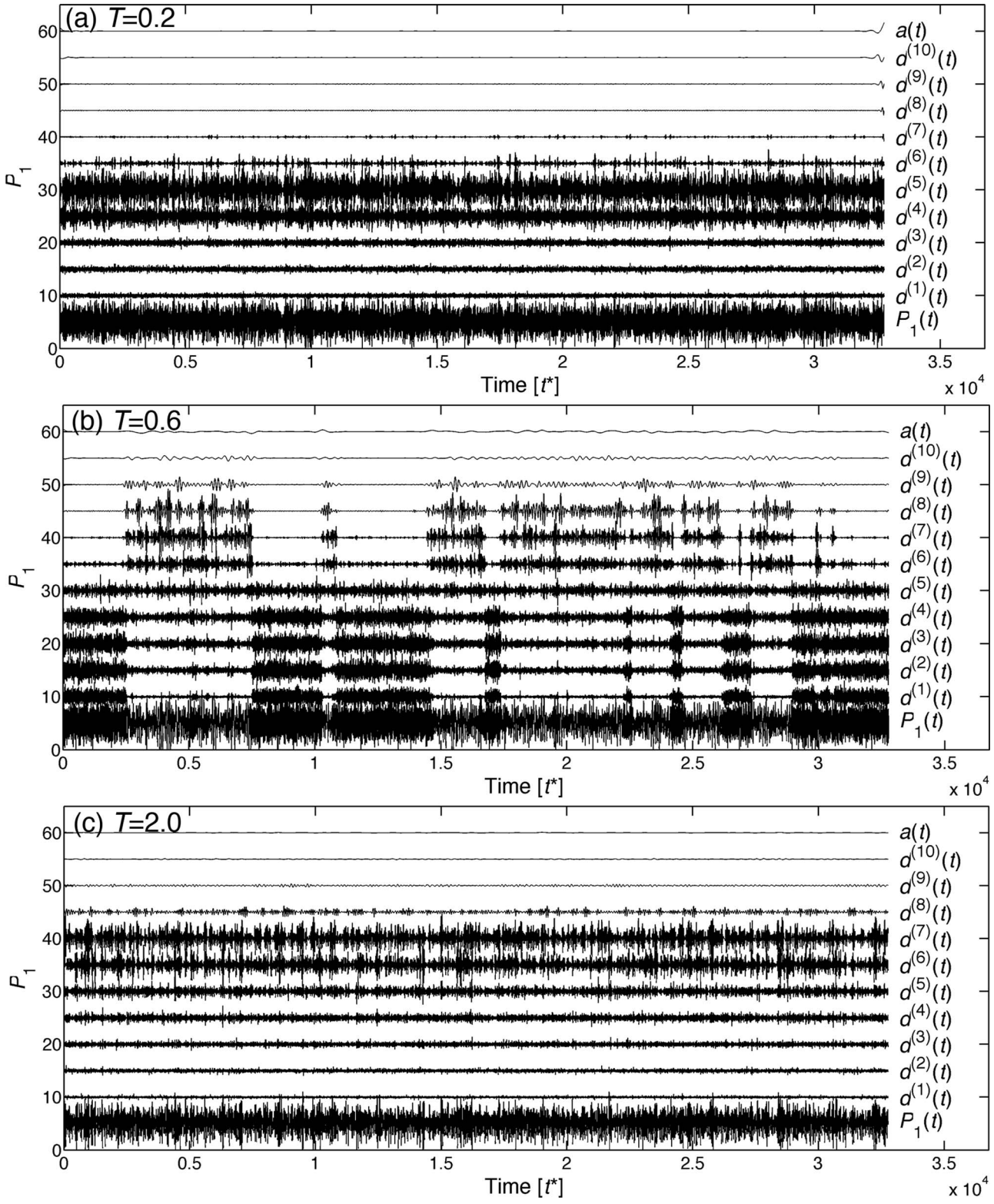

FIG. 3. Multiresolution decomposition of P_1 fluctuations, lowest line at $T = 0.2$ (a), 0.6 (b), and 2.0 (c). Higher lines show fluctuations in successively lower frequency bands, from lowest, d^{10} at top, to highest frequencies, d^1, next to bottom. At $T = 0.6$, a number of energy transfers between "fast" and "slow" scales coupled with folding-unfolding transitions are observed.

One notable characteristic, which may be a property of this specific model, is the sharpness of the transition between folded, fast-motion states and unfolded, slow-motion states. Whether this behavior appears in simulations of real proteins remains to be explored in future work. However, it is noted this "wavelet-PC" analysis can reveal the transition of vibrational energy between slow and fast modes, if it exists.

The fluctuations of approximation $a_1(t)$ are relatively small compared to other contributions to the time dependence irrespective of temperature. The same tendency was also

