Nanosecond Motions in Proteins Impose Bounds on the Timescale Distributions of Local Dynamics

Osman Burak Okan, Ali Rana Atilgan, and Canan Atilgan*
Faculty of Engineering and Natural Sciences, Sabanci University, Istanbul, Turkey

ABSTRACT   We elucidate the physics of protein dynamical transition via 10–100-ns molecular dynamics simulations at temperatures spanning 160–300 K. By tracking the energy fluctuations, we show that the protein dynamical transition is marked by a crossover from nonstationary to stationary processes that underlie the dynamics of protein motions. A two-timescale function captures the nonexponential character of backbone structural relaxations. One timescale is attributed to the collective segmental motions and the other to local relaxations. The former is well defined by a single-exponential, nanosecond decay, operative at all temperatures. The latter is described by a set of processes that display a distribution of timescales. Although their average remains on the picosecond timescale, the distribution is markedly contracted at the onset of the transition. It is shown that the collective motions impose bounds on timescales spanned by local dynamical processes. The nonstationary character below the transition implicates the presence of a collection of substates whose interactions are restricted. At these temperatures, a wide distribution of local-motion timescales, extending beyond that of nanoseconds, is observed. At physiological temperatures, local motions are confined to timescales faster than nanoseconds. This relatively narrow window makes possible the appearance of multiple channels for the backbone dynamics to operate.

INTRODUCTION

Protein function is made possible through fluctuations of varying amplitude and duration around the folded conformation of the molecule (1–3). In addition to its instrumental role in imposing the final folded state, the solvent forms a vicinal layer around the protein, mediating the functionality-related molecular fluctuations (4,5). Therefore, a functional protein might well be perceived as an effective protein-solvent complex (6,7). Proteins are characterized as soft materials (1,8) and the temperature window for effective softness is bracketed below by the dynamical transition and above by the unfolding event (2,9–11). The exact value of the dynamical transition temperature depends on environmental factors such as the solvent, the degree of solvation, and the amino acid composition of the protein of interest.

The microscopic interpretation of the dynamical transition is subtler than the unfolding counterpart, the latter being apparent with the overall loss of the folded topology. To interpret the former, many attempts have been made, some concentrating on the solvent around the protein (4,5,12), others on the protein itself (9,13). Furthermore, recent work has demonstrated that the dynamical transition occurs in the absence of protein tertiary or secondary structure, but is related to side-chain interactions with the solvent (14). For a fully hydrated protein, the collisions with the surrounding vicinal layer of solvent molecules excite the protein by transferring energy that in turn is also dissipated by the viscous drag of the solvent (15–17). Recently, it has been suggested that solvent-slaved α relaxations, due to

the strong coupling between surface residues and water, result in a large solvent reorganization energy, facilitating the functionality of the protein (18). The properties of the system below the dynamical transition are attributed to the existence of noncommunicating local substates, which, it is postulated, interconvert at physiological temperatures. For instance, for parvalbumin and using x-ray crystallography, well-defined alternate conformations have been observed within the hydrophobic core at low temperatures, whose population-weighted average corresponds to the single conformation observed at room temperature in the same crystal form (19).

The temporal memory loss in the system depends on the dynamical processes that annihilate local changes. These may be recovered from the correlation function formalism (20), coupled with an appropriate mathematical model describing the underlying physics. Single exponential decays are the natural starting point for analyzing relaxation phenomena, since they are the phenomenological descriptors of simple dashpots, i.e., dampers that display frictional resistance to motion (15). However, the existence of nonexponential decays in molecular systems is now well recognized, supported by experimental and simulation data (21–25). The multiexponential analysis, which extends the single-exponential scope by adding more terms of the same form, is the natural next step (26). Due to the difficulty of assigning a physical meaning to the timescales obtained from a multiexponential analysis, the stretched exponential function is adopted as an alternative (27,28). In this approach, the relaxation profile due to many distributed timescale processes, each contributing to the relaxation phenomena, are represented with a single characteristic timescale and a dimensionless exponent that captures their distribution profile. Such

doi: 10.1016/j.bpj.2009.07.036

behavior is observed when there are multiple channels for the process to decay (29).

In this study, we aimed to answer two major questions: 1), What are the time- and lengthscales of the processes that dynamically operate on the folded protein so as to render it functional above the dynamical transition? and 2), How do processes that operate on different timescales affect each other; in particular, do slower-timescale processes impose limitations on the fast, local motions so as to provide coupled processes and/or multiple decay channels? To answer these questions, we elucidate the physics of protein dynamical transition via 10–100-ns-long MD simulations on hen egg white lysozyme in water at a series of temperatures where the protein remains folded. By monitoring the energy fluctuations in the system, we show that the dynamical transition is marked by a crossover from nonstationary to stationary trajectories. By further monitoring the backbone dynamics, we show that a slower, nanosecond timescale motion in the protein imposes bounds on the distribution of the plethora of fast, local processes averaging in the picosecond range. The implications for controlling protein functionality are discussed.

METHODS

Simulation protocol

The NAMD package is used to model the dynamics of the protein-water systems. The protein is soaked in a solvent box such that there is layer of solvent of at least 5 Å in each direction from any atom of the protein to the edge of the box. The simulated protein-water complex is comprised of 2271 TIP3P water molecules soaking a single hen egg white lysozyme (Protein Data Bank (30) code 193L (31)) molecule, a 129-residue protein. Eight chlorine ions are added for charge neutrality. The system has a total of 10,281 atoms prepared using the VMD 1.8.6 program with solvate plug-in version 1.2 (32). The CharmM27 force field parameters are used (33). Electrostatic interactions are calculated via a particle mesh Ewald method (34). The cutoff distance for nonbonded van der Waals interactions was set to 12 Å with a switching-function cutoff of 10 Å. The RATTLE algorithm was used to fix the bond lengths to their average values.

The system is first subjected to energy minimization until the gradient tolerance is $<10^{-2}$ kcal/mol/Å with the conjugate gradients method. The resulting system with the orthorhombic periodic cell of starting dimensions $48 \times 45 \times 54$ Å is used as the initial structure in all MD simulations. All runs are made in the NPT ensemble at 1 atm with the NAMD software (35). Volumetric fluctuations are preset to be isotropic. Temperature control is carried out by Langevin dynamics with a dampening coefficient of 5/ps, and pressure control is attained by a Langevin piston. The Verlet algorithm is used for integrating the equations of motion with a time step of 2 fs (36). MD trajectories 10–15 ns long are produced, which span the temperature range 160–300 K in 10 K increments. In addition, 100-ns-long trajectories are calculated for systems at 160, 180, and 300 K.

