

Dynamic Fingerprints of Protein Thermostability Revealed by Long Molecular Dynamics

Enrique Marcos, Aurora Jiménez, and Ramon Crehuet*

Department of Biological Chemistry and Molecular Modelling, Institute of Advanced Chemistry of Catalonia (IQAC - CSIC), E-08034 Barcelona, Spain

S Supporting Information

ABSTRACT: The dynamical requirements for protein thermostability are a subject of intense debate since different techniques are sensitive to different dynamical processes. The present investigation arises from a neutron scattering experiment pointing to the lower temperature dependence of the flexibility of thermophilic proteins as a mechanism of enhanced thermostability. By means of 200 ns molecular dynamics simulations at different temperatures, we have investigated the differences in internal dynamics of the thermo-mesophilic pair of proteins studied in the experiment. The present work exceeds the time scales explored by the experiment and former studies on other thermo-mesophilic pairs by several orders of magnitude. Our simulations confirm the different thermal behavior observed in the experiment and suggest that both reduced coil segments and salt bridge interactions contribute to lowering the increase in flexibility with temperature. Moreover, the mesophilic protein exhibits a more heterogeneous distribution of residue mobilities involving more local motions. We suggest that the more collective motions of the thermophilic protein underlie a broader energy landscape.

INTRODUCTION

Enzymes achieve their outstanding efficiency under very specific environmental conditions. Factors like temperature, pressure, or salt concentration have a tremendous impact on enzyme activity, but the broad versatility of evolution allows for finding enzymes adapted to extremely diverse environments. Of particular interest are thermophilic and hyperthermophilic proteins from organisms that grow at very high temperatures ranging from 50 to 120 °C. Such adaptation to high temperatures requires proteins to be very resistant to heat denaturation, and elucidating the origin of such resistance could provide clues for designing proteins with enhanced thermal stability for a wide range of applications.

There has been a surge in experimental and computational studies comparing thermophilic proteins with their corresponding homologues working at room temperature, which are known as mesophilic. From studies in the past 20 years, it seems that there is not a unique strategy adopted by evolution to thermostabilize proteins. Structural and amino acid sequence comparisons among a vast range of thermophilic and mesophilic proteins point to some features that correlate with increased thermostability, namely, a larger proportion of charged residues, a more compact hydrophobic core, and shortened loops at the surface.^{1–3} The dynamical requirements for protein thermostability, however, are more controversial. Since thermophilic proteins unfold at a higher temperature and are less active than their mesophilic homologues at lower temperatures, thermophilic proteins have been traditionally considered as more rigid. According to the *corresponding states* hypothesis, thermophilic and mesophilic proteins achieve a similar flexibility at their respective temperatures for maximum activity. Nevertheless, experimental and simulation techniques able to explore atomic motions on different time scales have indicated that the panorama is more complex and that there is

not a unique strategy for protein thermostability. Some of these studies found thermophilic proteins to be more rigid than their mesophilic homologues,^{4–8} whereas others showed the opposite.^{9–14} This lack of consensus stems from the absence of a unique mechanism of thermostability and, on the other hand, from the fact that these techniques explore different aspects of protein dynamics among the vast diversity of events that occur in a broad range of time scales. Enhanced flexibility in thermophilic proteins^{9,11} can entail an increase in conformational entropy of the native state that lowers the entropy change upon unfolding providing more stability.¹⁵ On the other hand, NMR relaxation experiments⁸ show that large-amplitude motions in a thermophilic adenylate kinase do not occur as frequently as in a mesophilic homologue at low temperatures, which supports the corresponding states hypothesis. This is a clear indication that a proper definition of flexibility requires the specification of its time scale and type of motion. For this reason, the flexibility regarding different dynamic events is not directly comparable, and ultimately, its linkage to stability and function does not have to be necessarily the same.