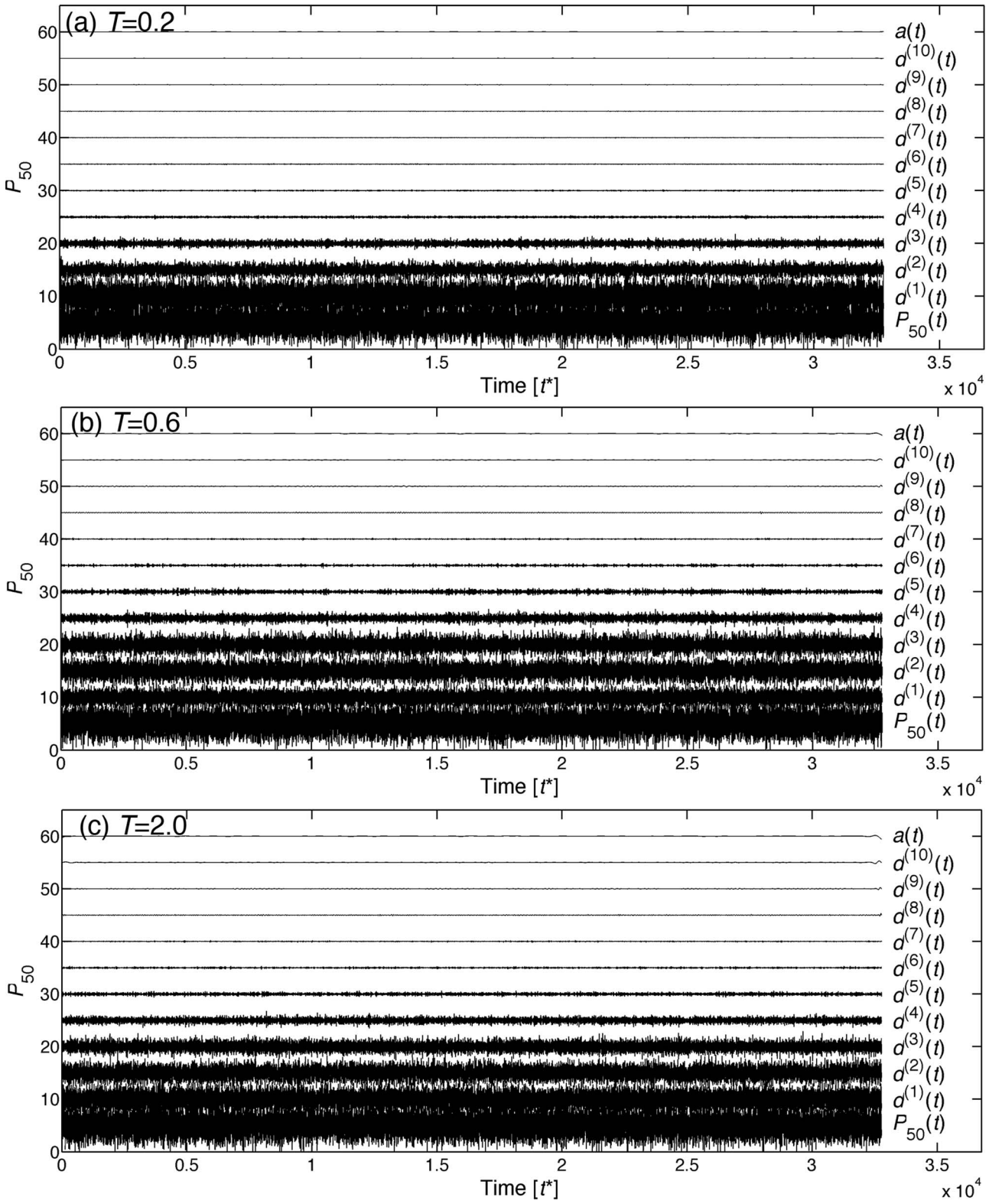

FIG. 4. Multiresolution decomposition of P_{50} at $T = 0.2$ (a), 0.6 (b), and 2.0 (c). Ordering is that of Fig. 3.

observed for the other collective momenta. As noted in Sec. II A, the approximation $a_i(t)$ includes all contributions to the time dependence of the collective momentum on levels higher (slower) than the truncation level 10. Thus, the small amplitudes of $a_i(t)$ indicate that (1) the multiresolution decomposition up to 10 levels captures the essence of a hierarchy of the collective dynamics along the chosen, collective momenta and (2) the contribution of $a_i(t)$ to aperiodic diffusive motion is negligible in our system.

Figure 4 shows the details and the approximation of the 50th largest amplitude collective motion, P_{50} (conjugate to the 50th PC, Q_{50}), at three different temperatures. In contrast to

the results of the largest amplitude motion, P_1 (Fig. 3), the amplitudes of the fluctuations of the details have large values around the fast time scales ($d_{50}^{(j)}(t)$, $j = 1, \ldots, 4$) irrespective of temperature, and the redistribution of the vibrational energy cannot be seen during the time course of evolution. From this observation, it is suggested that the harmonic approximation of the potential energy surface holds well in the subspace spanned by small amplitude PC modes. These harmonic modes are invariant during the folding/unfolding transition implying that these modes do not participate in the transition compared to the anharmonic modes, such as the largest amplitude motion, P_1.

In order to obtain further insights in terms of structure, the motions of the 1st PC (Q_1) in the folded and unfolded states were investigated by inverting typical structures from the average structure and the 1st PC (given in Fig. S5 in the supplementary material[54]) following the protocol of Ref. 55. It was found that, in the folded structure, the 1st PC mode "feels" the strong hydrophobic interactions between the strands while the unfolded state shows large opening/closing motions of the strands largely due to the hinge motions of the three flexible loop regions. On the other hand, the structures inverted from the 50th PC (Q_{50}) show little difference in the folded and unfolded states, as both states are found to be local deformations around the average structure (given in Fig. S6 in the supplementary material[54]).