The energy trajectories are monitored to select the correct equilibration period for the MD runs. We find that at the end of 0.4 ns, all trajectories obtained at temperatures ≥ 200 K display energy fluctuations occurring around a baseline average energy value. For temperatures spanning the 200–300 K range, all runs are therefore equilibrated for 400 ps, and structures are recorded for 9.6 ns for subsequent data analysis. At lower temperatures, such convergence occurs in a much longer time period, and we have monitored each trajectory to use 9.6 ns of energy-equilibrated data. This requires at least a 5-ns equilibration period at 160–190 K. The energy trajectories of all the equilibrated systems are displayed in Fig. S1 of the Supporting Material.

Characterization of the dynamics via C_α positional fluctuations

The temporal dynamics of lysozyme is analyzed through positional autocorrelation functions of C_α carbons of each residue (Fig. 1). We and others find that these represent well the picosecond to nanosecond dynamics undergone by the folded protein (5,37). As positional fluctuations occur through local deformations of the protein backbone, rigid-body translations and molecular rotations are discarded to monitor the internal positional fluctuations only. Thus, the x-ray structure coordinates of lysozyme are treated as the reference, and every recorded conformation from the MD trajectories is superimposed on that of lysozyme. The superimposition is based on a minimization of the quaternion norm pertaining to the differences in atomic coordinates between the target (i.e., reference) structure and the MD recorded structure (38). The $N \times 3$ coordinate matrix $\mathbf{R}(t)$ is thus generated at each time step, t. A mean structure is defined as the average over these coordinate matrices at each temperature, $\langle \mathbf{R}(t) \rangle$. One can then write the positional deviations for each residue j as a function of time and temperature:

$$\Delta\mathbf{R}_j(t) \,=\, \mathbf{R}_j(t) - \langle\mathbf{R}_j(t)\rangle. \tag{1}$$

The positional autocorrelation function of the protein at a given temperature is an average over those calculated for all residues j as

$$C(t) \,=\, \overline{C_j(t)} \,=\, \overline{\left[\frac{\langle \Delta\mathbf{R}_j(0) \cdot \Delta\mathbf{R}_j(t)\rangle}{\langle \Delta\mathbf{R}_j(0)^2 \rangle} \right]}, \tag{2}$$

where the overbar represents an average over all residues, and the brackets indicate time averages.

In previous work, the initial decay (up to 50 ps) of the relaxation curves was found to be well characterized by a stretched exponential form, which was originally proposed to represent the multiexponential character of the relaxation phenomena (2,9). In this approach, the relaxation profiles of m distributed timescale processes, each contributing to the relaxation phenomena with weight a_i and individual relaxation time τ_i, are approximated by

$$C(t) \,=\, \sum_{i=1}^{m} a_i \exp\left(\frac{t}{\tau_i}\right) \approx \exp\left(\frac{t}{\tau}\right)^{\beta}, \tag{3}$$

where τ is a characteristic timescale and β is a dimensionless exponent. The physical basis of β has not been unambiguously established, yet it is known
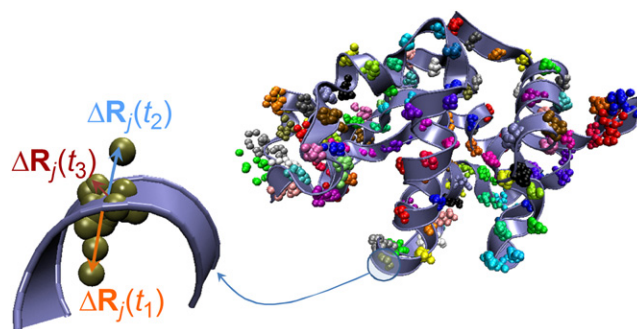

FIGURE 1 Sketch of the fluctuation vectors for C_α atoms. The protein conformations obtained from a chosen duration of the simulations are superimposed to obtain the average structure (depicted by the ribbon protein backbone); the instantaneous locations of the C_α atoms from 20 snapshots are shown as grouped spheres. The fluctuation vector of a given residue at time t is the vectorial difference in the current position of the C_α atom and its location in the average structure. At left, three vectors for a chosen atom j are magnified.

to carry the information of superposed single-exponential decays for the range $0 \leq \beta \leq 1$, following from its Laplace transform (39).

A time decay that may be interpreted via a stretched exponential function involves a dashpot with a time-dependent damping coefficient. The relaxation times obtained from fits to a stretched exponential model are scaled with the β-exponent; for $\beta < 1$, they effectively extend into timescales slower than τ. The initial infinite decay rate of a stretched exponential function follows from its power-law descriptor (40) and is physically ascribed to a purely elastic response preceding a viscous decay, accounting for the elastic contribution.

We find that it is not possible to model the total decay of the C_α relaxations using the stretched-exponential model as the single term. Although that model captures the initial decay profile faithfully, it misrepresents the long-time tail. The latter is well captured by a single-exponential decay model. Here, we propose what to our knowledge is a new functional form for nonexponential C_α relaxations by superposing a single-exponential term on the stretched-exponential model,

$$C(t) = a \exp\left[-\left(t/\tau_f\right)^\beta \right] + (1-a)\exp[-(t/\tau_s)]$$
$$= C_f(t) + C_s(t), \qquad (4)$$

to capture the overall relaxation profile. This treatment will provide information supplementary to that of our previous studies, where the focus was on the initial relaxations on the picosecond timescale (2,5,9). In Eq. 4, C_f and C_s are the contributions to the overall correlations from the fast and slow processes, respectively; τ_f and τ_s are the corresponding characteristic times. The positive front factor, $0 \leq a \leq 1$, describes the relative contributions from processes at these separated timescales. This composite model reflects the properties of both the short- and long-time behavior of the relaxations, as well as the crossover, as will be shown in the Results section.