One of the most intriguing dynamical mechanisms of protein thermostability was put forward few years ago by Zaccai and co-workers.¹¹ By means of elastic neutron scattering experiments, they observed that the mean-square displacement (MSD) on short time scales (~100 ps) of a thermophilic enzyme was less sensitive to temperature changes than that of the corresponding mesophilic homologue. These results motivated the authors to suggest that this can be a plausible mechanism for thermophilic proteins to control the structural fluctuations at high temperatures, and so avoid unfolding. In a recent work,¹⁶ however, we previously showed that such an amazing difference in the

Received: December 7, 2011

Published: January 30, 2012



thermal behavior of MSD between the two proteins, indeed, arises to a large extent from significant differences in diffusion under crowding conditions. After having rationalized the mechanism behind the experimental observation, in this work we aim to explore a broader range of time scales than in the experiment and characterize fundamental differences in the intramolecular dynamics of both proteins by means of molecular dynamics simulations (MD). In particular, we have focused on understanding the factors that can determine a different sensitivity of intramolecular motions to changes with the temperature. We have performed 200 ns simulations of each protein at three different temperatures (280, 300, and 320 K), adding to a total simulation time of 1.2 μ s. This exceeds the time scales explored in previous MD studies comparing other thermo-mesophilic pairs of proteins.

METHODS

Structural Data. The homotetrameric structures of the thermophilic enzyme (malate dehydrogenase from the *Methanococcus jannaschii* archaea) and its mesophilic homologue (lactate dehydrogenase from pig muscle) were obtained from the 1HYG¹⁷ and 9LDT¹⁸ entries in the Protein Data Bank, respectively. Both structures have 313 and 331 residues per chain, respectively, and have a 28% sequence identity.

Molecular Dynamics (MD). All MD simulations were performed with the Gromacs simulation package (version 4.0.5).¹⁹ Each protein was simulated under periodic boundary conditions using a rhombic dodecahedral water box. Standard protonation states were assigned to all protein residues, -16 and 8 electronic units being the total charge of the thermophilic and mesophilic proteins, respectively. The electric neutrality of the systems was achieved by adding Na⁺ counterions to the thermophilic protein system and Cl⁻ counterions to the mesophilic one. Substrates bound to the crystallographic structures were removed. The total numbers of atoms in the simulations of the thermophilic and mesophilic proteins is 89 229 and 90 393, respectively.

The two solvated systems under study were subjected to 200 ns simulations at temperatures set to 280, 300, and 320 K. We used the OPLS all-atom force field²⁰ for describing protein interactions and TIP3P²¹ to model water molecules. Short-range electrostatic interactions were calculated explicitly with a 10 Å cutoff, and long-range electrostatic interactions were computed via the Particle Mesh Ewald method²² using a grid spacing of 1.2 Å and a fourth order spline interpolation. Lennard-Jones interactions were calculated using a switch function between 0.8 and 0.9 nm. An integration time step of 2 fs was used by constraining all bonds with LINCS.²³ The temperature was controlled with the Berendsen thermostat²⁴ using a coupling time constant of 0.1 ps. The pressure was kept constant at 1 bar via the Berendsen barostat²⁴ using an isotropic compressibility of 4.5×10^{-5} bar⁻¹ and a coupling constant of 0.5 ps. Before production runs were started, the structure of the solvated protein was energy minimized with the steepest descent algorithm. Next, solvent surrounding the protein was equilibrated by running a MD simulation at the target temperature using harmonic position restraints on the heavy atoms of the protein with a force constant of 1000 kJ mol⁻¹ nm⁻². Finally, a 200 ns trajectory was carried out, and the last 160 ns were used for production. Snapshots were saved every 20 ps for subsequent analysis. In this work, we have only considered intramolecular protein motions. Translational and

rotational diffusive motions were removed by superimposing every frame into the first frame.

Time Scale Dependence of Protein Dynamics. The time-dependent mean-square displacement, $MSD(\Delta t)$, was calculated from time trajectories as

$$MSD(\Delta t) = \frac{1}{N(T - \Delta t + 1)} \sum_i \sum_{t=0}^{T-\Delta t} [\mathbf{R}_i(t + \Delta t) - \mathbf{R}_i(t)]^2 \quad (1)$$

where \mathbf{R}_i is the position vector of atom i , N is the number of atoms, T is the time length of the trajectory, and Δt is the time separation between saved frames.