B. Kinetic temperatures over scales

Then, how do the kinetic temperatures, $T_i^{(j)}$ in Eq. (4), associated with Q_i distribute over scales? As the contributions of details $d_i^{(j)}$ to the fluctuations of collective momentum P_i, we show the kinetic temperatures for P_1 and P_{50} over different scales at these three temperatures in Fig. 5. As one can infer from Fig. 3, for P_1, the scale corresponding to the maximum peak tends to shift toward slower scales (larger j in $d_i^{(j)}$) as the temperature increases although at $T = 0.6$ the kinetic temperature of P_1 is divided into two peaks around $j = 3$, and 8. On the other hand, for P_{50}, the contribution

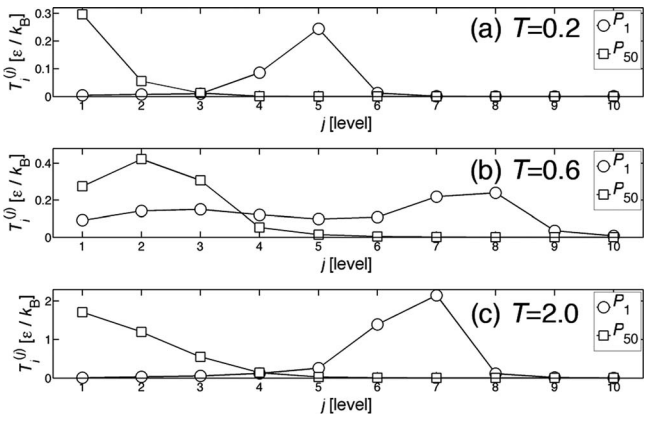

FIG. 5. Kinetic temperatures for P_1 (circles) and P_{50} (rectangles) at $T = 0.2$ (a), 0.6 (b), and 2.0 (c). The abscissa denotes the level j which has the time scale of $t = 2^j \Delta t$, where Δt is the sampling time. As the level increases, the corresponding time scale changes from the highest frequency (corresponding to the bond stretching motion) to low frequencies.

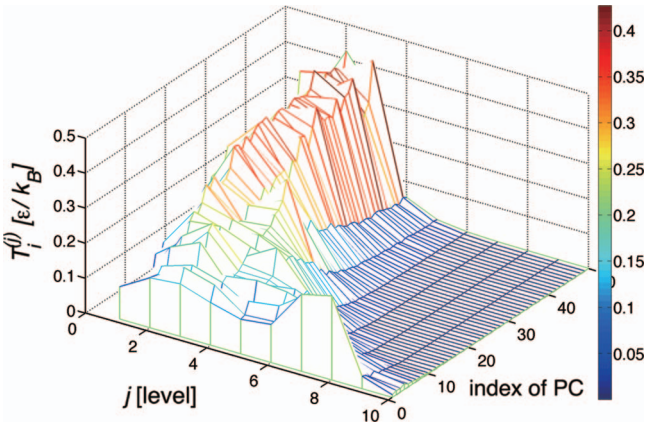

FIG. 6. Three-dimensional shaded surface whose z-components are the kinetic temperatures $T_i^{(j)}$ at $T = 0.6$, on the plane defined by the level j of the multiresolution decomposition and the index i of PC. The surface is colored according to the kinetic temperature from blue to red.

from the different scales always falls into a set of fast components around $j = 1$ and 2. Compared with $P_1(t)$ there exists neither temperature-dependence nor folding/unfolding event-dependence at the folding temperature for P_{50}. Again, this suggests that the energy redistribution due to the anharmonic dynamics during the folding/unfolding transition occurs in the relatively small subspace spanned by large-amplitude collective "modes."

The size of this subspace was investigated by calculating the kinetic temperatures from the 1st to 50th collective momenta at $T = 0.6$. Figure 6 plots a three-dimensional shaded surface whose z-components are the kinetic temperatures ($T_i^{(j)}$ in Eq. (4)) at $T = 0.6$, on the plane defined by our "wavelet-PCA" decomposition (the level j and the index i of PC). In the figure, the double peaks at fast and slow time scales at around $j = 3$, 7, and 8 are observed up to the 15th collective momentum, suggesting that the energy redistribution occurs in this 15-dimensional subspace, much smaller than the full dimensional space of the model (the number of full dimension is $46 \times 3 = 138$). To our knowledge, Fig. 6 provides us with a new insight about protein dynamics; the figure clearly reveals a set of coupled degrees of freedom in the dynamics, which is hard to detect with "static" analyses, such as the free energy landscape on the coordinate space. The figure suggests that the dynamical properties of the protein folding can be reduced to nonlinearly coupled vibrational oscillators in this small subspace, and the rest of degrees of freedom are regarded as thermal noise.

Considering the amplitudes of fluctuations in the coordinate space, the observed double peaks up to the 15th collective momentum at $T = 0.6$ are reasonable since the large-amplitude 15 PCs (conjugate to the collective momenta) represent 97% cumulative contributions to the total variance in the fluctuations of the coordinates. Thus, our finding implies that one could reduce both the ensemble and the dynamics of the folding/unfolding transitions into this relatively small subspace. As shown in the previous figure (Fig. 5), no double peaks were observed for low-indexed PCs at the other temperatures ($T = 0.2$ and 2.0) in the thee-dimensional plots (given in the supplementary material[54]).

C. Relaxation to a Gaussian distribution of collective momenta

How fast is the vibrational energy redistributed in each collective momentum? Does the time scale of the relaxation to a Gaussian distribution depend on the kinds of PCs or the states where the system resides? In Fig. 7, we present the variances of $\hat{\gamma}_i(n)$ as a function of the logarithm of the time window step n in the folded state (corresponding to the time region [11 250, 13 750] in Fig. 2), and the unfolded state (corresponding to the time region [5000, 7500] in Fig. 2), respectively. In the folded state (Fig. 7(a)), $\mathrm{Var}(\hat{\gamma}_i(n))$ tends to follow the scaling relation of Eq. (5) indicating that τ_{dep} is small compared with the observed time scales, that is, the dynamics is regarded as independent stochastic processes whose distribution is close to the Gaussian distribution. This tendency holds for all the P_i irrespective of temperatures in the folded state.