In each calculation, we first proceed to determine the relaxation curves from the MD trajectories, as defined in Eq. 2. The 9.6-ns-long trajectories are used in the calculations yielding smooth curves, which are found to have completely relaxed to zero within 2 ns at all temperatures. We then find the best curve fit using the four-parameter model in Eq. 4. The fitting procedure used is based on an exhaustive search in the space of permitted values of a, β, τ_f, and τ_s. Initially, a and β are incremented in units of 0.1 in the interval (0,1), τ_f is incremented in units of 2 ps in the interval (2,50), and τ_s is incremented in units of 50 ps in the interval (50,1500). The optimal solution is identified as the global least-squares error minimizer between the fitted and calculated $C(t)$ curves. At different temperatures, the number of points used from the calculated $C(t)$ curves in the fitting procedure is actively adjusted to be in the range 1.0–1.6 ns. This prevents excessive weighting of long-time tails at the expense of a poorly captured crossover region. Once different solutions are reached in this way, the final solution is verified to be the global least-squares error minimizer over the 2-ns fully relaxed $C(t)$ curve. For this purpose, different regions of optimality are bracketed for all four parameters, and they are further subjected to a finer-grained exhaustive search in increments of 0.01 for β, 1 ps for τ_f, and 10 ps for τ_s.

In closing, we note that we have also tried an extended five-parameter model, where the second term in Eq. 4 includes an additional stretched exponent. We found that the exponent due to the slow processes always converges to 1 at all temperatures, reducing to the model in Eq. 4. Thus, we find that the C_α positional fluctuations decay with complex dynamics at short (picosecond) times, whereas their longer (nanosecond) time tails are well represented by a single-exponential decay.

RESULTS

Stationary behavior and the dynamical transition

For a stationary process, both the averages (e.g., the energy) and the fluctuations of a given physical observable (e.g., the

heat capacity) should be independent of the window length of the trajectory used. However, especially at low temperatures, such stationarity may not have been achieved; thus, the equilibration observed in Fig. S1 does not directly imply stationarity. We therefore calculate the average energies over different window lengths, w, and find that these averages are predominantly w-dependent at temperatures <200 K and are independent of w at temperatures >200 K. Sample trajectories at 180 and 300 K are shown in Fig. S2, where the averages over window lengths of $w = 200$, 400, 800, and 1600 ps are marked.

We further investigate the fluctuations in energy through the heat capacity. Previously, we monitored the heat capacity of several protein-solvent systems to determine the dynamical transition (5,9). In those studies, trajectories were partitioned into 200-ps-long chunks, and the heat capacity was calculated for each chunk from the energy fluctuations, $(\langle E^2 \rangle - \langle E \rangle^2)/k_B T^2$, where E is the instantaneous energy at temperature T and k_B is the Boltzmann constant. The heat capacity at a given temperature was then reported as an average over heat capacities obtained in different chunks. To characterize the presence/absence of stationary behavior in the system at different temperatures, we display in Fig. 2 the heat capacity calculated from trajectories of length 9600 ps by partitioning the trajectories into lengths of $w = 200$ ps (48 sets), $w = 400$ ps (24 sets), $w = 800$ ps (12 sets), and $w = 1600$ ps (6 sets). We find that at ≥ 200 K, the value of the heat capacity is the same for all trajectory lengths within the error bounds (the standard error of the mean of the data, ε, averaged over $n = 48$ sets obtained with $w = 200$ ps is represented by the shaded area,
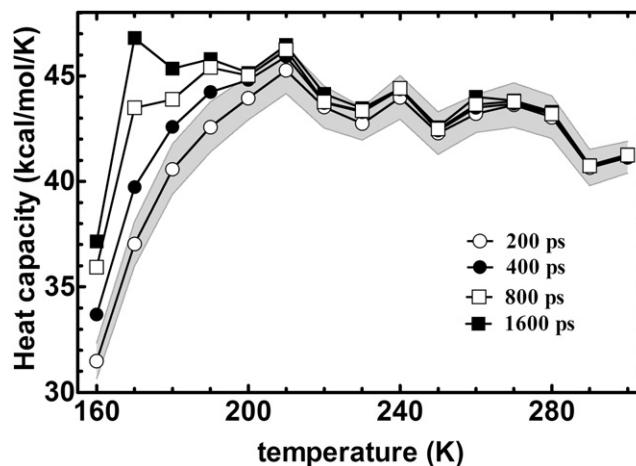

FIGURE 2 Heat capacity calculated from the energy trajectories using $(\langle E^2 \rangle - \langle E \rangle^2)/k_B T^2$. The different curves are obtained by partitioning the 9.6 ns trajectories into chunks of length 200, 400, 800, and 1600 ps (48, 24, 12, and 6 sets, respectively). The error bounds on the 200 ps window size calculations are shown by the shaded area. At physiological temperatures that emerge with the onset of the dynamical transition at 200 K, the length of the trajectory used does not affect the heat capacity values, within error bounds. Below 200 K, calculations depend on the length of data used, due to the nonstationary nature of the trajectories.

with $\varepsilon = \sigma/\sqrt{(n-1)}$, where σ is the standard deviation and n is the sample size). Conversely, at lower temperatures, the value of the heat capacity depends strongly on the window length of the trajectory used in the calculation, with larger fluctuations for larger w.

We postulate that at these lower temperatures, the system may only achieve piecewise stationarity. For the energy landscape, this would imply a system settling at local minima corresponding to metastable states, with occasional jumps between those states; these local minima, however, decorate a larger energy well that is the attractor for the whole system (41). This type of nonstationary behavior may be associated with glassiness for which the concept of metastable states is central (22).

Thus, we find that it is possible to define the dynamical transition as the point where the nonstationarity in the data disappears. This transition may be uniquely distinguished by probing the temperature below which the heat capacities obtained from various time windows do not display a convergent behavior. Note that such a measure of the dynamical transition is suggested for the first time that we know of in the literature. We will show that the transition from nonstationarity to stationarity in the time evolution of states has implications for the timescales governing the backbone dynamics of the protein. We next set out to characterize the origin of the various contributions to the protein-water dynamics below and above this transition temperature.