Principal Component Analysis (PCA). It is possible to describe the conformational fluctuations in terms of collective variables that concentrate the most important dynamic information from the trajectory and filter out the noise from irrelevant local motions. This can be done by doing a principal component analysis (PCA) of MD trajectories, also known as *essential dynamics*.²⁵ First, the covariance matrix from the fluctuations of atomic positions is built as

$$C_{ij} = \langle (q_i - \langle q_i \rangle)(q_j - \langle q_j \rangle) \rangle \quad (2)$$

where q_k is the k component of vector $\mathbf{q} = \{q_1, \dots, q_{3N}\}$ which defines the coordinates of the system of N atoms. C is a symmetric $3N \times 3N$ matrix, whose diagonal elements represent the atomic mean-square fluctuations and the off-diagonal elements, the correlation between two variables. The eigenvectors of C are $3N$ -dimensional vectors that indicate the direction of motion of the principal components or *essential modes*, and the corresponding eigenvalues correspond to the mean-square fluctuations associated with the mode. Each eigenvector defines the direction of motion as a displacement from the average structure. The trajectory can be projected into a principal mode k (\mathbf{v}_k) as: $p_k(t) = \mathbf{v}_k^t(\mathbf{q}(t) - \langle \mathbf{q} \rangle)$. Such analysis is particularly useful for characterizing conformational transitions with different amplitudes occurring in the course of the simulation.

The projection $p_k(t)$ has short-term fluctuations around the global trend determined by the direction of principal component k . To quantify the amplitude of these fluctuations, we have first determined the trend of the projection at each time t by calculating a moving average $\bar{p}_k(t)$, which have been calculated as

$$\bar{p}_k(t) = \frac{1}{2n} \sum_{m=t-n}^{t+n} p_k(m) \quad (3)$$

where $\bar{p}_k(t)$ is an average over all frames between $t + n$ and $t - n$. We have set n to 227 points, which implies taking the average over 4.54 ns. Subsequently, we have calculated the deviation of $p_k(t)$ from the trend $\bar{p}_k(t)$ as a variance (σ^2):

$$\sigma^2 = \frac{1}{T - 2n} \sum_{t=n}^{T-n} [p_k(t) - \bar{p}_k(t)]^2 \quad (4)$$

where T is the total number of frames

Calculation of Degree of Packing. To describe the packing density at each protein residue, we have adopted the definition of Chennubhotla and Bahar²⁶ originally introduced in a Markov model of communication among residues. The

interaction between pairs of residues is defined with an affinity matrix, \mathbf{A} , whose a_{ij} elements are calculated as

$$a_{ij} = \frac{N_{ij}}{\sqrt{N_i N_j}} \quad (5)$$

where N_{ij} is the number of atom–atom contacts between residues i and j within a cutoff distance of 4 Å and N_i and N_j are the number of heavy atoms of both residues. From the affinity matrix, the degree diagonal matrix $\mathbf{D} = \{d_i\}$ is defined, where $d_i = \sum_{j=1}^N a_{ij}$, and reflects the packing density at each residue.

Collectivity of Modes of Motion. The collectivity index²⁷ of a given mode of motion, e.g., normal mode or principal component, gives a measure of how many atoms contribute to this mode. We have used this index to characterize the dynamical heterogeneity associated with the principal components. The collectivity (k_i) of a given mode \mathbf{r}_i is defined as

$$k_i = \frac{1}{N} \exp \left\{ - \sum_{n=1}^N u_{i,n}^2 \log u_{i,n}^2 \right\} \quad (6)$$

where i is the mode index, n runs over N atoms, and

$$u_{i,n}^2 = \alpha \frac{|\mathbf{r}_{i,n}|^2}{m_n} \quad (7)$$

where m_n is the mass of atom n and α is a normalization constant to give $\sum_{n=1}^N u_{i,n}^2 = 1$. The collectivity tends to decrease for local motions (high-frequency modes), which means that a lower number of atoms contribute to the motion described by the mode. Highly collective or *cooperative* motions, on the other hand, are those in which a large proportion of atoms participate.

RESULTS AND DISCUSSION

Overall Flexibility at Different Time Scales and Temperatures. First, we have explored how the thermophilic and mesophilic proteins differ in both the time scale and temperature dependence of their conformational fluctuations. In the present investigation, this has been determined with the mean-square-displacement (MSD), as given by eq 1, on time scales ranging from 20 ps to 160 ns at temperatures set to 280, 300, and 320 K. The results are summarized in Figure 1. Panel a shows the MSDs taking into account the full trajectories, whereas panel b focuses on time scales up to 50 ns, including an approximation to the error. The errors of each curve have been determined with the bootstrapping method by calculating the MSD of 100 randomly generated 50 ns chunks of each 160 ns trajectory.