On the other hand, in the unfolded state (Fig. 7(b)), the scaling relation starts to deviate for the momenta conjugate to the large-amplitude coordinates and they do not appear to follow the relation even around the time scale of $t = 2^{10}\Delta t$ ($2^9 = 512$ times longer than the time scale of one oscillation of bond stretching between two beads $t \approx 2^1\Delta t$). This means that the dynamics of these momenta preserve memories at the observed time scales, and more importantly, they do not simply follow the Gaussian distribution based on the central limit theorem. Interestingly, the deviations from the scaling relation are pronounced at $T = 0.6$ rather than those of $T = 2.0$. Non-Gaussianity seems to be strong at the transition temperature. These imply that the time scale of vibrational energy relaxation for the process of collective motions with large amplitudes is relatively slower in the unfolded state than in the folded state, although small amplitude motions are simply regarded as spectator modes of "thermal noise" in the protein system irrespective of which state the system visits.

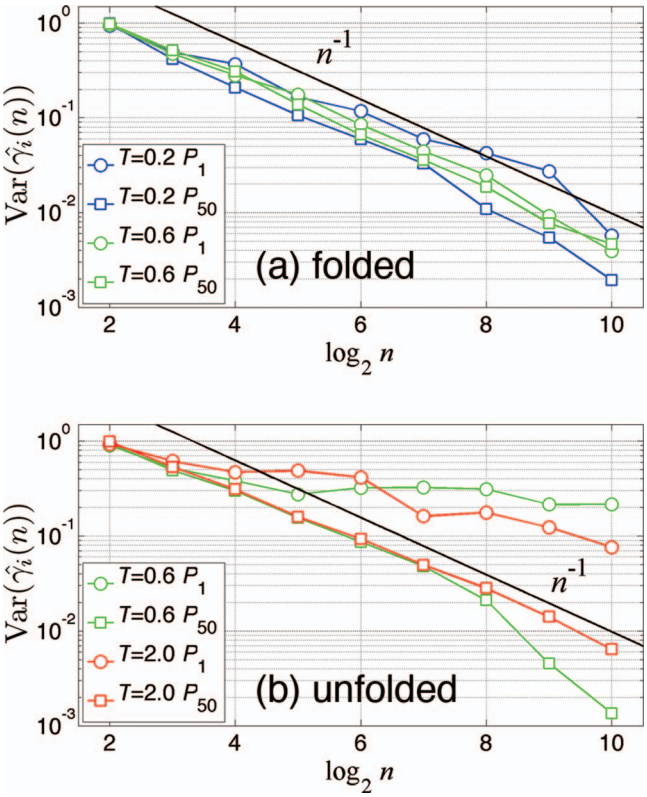

FIG. 7. The variances of the skewness estimation $\hat{\gamma}_i(n)$ as a function of the logarithm of the time window step n in the folded state (a), and the unfolded state (b). The logarithm of n, $\log_2 n$, is equivalent to the level j of multiresolution decomposition, both of which have the time scale of $t = 2^j\Delta t$. The values of P_1 (circles) and P_{50} (rectangles) are plotted at $T = 0.2$ (blue), 0.6 (green), and 2.0 (red).

V. DISCUSSION AND CONCLUSION

In this article, we have applied the multiresolution decomposition using wavelets to a small coarse-grained protein model. The momenta conjugate to the collective coordinates captured by the PCA and the kinetic energy under general coordinate transformation were formulated, decomposing various space scale motions into contributions from large-amplitude and small-amplitude PC "modes."

Unlike the conventional Fourier-based approaches, the wavelet-based approach provides the time-localized decomposition of the collective momenta, which reveals the vibrational energy transfers depending on the folding/unfolding transitions of the system. The contributions or details from a hierarchy of time scales are further characterized by the effective kinetic temperatures of the system in analogy with conventional Fourier spectra, and the Gaussianity measure based on the skewness of the shape of the distribution function. From these analyses, we have made the following observations regarding the conformational transition of a model protein in the low viscosity regime in condensed phase: (i) The vibrational energy redistribution due to the anharmonic nature of the conformational transition occurs in a relatively small subspace spanned by the large-amplitude collective coordinates. (ii) In the subspace, the vibrational energies are transferred between the slow and fast "modes" and these transfers are well correlated with the events of folding/unfolding transitions. (iii) The time scales of the energy redistribution along the large-amplitude collective coordinates are slow compared to the fluctuations of the other small-amplitude coordinates. (iv) Thus, for the reactive large-amplitude collective coordinates, the effect of the coupling with the non-reactive small-amplitude coordinates can be regarded as thermal noise.

In previous studies, the anharmonic nature of protein dynamics has been investigated by various simulation studies, most of which have mainly focused on the anharmonic aspects of the potential or free energy landscapes along the large-amplitude collective coordinates. The characterization of the energy landscape tells us the "static" property of the system, such as the conformational distribution under the equilibrium ensemble. In this study, we have combined wavelet and principal component analyses to give a new perspective that identifies functionally important vibrational states and elucidates the exchange of energy between slow and fast vibrational modes, as it occurs in the temporal behavior of a protein model. This was accomplished by applying our methods to the time series of the momenta instead of the atomic coordinates. The momentum space is not only suitable to characterize the "dynamic" aspects of the anharmonic

system, such as vibrational energy transfer, but is also appropriate to detect the transitions between a number of minima on the energy landscape as demonstrated in the correlations between the energy transfers and the conformational transitions by this study.

It is noted that the Gaussianity measure based on the skewness of the shape of the distribution function evaluates the asymmetry of the shape. Thus, for example, this measure lacks to differentiate Lorentzians (symmetric shapes but longer tails than Gaussians) from Gaussians. In order to explore symmetric but non-Gaussian shapes of the distributions, an application of the fourth moment (kurtosis), which primarily measures the heavy tails of the distribution, will be useful in future work.