Characterization of the backbone dynamics

In the folded protein, although the backbone goes through incessant fluctuations, there are no conformational changes, as would be characterized by jumps in the backbone torsional angles (φ and ψ). This is contrary to the side-chain torsional angles, χ, where occasional jumps between the conformational states are recorded, in particular in the surface groups and at temperatures above the dynamical transition (5,9). The former is prohibited by the free-energetic cost of a required counterrotation in a nearby torsion in polymeric chains to prevent large displacements of chain ends (see (42) and references cited therein). Thus, the backbone dynamics provides an ideal reflection for the contributions of the picosecond-nanosecond timescales to the overall dynamics of the system due to the well-defined "average" for this observable.

By applying Eq. 2, we first study the average C_α relaxations over different time windows, w; i.e., the time averaging is done for different trajectory lengths. We make use of the $W = 100$-ns-long trajectories generated at 180 K and 300 K, and partition them into $m = \text{int}(W/w)$ subtrajectories, each of which we consider to be independent so far as the backbone relaxations are concerned. We find that the relaxation depends on the time window of observations, as displayed in Fig. 3 for the system at 300 K. At short time windows of observation, up to the order of a few nanoseconds, the relaxation curves shift
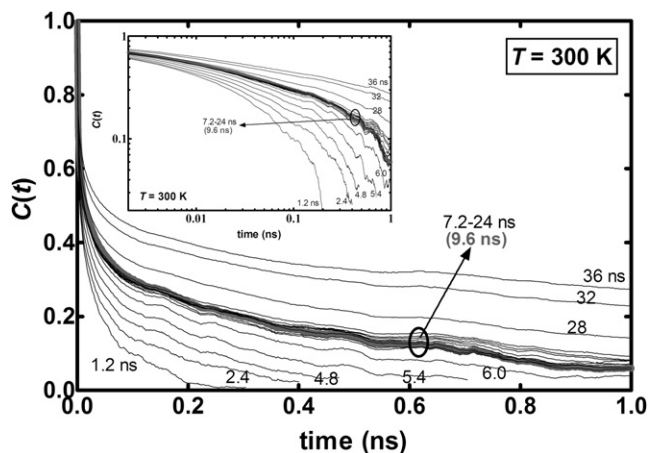

FIGURE 3 Relaxation curves, $C(t)$, at 300 K obtained for time windows of various length, $w = 1.2$–36 ns. We observe that the relaxation curves shift in regions below $ca.$ $w = 7.2$ ns and above $ca.$ $w = 24$ ns. In the interim values of w, the general behavior of the relaxation profile remains the same. We therefore choose $w = 9.6$ ns (*gray curve*) as the observation window in all the relaxation curve calculations. Note that the same results are reproduced at 180 K. The same data are plotted on the log-log scale in the inset.

to incorporate higher timescales. In the observation window of $w = 7.2$–24 ns, however, the relaxation profiles are nearly the same. This observation is interpreted as follows: at certain observation windows, there are some partially relaxed processes contributing to the decays. Thus, by elongating the observation time, w, a larger portion of these are included in the decay profiles. For certain w, however, all contributing processes have essentially relaxed, forming a plateau. Above $w = 24$ ns, the relaxation curves display shifts to higher relaxation times, indicating that longer-timescale processes begin to influence the dynamics in these longer windows. We note that the same results are reproduced at 180 K, indicating that this behavior is independent of the state of the system. To better treat the physics of fully relaxed processes, we choose $w = 9.6$ ns at all temperatures studied. We find that Eq. 4 provides an excellent model fit to all the curves. It represents the overall relaxation behavior, including the initial relaxation, the long-time tail, and the crossover between the two regimes. A sample relaxation curve of the backbone fluctuations is provided in Fig. S3. We display $C(t)$ at 180 K (below the dynamical transition) computed via Eq. 2, and the model fit of Eq. 4; the short-time region is magnified in the inset. The short-time decay depicts a sharp elbow below the dynamical transition, which smoothes out at higher temperatures.

The results of the fits to Eq. 4 are displayed for the exponent β in Fig. 4 a. MD results imply a biphasic character for the exponent, best described by a Boltzmann sigmoidal function as in our previous work (9). Levels are at $ca.$ $\beta = \sim0.2$ below the dynamical transition and at $ca.$ $\beta = \sim0.4$ above it; the transition temperature is predicted at 192 $\pm$ 2 K. Inasmuch as β describes the initial decay characteristics of the relaxation, these results are the same as those obtained
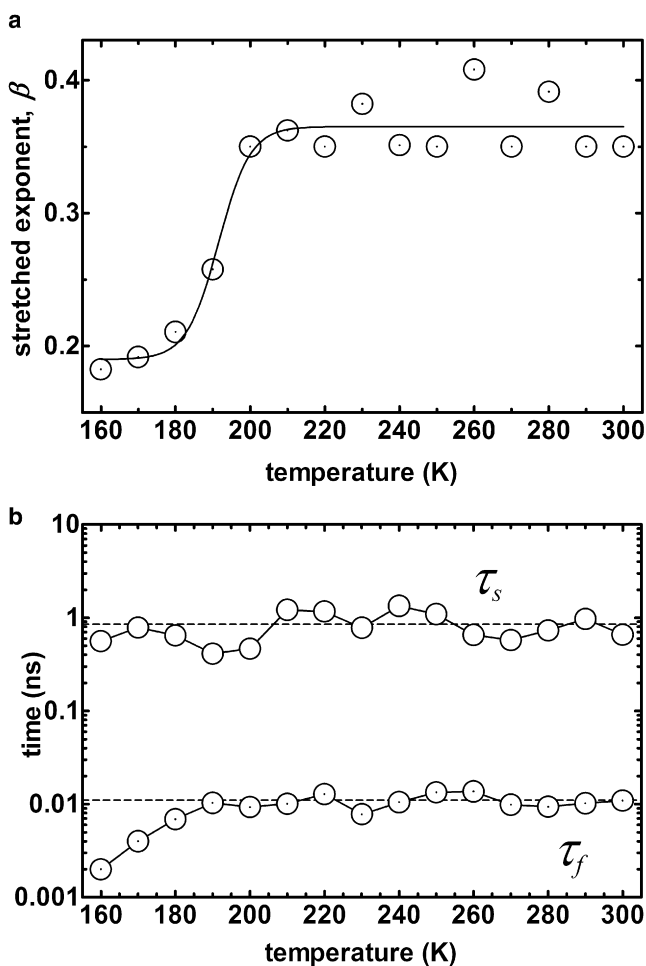
FIGURE 4 (*a*) Variation of the β-exponent with temperature, obtained from the fit of Eq. 4 to the calculated $C(t)$ curves. A Boltzmann sigmoidal function fit to the data predicts the dynamical transition at 192 $\pm$ 2 K, with β levels at 0.19 $\pm$ 0.01 below the dynamical transition and 0.37 $\pm$ 0.01 above it. (*b*) The characteristic decay times for the fast and slow processes, τ_f and τ_s, respectively, obtained from Eq. 4. The front factor, *a*, that dictates the relative contributions of the two timescales to the overall relaxation phenomenon is temperature-independent with $a \approx 0.74 \pm 0.02$.

in our previous work for the proteins BPTI (9) and endoglu-conase in water and glycerol (5), wherein only the initial decay characteristics were studied.