On time scales below 5 ns (see Figure 1b), the MSD values of the thermophilic protein are above the mesophilic one at the three temperatures studied. In other words, the thermophilic protein is more flexible than the mesophilic one on short time scales. This shows that the observation in our previous theoretical study¹⁶ that the thermophilic protein is more mobile than the mesophilic one at the 100 ps time scale also applies to a broader range of time scales. Interestingly, at the highest temperature (320 K) and beyond the 5 ns time scale, the MSD of the mesophilic protein approaches that of the thermophilic protein. Thus, the increase in flexibility with temperature and time scale differs between the two proteins. Nevertheless, for longer time scales, it is not possible to ascertain if the mesophilic protein becomes more flexible than

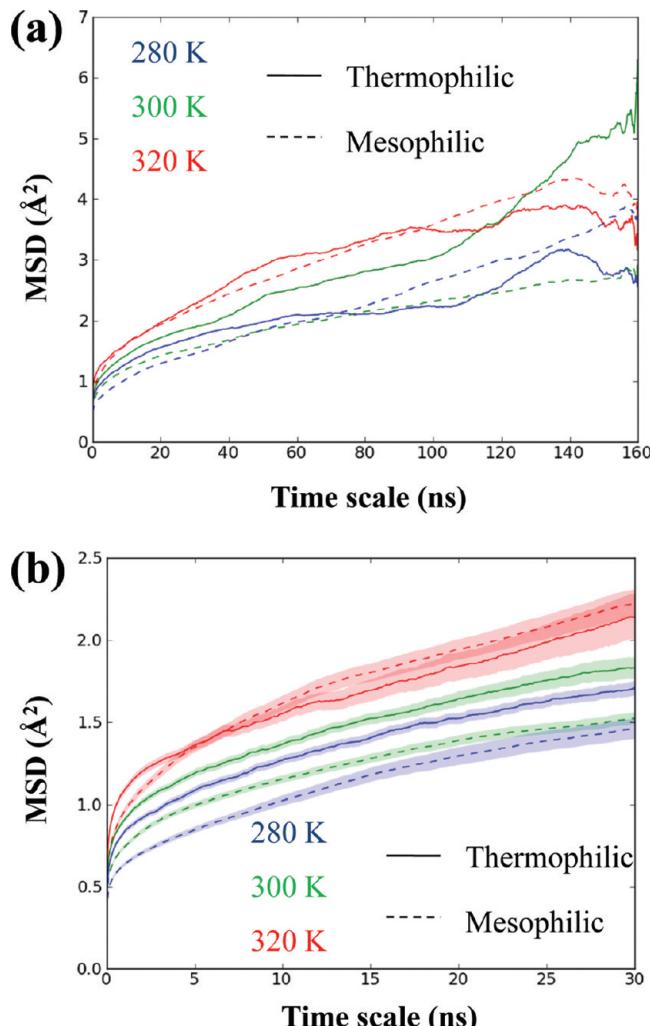


Figure 1. Mean square displacements of the thermophilic (solid lines) and mesophilic (dashed lines) over the 0–160 ns time scale range at three different temperatures: 280 K (blue), 300 K (green), and 320 K (red). Panel b shows the 0–30 ns range of panel a. MSD values are averaged over all backbone atoms. We have computed the errors bars in panel b with the bootstrapping statistical method, which implies computing the MSD from randomly generated chunks of the trajectory.

the thermophilic one, since a 160 ns trajectory does not provide enough statistics to characterize atomic motions beyond the 10–20 ns time scale. This could be done by lengthening the simulation by an order of magnitude at least.