The methods applied here can be further developed in several directions. One is to introduce other multivariate statistical methods to interpret the multivariate data sets instead of the PCA. For example, the relaxation mode analysis[56,57] or the time-structure based independent component analysis (tICA)[58] would be superior to the PCA in the sense that it can capture the slow hidden collective variables which are often unrecognized in analyses based on PCA. In particular, the wavelet analysis of coordinates (instead of momenta) extracted from these methods would be suitable for the investigation of slow *diffusive* motions since the coordinates can be regarded as a low-pass filtered process of velocity in terms of signal processing. A second development would be to extract time-localized coordinates (or their conjugate momenta) from the pattern of the multiresolution decomposition of the original Cartesian coordinates (or momenta). This can be done with singular value decomposition of the wavelet coefficient matrix,[39] in analogy with the conventional spectral density matrix.[59] Using this approach, Kamada and co-workers revealed that the clusters of the collective motions of *Thermomyces lanuginosa* lipase drastically change during the time evolution, involving single or multiple secondary structures.[39] Still one more future direction would be using our approach to identify the degree of sharpness of temporal changes in vibrational energy distributions associated with formation of key contacts in a protein's history. Coupling the power of wavelets with the capability of a clustering algorithm in the time-frequency domain would provide a new pathway to deeper understanding of the dynamics of a complex system.

ACKNOWLEDGMENTS

The authors thank D. M. Leitner for helpful comments and discussions, and wish to thank the Telluride Science Research Center for its hospitality, enabling authors to work together to complete this work. T.K. acknowledges financial support from JSPS, JST/CREST, Grant-in-Aid for Research on Priority Areas "Systems Genomics," "Molecular Theory for Real Systems," MEXT. M.T. is supported by a Grant-in-Aid for Research on Priority Area "Real Molecular Theory" from MEXT, by a Grant-in-Aid for Scientific Research (Grant No. 21540413) and for Challenging Exploratory Research (Grant No. 22654047) from JSPS, and Nara Women's University Intramural Grant for Project Research. Y.M. is supported by the Special Postdoctoral Researcher Program of RIKEN, and a Grant-in-Aid for Young Scientists (B) (Grant No. 24770159) from JSPS.

APPENDIX A: KINETIC ENERGY UNDER GENERAL COORDINATE TRANSFORMATION

1. Kinetic energy in general coordinates

In the following, the Cartesian coordinates of the jth particle are denoted as $(q_{3(j-1)+1}, q_{3(j-1)+2}, q_{3j})$. The mass of the jth particle is denoted by m_{3j} $(j = 1, \ldots, N)$, and we use here the convention of $m_{3(j-1)+1} = m_{3(j-1)+2} = m_{3j}$. Then, the kinetic energy of the total system is given by

$$K = \frac{1}{2}\sum_{i=1}^{3N} m_i \dot{q}_i^2.$$

Suppose that a general transformation is given to another coordinate system Q_l $(l = 1, \ldots, M)$

$$q_i = f_i(Q_1, Q_2, \ldots, Q_M) \quad (i = 1, \ldots, 3N).$$

Here, we include a possibility of holonomic constraints, i.e., $M < 3N$. After differentiating q_i by t as

$$\dot{q}_i = \sum_{l=1}^{M} \frac{\partial f_i}{\partial Q_l}(Q_1, Q_2, \ldots, Q_M)\dot{Q}_l \quad (i = 1, \ldots, 3N),$$

the kinetic energy is represented by

$$\begin{aligned} K &= \frac{1}{2}\sum_{i=1}^{3N} m_i \left(\sum_{l=1}^{M} \frac{\partial f_i}{\partial Q_l}\dot{Q}_l\right)\left(\sum_{k=1}^{M} \frac{\partial f_i}{\partial Q_k}\dot{Q}_k\right) \\ &= \frac{1}{2}\sum_{l,k=1}^{M} \dot{Q}_l \dot{Q}_k \sum_{i=1}^{3N} m_i \frac{\partial f_i}{\partial Q_l}\frac{\partial f_i}{\partial Q_k} \\ &= \frac{1}{2}\sum_{l,k=1}^{M} m_{l,k}\dot{Q}_l \dot{Q}_k, \end{aligned}$$

where we define

$$m_{l,k} \equiv \sum_{i=1}^{3N} m_i \frac{\partial f_i}{\partial Q_l}\frac{\partial f_i}{\partial Q_k}.$$

If the masses are not uniform, we can define a mass-scaled orthogonal transformation $\{q_i\} \to \{Q_l\}$ so that

$$\begin{aligned} m_{l,k} &= \sum_{i=1}^{3N} m_i \frac{\partial f_i}{\partial Q_l}\frac{\partial f_i}{\partial Q_k} \\ &= m'_l \delta_{l,k}, \end{aligned}$$

where the matrix $\{m_{l,k}\}$ becomes diagonal, and the kinetic energy is

$$K = \frac{1}{2}\sum_{k=1}^{M} m'_k \dot{Q}_k^2.$$

In particular, if the masses of the particles are the same, i.e., $m \equiv m_i$, and the transformation is orthogonal,

$$\sum_{i=1}^{3N} \frac{\partial f_i}{\partial Q_l} \frac{\partial f_i}{\partial Q_k} = \delta_{l,k},$$

then, the kinetic energy becomes simply

$$K = \frac{m}{2} \sum_{k=1}^{M} \dot{Q}_k^2.$$

It is found that, in the cases of linear transformation, as the PCs used in this study, the orthogonality (mass-scaled orthogonality) is required for the coefficients of the transformation. However, for nonlinear transformation, these conditions should hold for any values of the generalized coordinates.