We further display the characteristic times of the fast and slow processes in Fig. 4 *b*. We find that τ_s is independent of temperature, with an average value of 0.8 $\pm$ 0.1 ns. On the other hand, τ_f shows an increase from ~2 ps to 11 $\pm$ 1 ps below the dynamical transition. We have previously explained this slowing down in the local relaxations using a model that incorporates the onset of the coupling between the local internal motions of the protein and the vicinal solvent layer with temperature (5). We have shown that such behavior is due to interplay between the decreased stiffness and the modified effective friction coefficient.

In sum, we characterize the backbone dynamics of the folded protein in the temperature range of interest and up

to timescales of ~20 ns. The backbone dynamics accommodates the superposition of fast local relaxations with a wide distribution of timescales ($\tau_f \approx$ 2–10 ps, $\beta \approx$ 0.2–0.4) and a more collective, single-timescale motion ($\tau_s \approx$ 0.8 ns). The relative contribution of the slow and fast timescales to the overall relaxation remains the same at all temperatures (i.e., $a \approx 0.74 \pm 0.02$).

In the previous subsection, we have shown that before the onset of the dynamical transition, the trajectories are nonstationary. Communication among the different local substates of the protein occur in a discrete manner, displaying a widely distributed range of relaxation times, characterized by $\beta \approx$ 0.2 and a picosecond timescale of τ_f. With the onset of the transition, continuous communications among the different local substates emerge, accompanied by the increase in β to ~0.4. The latter indicates a change in the distribution of relaxation times of the fast processes. The more collective motions are ever-present on the nanosecond timescale (τ_s). The question arises of how these slow-timescale motions influence the fast internal processes, if indeed they do, and we elaborate upon this in the next subsection.

Slow collective motions in proteins impose bounds on the distribution of timescales of fast processes

The interpretation of the long-observed trend for the β-increase from ~0.2 to 0.4 for the distribution of the local timescale motions with the onset of the dynamical transition (2,5,9) can now be rationalized. The overall effect of β is made transparent when the characteristic relaxation time of the total decay, τ^*, is studied. We define the integrated area τ^* of the autocorrelation function as the average time constant of the relaxation (40). Using the model in Eq. 4, this yields

$$\tau^* = \int_0^\infty C(t)dt = a \frac{\tau_f}{\beta} \Gamma\left(\frac{1}{\beta}\right) + (1-a)\tau_s, \quad (5)$$

where the solution involves the gamma function, $\Gamma(1/\beta)$. Due to the mild dependence of τ_f, as well as the independence of *a* and τ_s on temperature, the change in the overall relaxation behavior is governed by the stretched exponent, as displayed in Fig. 4. Thus, τ^* substantially decreases with onset of the dynamical transition, from ~1 ns to 200 ps (Fig. 5).

β-values near 0.2 at low temperatures render much faster timescale motions compatible with the slow one via an effective stretching of time. From another viewpoint, the distribution of the timescales of the fast processes at these temperatures are such that their effect is not felt by the overall system, so that the slow-process timescale emerges as the controlling factor, decoupled from the local motions, so that $\tau^* \approx \tau_s$. At temperatures where proteins are functional, on the other hand, there is substantial coupling between the collective motions and the more localized picosecond timescale degrees of freedom, so that $\tau_f < \tau^* < \tau_s$.
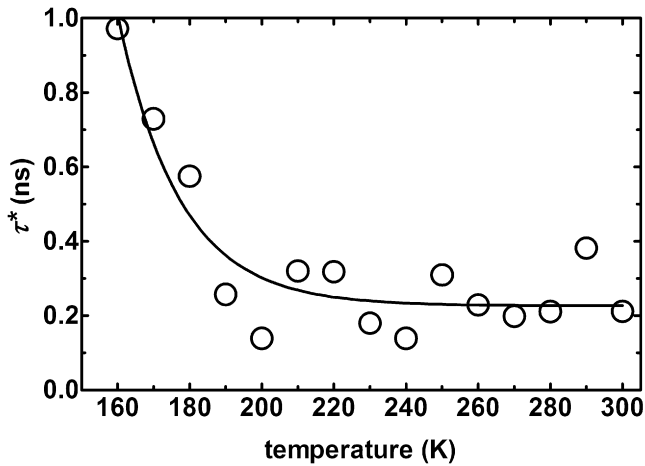
FIGURE 5   Characteristic timescale, τ^*, obtained from the combination of the fast and slow processes (Eq. 5). The line is an exponential fit to the data and is provided only as a guide to the eye.

FIGURE 6   Distribution function $G(\tau,\beta) = \tau\rho(\tau,\beta)$ for the fast-timescale contribution to the relaxation with the characteristic timescale of $\tau_f = 10$ ps. Curves for $\beta = 0.2$ and 0.4 are displayed. We find that the spectrum of relaxation times at $\beta = 0.4$ effectively vanishes at $\tau/\tau_f \cong 100$. On the other hand, for $\beta = 0.2$, the distribution has appreciable contributions from a much wider range, scanning the femtosecond to microsecond timescales. The area under the curves is normalized so that $\int_{-\infty}^{\infty} G(\tau,\beta)d\ln\tau = 1$.

We find that the coexistence of the slow timescale belonging to the collective motions with a couple of orders of magnitude separation imposes both an upper and a lower bound on these values, along with an understanding of how the protein and its environment operate. These are discussed below.