2. Introduction of momenta in general coordinates

The Lagrangian in general coordinates Q_k $(k = 1, \ldots, M)$ is given by

$$L\left(Q_1, \ldots, Q_M, \dot{Q}_1, \ldots, \dot{Q}_M\right) = \frac{1}{2} \sum_{l,k=1}^{M} m_{l,k} \dot{Q}_l \dot{Q}_k - V(Q_1, \ldots, Q_M),$$

where $V(Q_1, \ldots, Q_M)$ is the potential. The momenta P_k $(k = 1, \ldots, M)$ are defined by $P_k \equiv \frac{\partial L}{\partial \dot{Q}_k}$

$$P_k = \frac{\partial}{\partial \dot{Q}_k} \left(\frac{1}{2} \sum_{l,k'=1}^{M} m_{l,k'} \dot{Q}_l \dot{Q}_{k'} \right) = \sum_{l=1}^{M} m_{l,k} \dot{Q}_l = \sum_{l=1}^{M} m_{k,l} \dot{Q}_l.$$

Denoting the inverse matrix of $\{m_{j,k}\}_{j,k=1,\ldots,M}$ as $\{\tilde{m}_{j,k}\}_{j,k=1,\ldots,M}$, i.e., $\sum_{k=1}^{M} m_{j,k} \tilde{m}_{k,l} = \delta_{j,l}$, we have

$$\dot{Q}_j = \sum_{k=1}^{M} \tilde{m}_{j,k} P_k.$$

Then, the kinetic energy is expressed as

$$K = \frac{1}{2} \sum_{l,k=1}^{M} m_{l,k} \left(\sum_{j=1}^{M} \tilde{m}_{l,j} P_j \right) \left(\sum_{j'=1}^{M} \tilde{m}_{k,j'} P_{j'} \right) = \frac{1}{2} \sum_{l,j=1}^{M} \tilde{m}_{l,j} P_j \sum_{j'} P_{j'} \sum_{k=1}^{M} m_{l,k} \tilde{m}_{k,j'} = \frac{1}{2} \sum_{l,j=1}^{M} \tilde{m}_{l,j} P_j P_l.$$

When the matrix $\{m_{l,k}\}_{l,k=1,\ldots,M}$ is diagonal (i.e., $m_{l,k} = m'_k \delta_{l,k}$), $\tilde{m}_{l,j} = \frac{\delta_{l,j}}{m'_l}$ holds. The kinetic energy becomes diagonal

$$K = \sum_{j=1}^{M} \frac{{P_j}^2}{2m'_j}.$$

In particular, if the masses are uniform, and the transformation is orthogonal, the kinetic energy becomes

$$K = \frac{1}{2m} \sum_{j=1}^{M} {P_j}^2.$$

APPENDIX B: GAUSSIANITY MEASURE

Skewness is a measure of the asymmetry of the distribution function of a random variable X. Since one of the most significant properties of a Gaussian function is its symmetric shape, the skewness is often used to evaluate whether the given distribution function is Gaussian or not. In the following, we formulate the convergence of the estimates of the skewness as a function of time.

Here, as a measure of the skewness, we use the third central moment μ_3 of the random variable X,

$$\gamma \equiv \langle (X - \mu)^3 \rangle,$$

where $\langle \cdot \rangle$ means the expectation, or the ensemble average, and μ is the expectation of X. Since we use the momentum P_i as X, $\mu = 0$ and $\gamma = \langle X^3 \rangle = 0$. For time series $\{X_j\}$ $(j \in \mathbb{N})$ with length n, an estimator for γ is

$$\hat{\gamma} = \frac{1}{n} \sum_{j=1}^{n} X_j^3.$$

$\hat{\gamma}$ is an unbiased estimator for γ, i.e., $\langle \hat{\gamma} \rangle = \gamma = 0$. To investigate the convergence of $\hat{\gamma}$, we define the variance of the estimates for the skewness,

$$\mathrm{Var}(\hat{\gamma}) = \langle (\hat{\gamma} - \langle \hat{\gamma} \rangle)^2 \rangle = \frac{1}{n^2} \sum_{j=1}^{n} \sum_{k=1}^{n} \left\langle X_j^3 X_k^3 \right\rangle. \tag{B1}$$

Let us assume that the stochastic process $\{X_n\}$ follows a Gaussian process characterized by a covariance matrix $M_{jk} = \langle X_j X_k \rangle$ and $\langle X_j \rangle = 0$ for all j. Then the expectation of $X_j F(X_1, \ldots, X_n)$, where F is an arbitrary function of $\{X_i\}$, is given by[60]

$$\langle X_j F(X_1, \ldots, X_n) \rangle = \sum_{k=1}^{n} M_{jk} \left\langle \frac{\partial}{\partial X_k} F(X_1, \ldots, X_n) \right\rangle.$$

This relation can be used to work out the expectations of the products of Eq. (B1). Applying the above relation to Eq. (B1) yields

$$\mathrm{Var}(\hat{\gamma}) = \frac{9}{n^2} \sum_{j=1}^{n} \sum_{k=1}^{n} M_{jj} M_{kk} M_{jk} + \frac{6}{n^2} \sum_{j=1}^{n} \sum_{k=1}^{n} M_{jk}^3. \tag{B2}$$

When the covariance matrix is diagonal, i.e., $M_{jk} = \sigma^2 \delta_{jk}$, where δ_{jk} is the Kronecker delta, the variance of the estimates

for the skewness is simply given by

$$\mathrm{Var}(\hat{\gamma}) = \frac{\mu_6}{n}, \tag{B3}$$

where μ_6 is the sixth order moment of X ($\mu_6 = 15\sigma^6$).

On the other hand, let us assume that the covariance matrix element M_{jk} is expressed by a single exponential function,

$$M_{jk} = \sigma^2 e^{-|k-j|/\tau},$$

where τ is the decay time of X. Then, the first term of the r.h.s. of Eq. (B2) can be calculated explicitly

$$\begin{aligned}
&\frac{9}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} M_{jj}M_{kk}M_{jk} \\
&= \frac{9\sigma^6}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} e^{-|k-j|/\tau} \\
&= \frac{9\sigma^6}{n^2}\left[n + 2\sum_{j=1}^{n}\sum_{k>j}^{n} e^{-(k-j)/\tau}\right] \\
&= \frac{9\sigma^6}{n^2}\left[n + 2\sum_{l=1}^{n-1}(n-l)e^{-l/\tau}\right] \\
&= \frac{9\sigma^6}{n}\left[\frac{1+e^{-1/\tau}}{1-e^{-1/\tau}} - \frac{2e^{-1/\tau}(1-e^{-n/\tau})}{n(1-e^{-1/\tau})^2}\right].
\end{aligned}$$

Also, the second term is

$$\begin{aligned}
&\frac{6}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} M_{jk}^3 \\
&= \frac{6\sigma^6}{n}\left[\frac{1+e^{-3/\tau}}{1-e^{-3/\tau}} - \frac{2e^{-3/\tau}(1-e^{-3n/\tau})}{n(1-e^{-3/\tau})^2}\right].
\end{aligned}$$

We find the variance of the estimates for the skewness,

$$\mathrm{Var}(\hat{\gamma}) = \frac{\mu_6}{n}\left(\frac{9}{15}A_\tau + \frac{6}{15}A_{\tau/3}\right),$$

where

$$A_\tau = \frac{1+e^{-1/\tau}}{1-e^{-1/\tau}} - \frac{2e^{-1/\tau}(1-e^{-n/\tau})}{n(1-e^{-1/\tau})^2}.$$

Now, we introduce a quantity with a dimension of time, called statistical dependence time τ_{dep},[61] defined by

$$\tau_{\mathrm{dep}}(\tau, n) = \left(\frac{9}{15}A_\tau + \frac{6}{15}A_{\tau/3}\right), \tag{B4}$$

and the variance is written as

$$\mathrm{Var}(\hat{\gamma}) = \frac{\mu_6 \tau_{\mathrm{dep}}(\tau, n)}{n}. \tag{B5}$$

As was discussed in Ref. 61, Eq. (B5) expresses that the statistical degrees of freedom, or the number of independent measurement of the skewness is reduced by the factor $1/\tau_{\mathrm{dep}}$ due to the correlations among the successive measurements. Thus, the time scale of τ_{dep} can be interpreted as the mean interval of successive statistically independent measurements.

It is noted that $\tau_{\mathrm{dep}}(\tau, n)$ depends not only on the decay time τ but also on the length of the time series n. Thus, the variance of the estimates, Eq. (B5), deviates from the simple scaling relation as a function of n (Eq. (B3)) when the value of τ is large compared with the time step of sampling.

In the large n limit, the value of τ_{dep} can be related to the decay time τ. Taylor expansion of Eq. (B4) with regard to $1/\tau$ yields

$$\lim_{n\to\infty} \tau_{\mathrm{dep}} = \tau\left[\frac{22}{15} + \frac{3}{10\tau^2} + O\left(\frac{1}{\tau^4}\right)\right]. \tag{B6}$$