The upper bound on β is observed at physiological temperatures. Since there are a large number of fast exponential decaying processes contributing to the local dynamics in the protein environment, we assume a continuous spectrum of relaxation processes for the stretched-exponential component. The distribution function of the stretched exponential, $\rho(\tau,\beta)$, may be expressed as a convergent infinite summation (43):

$$\exp\left(-t/\tau_f\right)^\beta = \int_0^\infty \exp(t/\tau)\rho(\tau,\beta)d\tau$$
$$\rho(\tau,\beta) = -\frac{\tau_f}{\pi\tau^2}\sum_0^\infty \frac{(-1)^k}{k!}\sin(\pi\beta k)\Gamma(\beta k + 1)\left(\frac{\tau}{\tau_f}\right)^{\beta k + 1}.$$
(6)

Here, τ covers the processes distributed with all possible characteristic times. For $\beta = 1$, the sum reproduces the Dirac Δ function centered on τ_f. The probability of finding processes with timescales extending into very slow regimes increases as $\beta \to 0$. In Fig. 6, we display, for ($\beta = 0.2$, $\tau_f = 2$ ps) and ($\beta = 0.4$, $\tau_f = 11$ ps), the distribution functions $G(\tau, \beta) = \tau\rho(\tau, \beta)$, which is the form customarily used to present these distribution functions (43). The area under the curves are normalized, so that $\int_{-\infty}^{\infty} G(\tau,\beta)d\ln\tau = 1$.

We find that the spectrum of fast relaxation times at $\beta = 0.4$ effectively vanishes at $\tau = 1$ ns $\approx \tau_s$. On the other hand, for $\beta = 0.2$, the distribution has appreciable contributions from a much wider range, extending into the 10-ns timescales. Therefore, $\beta = 0.4$ forms an upper bound for delimiting the distribution of timescales at nanoseconds.
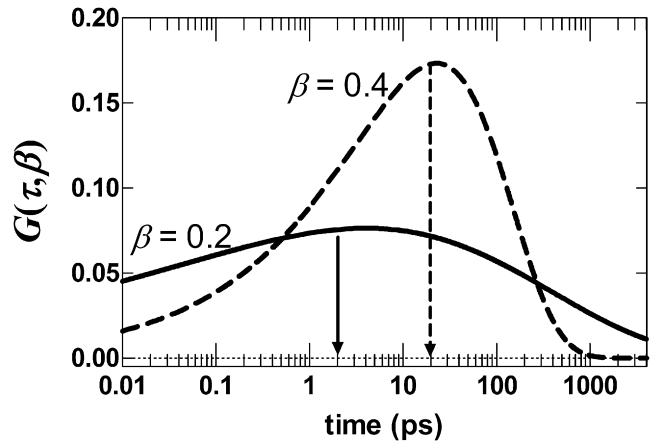
Note that Eq. 6 assumes a continuous distribution over an infinite number of contributing processes. In practice, a low β-value is attainable even with two processes, with the condition that their timescales are well separated (9). In proteins, we expect to find a large, but finite, number of fast-timescale motions, and the above discussion merely points to the fact that these are widely distributed over many orders of magnitude below the dynamical transition. On the other hand, they are confined to a smaller region at higher temperatures, enabling their efficient coupling with the collective motions. In other words, instead of there being very slow or very fast isolated relaxation events, they are confined to a narrower region where communication among different processes is made possible.

To rationalize the lower bound of $\beta \approx 0.2$ observed at low temperatures, we rely on a model of hierarchically constrained dynamics, accounting for the slowing down in hierarchically ordered materials such as glasses and proteins (44–46). In this model, the local energy landscape is viewed as consisting of microstates. The system has to overcome a succession of energy barriers of increasing height, belonging to the n substates decorating each microstate. From the ongoing discussion related to Fig. 2, we conjecture that below the dynamical transition, the noncommunicating local substates lead to such hierarchy with a wide distribution of timescales. The assumption of noncommunication is based on the observed nonstationary behavior of the trajectories.

The characteristic time of each contributing process i must overcome the barrier ΔE_i by the Arrhenius law $\tau_i \sim \exp(\Delta E_i/kT)$, where k is the Boltzmann factor. Since the dynamical model relies on the hierarchical sampling of states, the total energy spent in a microstate with n substates is $\Delta E_n = \Sigma_i \Delta E_i$.

The total time spent in a given microstate, T_n, is then a series of multiplications of the timescales of the contributing substates, $T_n \sim \Pi_i \tau_i$. An observable relaxing macroscopically is given by $\langle \exp(t/T_n) \rangle$, where the average is carried out over the distribution of T_n, due to the presence of the many microstates, each decorated with n substates. Using extreme deviation theory, it is shown (44,45) that this average may be expressed as the stretched exponential, $\exp(t/\tau)^\beta$. Furthermore, the integral in Eq. 5 becomes

$$\tau^* = a \frac{\tau_f}{\beta} \Gamma\left(\frac{1}{\beta}\right) + (1-a)\tau_s = a\langle T_n \rangle + (1-a)\tau_s = \tau_s + a[\langle T_n \rangle - \tau_s].$$

Thus, one can write

$$\tau^* \sim \exp \frac{\Delta E_s}{kT} + a \left[\left\langle \exp \frac{\Delta E_n}{kT} \right\rangle - \exp \frac{\Delta E_s}{kT} \right],$$

where, again, the average is carried out over the distribution of T_n. The energy associated with the convex hull of the n substates is maintained by the prevalent nanosecond-timescale collective motions of the protein. The situation where the average energy spent in the microstates is compensated by that of the slower relaxations of the hull forms a limiting case. This limiting case sets the last term in square brackets to zero. Equating the two contributing terms, with $\tau_f = 2$ ps, and $\tau_s = 0.8 \pm 0.1$ ns, gives the limiting value of β:

$$\frac{\tau_f}{\beta} \Gamma\left(\frac{1}{\beta}\right) \approx \tau_s \rightarrow \beta \cong 0.18 \pm 0.01.$$

Smaller values of β give explosively slower overall relaxation times (τ^*), which would let the fast processes take control over the whole dynamics; values >0.2 yield small contributions to the sum in Eq. 5, so that τ^* is driven toward the subnanosecond timescale. Thus, the protein-water system below the dynamical transition may be assumed to have converged to this limiting behavior. Note that, above the dynamical transition, the hierarchical model no longer applies, and the observed multiexponential decay may be explained by, for example, subdiffusive behavior (47).