[1] *Proteins: Energy, Heat and Signal Flow*, edited by D. M. Leitner and J. E. Straub (CRC Press, 2009).
[2] D. M. Leitner, Annu. Rev. Phys. Chem. **59**, 233–259 (2008).
[3] P. Kukura, D. W. McCamant, S. Yoon, D. B. Wandschneider, and R. A. Mathies, Science **310**, 1006–1009 (2005).
[4] J. T. Kennis, D. S. Larsen, K. Ohta, M. T. Facciotti, R. M. Glaeser, and G. R. Fleming, J. Phys. Chem. B **106**, 6067–6080 (2002).
[5] T. Ishikura and T. Yamato, Chem. Phys. Lett. **432**, 533–537 (2006).
[6] D. Xu, C. Martin, and K. Schulten, Biophys. J. **70**, 453–460 (1996).
[7] G. S. Engel, T. R. Calhoun, E. L. Read, T.-K. Ahn, T. Mančal, Y.-C. Cheng, R. E. Blankenship, and G. R. Fleming, Nature (London) **446**, 782–786 (2007).
[8] A. Sato and Y. Mizutani, Biochemistry **44**, 14709–14714 (2005).
[9] P. M. Champion, Science **310**, 980–982 (2005).
[10] R. D. Miller, Annu. Rev. Phys. Chem. **42**, 581–614 (1991).
[11] A. Nagy, V. Raicu, and R. Miller, Biochim. Biophys. Acta **1749**, 148–172 (2005).
[12] D. E. Sagnella and J. E. Straub, J. Phys. Chem. B. **105**, 7057–7063 (2001).
[13] L. Bu and J. E. Straub, J. Phys. Chem. B **107**, 10634–10639 (2003).
[14] S. W. Lockless and R. Ranganathan, Science **286**, 295–299 (1999).
[15] N. Ota and D. A. Agard, J. Mol. Biol. **351**, 345–354 (2005).
[16] K. Sharp and J. J. Skinner, Proteins **65**, 347–361 (2006).
[17] K. Moritsugu, O. Miyashita, and A. Kidera, Phys. Rev. Lett. **85**, 3970–3973 (2000).
[18] D. M. Leitner, Phys. Rev. Lett. **87**, 188102 (2001).
[19] H. Fujisaki and J. E. Straub, Proc. Natl. Acad. Sci. U.S.A. **102**, 6726–6731 (2005).
[20] P. Hamm, M. Lim, and R. M. Hochstrasser, J. Phys. Chem. B **102**, 6123–6138 (1998).
[21] A. Xie, L. van der Meer, W. Hoff, and R. H. Austin, Phys. Rev. Lett. **84**, 5435–5438 (2000).
[22] B. J. Berne, M. Borkovec, and J. E. Straub, J. Phys. Chem. **92**, 3711–3725 (1988).
[23] J. N. Onuchic and P. G. Wolynes, J. Phys. Chem. **92**, 6495–6503 (1988).
[24] J. E. Straub and D. Thirumalai, Proc. Natl. Acad. Sci. U.S.A. **90**, 809–813 (1993).
[25] J. E. Straub and J. K. Choi, J. Phys. Chem. **98**, 10978–10987 (1994).
[26] K. Moritsugu and A. Kidera, J. Phys. Chem. B **108**, 3890–3898 (2004).
[27] C. C. Martens, Phys. Rev. A **45**, 6914–6917 (1992).
[28] L. Finney, A. Borrmann, and C. Martens, Chem. Phys. Lett. **214**, 159–165 (1993).
[29] I. Daubechies, *Ten Lectures on Wavelets* (Society for Industrial Mathematics, 1992).
[30] R. S. Berry, N. Elmaci, J. P. Rose, and B. Vekhter, Proc. Natl. Acad. Sci. U.S.A. **94**, 9520–9524 (1997).
[31] M. A. Miller and D. J. Wales, J. Chem. Phys. **111**, 6610–6616 (1999).
[32] G. J. Rylance, R. L. Johnston, Y. Matsunaga, C. B. Li, A. Baba, and T. Komatsuzaki, Proc. Natl. Acad. Sci. U.S.A. **103**, 18551–18555 (2006).
[33] Y. Matsunaga, K. S. Kostov, and T. Komatsuzaki, J. Phys. Chem. A **106**, 10898–10907 (2002).
[34] Y. Matsunaga, C. B. Li, and T. Komatsuzaki, Phys. Rev. Lett. **99**, 238103 (2007).
[35] L. Ye, H. Chen, T. Liu, Z. Wu, J. Li, and R. Zhou, J. Bioinf. Comput. Biol. **03**, 1351–1370 (2005).
[36] L. Ye, Z. Wu, M. Eleftheriou, R. Zhou *et al.*, Biochem. Soc. Trans. **35**, 1551 (2007).
[37] N. C. Benson and V. Daggett, J. Phys. Chem. B **116**, 8722–8731 (2012).
[38] N. C. Benson and V. Daggett, Int. J. Wavelets, Multiresolut. Inf. Process. **10**, 1250040 (2012).
[39] M. Kamada, M. Toda, M. Sekijima, M. Takata, and K. Joe, Chem. Phys. Lett. **502**, 241–247 (2011).

[40]T. Ichiye and M. Karplus, Proteins **11**, 205–217 (1991).
[41]A. E. García, Phys. Rev. Lett. **68**, 2696–2699 (1992).
[42]A. Amadei, A. B. M. Linssen, and H. J. C. Berendsen, Proteins **17**, 412–425 (1993).
[43]A. Kitao, S. Hayward, and N. Go, Proteins **33**, 496–517 (1998).
[44]A. Lichtenberg and M. Lieberman, *Regular and Chaotic Dynamics* (Springer, New York, 1992).
[45]Z. Li, A. Borrmann, and C. C. Martens, Chem. Phys. Lett. **214**, 362–366 (1993).
[46]J. Honeycutt and D. Thirumalai, Biopolymers **32**, 695 (1992).
[47]J. D. Bryngelson, J. N. Onuchic, N. D. Socci, and P. G. Wolynes, Proteins **21**, 167–195 (1995).
[48]H. Nymeyer, A. E. García, and J. N. Onuchic, Proc. Natl. Acad. Sci. U.S.A. **95**, 5921–5928 (1998).
[49]N. Gō, Annu. Rev. Biophys. Bioeng. **12**, 183–210 (1983).
[50]H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, J. Chem. Phys. **81**, 3684–3690 (1984).
[51]J. E. Basconi and M. R. Shirts, J. Chem. Theory Comput. **9**, 2887–2899 (2013).
[52]D. Klimov and D. Thirumalai, Phys. Rev. Lett. **79**, 317–320 (1997).
[53]R. B. Best and G. Hummer, Phys. Rev. Lett. **96**, 228104 (2006).
[54]See supplementary material http://dx.doi.org/10.1063/1.4834415 for supplementary figures.
[55]T. Komatsuzaki, K. Hoshino, Y. Matsunaga, G. J. Rylance, R. L. Johnston, and D. J. Wales, J. Chem. Phys. **122**, 084714 (2005).
[56]H. Takano and S. Miyashita, J. Phys. Soc. Jpn. **64**, 3688–3698 (1995).
[57]A. Mitsutake, H. Iijima, and H. Takano, J. Chem. Phys. **135**, 164102 (2011).
[58]Y. Naritomi and S. Fuchigami, J. Chem. Phys. **134**, 065101 (2011).
[59]Y. Matsunaga, S. Fuchigami, and A. Kidera, J. Chem. Phys. **130**, 124104 (2009).
[60]R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, 2001).
[61]M. Kikuchi and N. Ito, J. Phys. Soc. Jpn. **62**, 3052–3061 (1993).