CONCLUDING REMARKS

In this study, we have studied the backbone dynamics of folded proteins spanning timescales on five orders of magnitude and in a wide temperature window, to gain a general understanding of the dynamics. We show that the dynamical transition in the protein results in part from the various timescales coming into concert, allowing large structural motions to occur through the cooperativity of the various accompanying local motions.

The view that proteins exhibit a hierarchical arrangement of conformational substates due to structural constraints is well established (48,49). The presence of tiers along the conformational energy surface of proteins has also been exploited experimentally to study various dynamic properties. For instance, Samuni et al. have studied kinetic allosteric intermediates and equilibrium populations of hemoglobin (50), as well as the sequential onset of different tiers of dynamics in myoglobin (51). Their results were obtained under the assumption that in an environment that leads to the slowing down of timescales (e.g., by the effect of viscosity, temperature, or mutants), there will be a separation of timescales between the different processes contributing to the protein dynamics; those that are slaved the most will influence the dynamics with the largest delay.

In a similar way, in this work, as the system is cooled below the dynamical transition temperature, it is possible to resolve various substates that contribute to the dynamics, captured in the nonstationary nature of the trajectories. At these temperatures, where functionality has not been onset, the localized "fast" processes display heterogeneous dynamics with a characteristic timescale on the order of 2–10 ps. Their distribution covers several orders of magnitude (Fig. 6) and even accommodates very slow relaxations with timescales beyond that of nanoseconds with low probabilities.

We further find that below the dynamical transition temperature, the partitioning and the miscommunication of the local motions manifest themselves in the nonstationary behavior of the total energy trajectories. Communication among the substates is established above the transition temperature with the redistribution of the timescales of these local motions, as evidenced by an increase in the stretched exponent that characterizes these distributions (Fig. 4), and as detected in the heat capacity calculations (Fig. 2). A slower timescale on the order of 1 ns is present at all temperatures. We conjecture that this timescale belongs to segmental motions displaying homogenous dynamics.

Certain motions that only appear with the proper coupling of the local protein motions and the vicinal layer of solvent, such as side-chain motions (19,52,53), belong to the group of fast processes and take part in the redistribution of timescales (5). The ever-present slower dynamics acts as an agent that bridges the timescales. We argue that the coexistence of the slow timescale belonging to the collective motions imposes both an upper and a lower bound on these values, along with an understanding of how the protein and its environment operate. Thus, in environments that lead to the slowing down of timescales, certain motions may be not just delayed but completely hindered.

SUPPORTING MATERIAL

REFERENCES

1. Zaccai, G. 2000. How soft is a protein? A protein dynamics force constant measured by neutron scattering. *Science*. 288:1604–1607.

2. Baysal, C., and A. R. Atilgan. 2002. Relaxation kinetics and the glassiness of proteins: the case of bovine pancreatic trypsin inhibitor. *Biophys. J.* 83:699–705.

3. Tarek, M., and D. J. Tobias. 2002. Role of protein-water hydrogen bond dynamics in the protein dynamical transition. *Phys. Rev. Lett.* 88:138101–138104.

4. Pal, S. K., J. Peron, and A. H. Zewail. 2002. Biological water at the protein surface: dynamical solvation probed directly with femtosecond resolution. *Proc. Natl. Acad. Sci. USA.* 99:1763–1768.

5. Atilgan, C., A. O. Aykut, and A. R. Atilgan. 2008. How a vicinal layer of solvent molecules modulates the dynamics of proteins. *Biophys. J.* 94:79–89.

6. Barron, L. D., L. Hecht, and G. Wilson. 1997. The lubricant of life: a proposal that solvent water promotes extremely fast conformational fluctuations in mobile heteropolypeptide structure. *Biochemistry*. 36:13143–13147.

7. Oleinikova, A., N. Smolin, and I. Brovchenko. 2007. Influence of water clustering on the dynamics of hydration water at the surface of a lysozyme. *Biophys. J.* 93:2986–3000.

8. Gabel, G., M. Weik, B. P. Doctor, A. Saxena, D. Fournier, L. Brochier, F. Renault, P. Masson, I. Silman, and G. Zaccai. 2004. The influence of solvent composition on global dynamics of human butyrylcholinesterase powders: a neutron scattering study. *Biophys. J.* 86:3152–3165.

9. Baysal, C., and A. R. Atilgan. 2005. Relaxation kinetics and the glassiness of proteins: coupling of time scales. *Biophys. J.* 88:1570–1576.

10. Duan, Y., and P. A. Kollman. 1998. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science*. 282:740–743.

11. Murphy, K. P., P. L. Privalov, and S. J. Gill. 1990. Common features of protein unfolding and dissolution of hydrophobic compounds. *Science*. 247:559–561.

12. Tsai, A. M., T. J. Udovic, and D. A. Neumann. 2001. The inverse relationship between protein dynamics and thermal stability. *Biophys. J.* 81:2339–2343.

13. Bicout, D. J., and G. Zaccai. 2001. Protein flexibility from the dynamical transition: a force constant analysis. *Biophys. J.* 80:1115–1123.

14. He, Y. F., P. I. Ku, J. R. Knab, J. Y. Chen, and A. G. Markelz. 2008. Protein dynamical transition does not require protein structure. *Phys. Rev. Lett.* 101:178103.

15. Doi, M., and S. F. Edwards. 1986. The Theory of Polymer Dynamics. Clarendon Press, Oxford, UK.

16. Qian, H. 2000. Single-particle tracking: Brownian dynamics of viscoelastic materials. *Biophys. J.* 79:137–143.

17. Kurchan, J. 2005. In and out of equilibrium. *Nature*. 433:222–225.

18. LeBard, D. N., and D. V. Matyushov. 2008. Glassy protein dynamics and gigantic solvent reorganization energy of plastocyanin. *J. Phys. Chem. B.* 112:5218–5227.

19. Declercq, J. P., C. Evrard, V. Lamzin, and J. Parello. 1999. Crystal structure of the EF-hand parvalbumin at atomic resolution (0.91 Å) and at low temperature (100 K). Evidence for conformational multistates within the hydrophobic core. *Protein Sci.* 8:2194–2204.

20. Boon, J.P., and S. Yip. 1991. Molecular Hydrodynamics. Dover, Mineola, NY.

21. Xia, X., and P. G. Wolynes. 2001. Microscopic theory of heterogeneity and nonexponential relaxations in supercooled liquids. *Phys. Rev. Lett.* 86:5526–5529.

22. Debenedetti, P. 1996. Metastable Liquids. Concepts and Principles. Princeton University Press, Princeton, NJ.

23. Kneller, G. R., and K. Hinsen. 2004. Fractional Brownian dynamics in proteins. *J. Chem. Phys.* 121:10278–10283.

24. Piazza, F., and P. De Los Rios. 2005. Slow energy relaxation of macromolecules and nanoclusters in solution. *Phys. Rev. Lett.* 94:145502.

25. Kakalios, J., R. A. Street, and W. B. Jackson. 1987. Stretched-exponential arising from dispersive diffusion of hydrogen in amorphous silicon. *Phys. Rev. Lett.* 59:1037–1040.

26. Istratov, A. A., and O. F. Vyvenko. 1999. Exponential analysis in physical phenomena. *Rev. Sci. Instrum.* 70:1233–1257.

27. Williams, G., and D. C. Watts. 1970. Non-symmetrical dielectric relaxation behavior arising from a simple empirical decay function. *Trans. Faraday Soc.* 66:80–85.

28. Kohlrausch, R. 1854. Theorie des elektrischen Ruckstandes in der Leidener Flasche. *Ann. Phys Chem. (Leipzig).* 91:179–214.

29. Wu, F., and B. A. Huberman. 2007. Novelty and collective attention. *Proc. Natl. Acad. Sci. USA.* 104:17599–17601.

30. Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, et al. 2000. The Protein Data Bank. *Nucleic Acids Res.* 28:235–242.

31. Vaney, M. C., S. Maignan, M. Riès-Kautt, and A. Ducruix. 1996. High resolution structure (1.33 Å) of a HEW lysozyme tetragonal crystal grown in the APCF apparatus. Data and structural comparison with a crystal grown under microgravity from SpaceHab-01 mission. *Acta Crystallogr. D Biol. Crystallogr.* 52:505–517.

32. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38.

33. Brooks, B. R., R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, et al. 1983. Charmm: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* 4:187–217.

34. Darden, T., L. Perera, L. Li, and L. Pedersen. 1999. New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure.* 7:R55–R60.

35. Phillips, J. C., R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, et al. 2005. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26:1781–1802.

36. Swope, W. C., H. C. Andersen, P. H. Berens, and K. R. Wilson. 1982. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: application to small water clusters. *J. Chem. Phys.* 76:637–649.

37. Laine, E., J. D. Yoneda, A. Blondel, and T. Malliavin. 2008. The conformational plasticity of calmodulin upon calcium complexation gives a model of its interaction with the oedema factor of *Bacillus anthracis. Proteins.* 71:1813–1829.

38. Kearsley, S. K. 1989. On the orthogonal transformation used for structural comparisons. *Acta Crystallogr. A.* 45:208–210.

39. Pollard, H. 1946. The representation of $\exp(-X^\lambda)$ as a Laplace integral. *Bull. Am. Math. Soc.* 52:908–910.

40. Berberan-Santos, M. N., E. N. Bodunov, and B. Valeur. 2005. Mathematical functions for the analysis of luminescence decays with underlying distributions 1. Kohlrausch decay function (stretched exponential). *Chem. Phys.* 315:171–182.

41. Frauenfelder, H., S. G. Sligar, and P. G. Wolynes. 1991. The energy landscapes and motions of proteins. *Science*. 254:1598–1603.

42. Baysal, C., B. Erman, and I. Bahar. 1994. Contribution of short-range intramolecular interactions to local chain dynamics. *Macromolecules*. 27:3650–3657.

43. Lindsey, C. P., and G. D. Patterson. 1980. Detailed comparison of the Williams-Watts and Cole-Davidson functions. *J. Chem. Phys.* 73:3348–3357.

44. Frisch, U., and D. Sornette. 1997. Extreme deviations and applications. *J. Phys. I.* 7:1155–1171.

45. Klafter, J., and M. F. Shlesinger. 1986. On the relationship among three theories of relaxation in disordered systems. *Proc. Natl. Acad. Sci. USA.* 83:848–851.

46. Metzler, R., J. Klafter, and J. Jortner. 1999. Hierarchies and logarithmic oscillations in the temporal relaxation patterns of proteins and other complex systems. *Proc. Natl. Acad. Sci. USA.* 96:11085–11089.

47. Bouchaud, J.-P. 2008. Anomalous relaxation in complex systems: from stretched to compressed exponentials. *In* Anomalous Transport: Foundations and Applications. R. Klages, G. Radons, and I. M. Sokolov, editors. Wiley, VCH, Berlin. pp. 327–345.

48. Frauenfelder, H., B. H. McMahon, R. H. Austin, K. Chu, and J. T. Groves. 2001. The role of structure, energy landscape, dynamics, and allostery in the enzymatic function of myoglobin. *Proc. Natl. Acad. Sci. USA.* 98:2370–2374.

49. Nienhaus, G. U., J. D. Muller, B. H. McMahon, and H. Frauenfelder. 1997. Exploring the conformational energy landscape of proteins. *Physica D.* 107:297–311.

50. Samuni, U., C. J. Roche, D. Dantsker, and J. M. Friedman. 2007. Conformational dependence of hemoglobin reactivity under high viscosity conditions: the role of solvent slaved dynamics. *J. Am. Chem. Soc.* 129:12756–12764.

51. Samuni, U., D. Dantsker, C. J. Roche, and J. M. Friedman. 2007. Ligand recombination and a hierarchy of solvent slaved dynamics: the origin of kinetic phases in hemeproteins. *Gene.* 398:234–248.

52. Zanotti, J. M., M. C. Bellissent-Funel, and J. Parello. 1999. Hydration-coupled dynamics in proteins studied by neutron scattering and NMR: the case of the typical EF-hand calcium-binding parvalbumin. *Biophys. J.* 76:2390–2411.

53. Markelz, A. G., J. R. Knab, J. Y. Chen, and Y. F. He. 2007. Protein dynamical transition in terahertz dielectric response. *Chem. Phys. Lett.* 442:413–417.