The fact that at 320 K the mesophilic protein achieves similar flexibility to that of the thermophilic protein on the ~5 ns time scale is indicative of a different sensitivity to temperature changes in the two proteins. Figure 2 shows the temperature dependence of MSD (dMSD/dT) on different time scales. The dMSD/dT quantity has been determined by taking the slope of a linear fit of MSD versus temperature on each time scale. Figure 2 clearly illustrates how the temperature dependence of MSD increases with the time scale in both proteins, this increase of dMSD/dT being more accentuated in the mesophilic protein. On time scales ranging from 20 ps to 100–200 ps, dMSD/dT is very similar in the two proteins, but it is beyond the 300 ps time scale that the difference in dMSD/dT is manifested. The key message conveyed by this figure is that experimental techniques probing atomic motions on

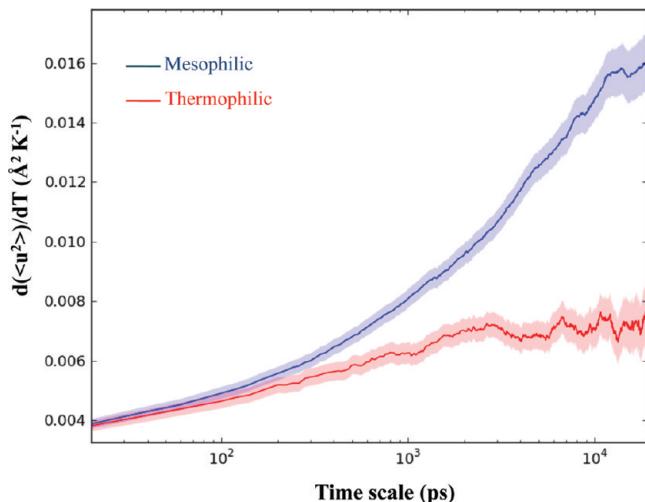


Figure 2. Representation of the temperature-dependence of MSD from backbone atoms of the thermophilic (red) and mesophilic proteins (blue) on different time scales (from 20 ps to 10 ns). We have computed the errors bars with the bootstrapping statistical method, which implies computing the MSD from randomly generated chunks of the trajectory.

different time scales would show a different thermal behavior of intramolecular dynamics, which is a key aspect for studying the basis of protein thermostability. It is worth mentioning that $d\text{MSD}/dT$ is a quantity routinely determined by elastic neutron scattering experiments on the picosecond time scale, and thus information on the dynamics on longer time scales would be an excellent complement.

The lower tendency of the thermophilic enzyme of increasing its mobility with temperature can have important implications for thermostability, since this can protect the structure from undergoing conformational motions related to unfolding events. Zaccai and co-workers, on the basis of neutron scattering experiments¹¹ on the 100 ps time scale, had already proposed this idea as the dynamical basis underlying the higher thermal stability of thermophilic proteins. Although in our previous study¹⁶ we showed that a contribution of diffusion to the observed overall dynamics plays a role in interpreting their data, here we support the fundamental essence of their idea. Below, we address some of the reasons for such a difference in $d\text{MSD}/dT$ between the two proteins. It is intriguing that the different $d\text{MSD}/dT$ observed for both proteins on the picosecond to nanosecond time scale apparently indicates a fingerprint of thermostability, taking into account that the unfolding process, which is a rare event, takes place on much longer time scales.

Dynamical Heterogeneity. Does the higher flexibility of the thermophilic protein apply to all residues or some specific regions? Of course, the protein surface tends to be more mobile than those regions buried in the protein core, which is subjected to larger spatial constraints. This makes the protein structure dynamically heterogeneous. Is the dynamical heterogeneity of the thermophilic protein different from that of the mesophilic homologue? To gain insight into this matter, now we turn our attention to the contribution of residues to the overall flexibility.

Figure 3 shows the MSD of each residue (averaged over the four chains) of the thermophilic and mesophilic proteins (panel a). The most flexible parts of the two proteins correspond to surface loops, turns connecting secondary structure elements,

and chain extremes, whereas the most rigid residues are located at the protein core. In particular, the N and C termini exhibit much higher mobilities in the mesophilic protein (see arrows in Figure 3a) in support of the general observation that chain extremes are better anchored in thermophilic homologues.¹ Figure 3c and d illustrate this heterogeneous distribution of residue mobilities with a color-coded diagram of both protein structures. Despite the lower overall flexibility of the mesophilic protein, some of its regions exhibit very high mobilities not accessible by the thermophilic homologue, which is indicative of a different distribution of residues' flexibilities, as will be shown below.

To compare the flexibility of structurally analogous regions of the two proteins, we have structurally aligned both protein structures with DALI.²⁸ Figure 3b shows the MSD of aligned residues. The mobility of the thermophilic protein is higher in all aligned residues except for the C terminal tail. Note that the whole N terminal part of the mesophilic protein and the loop defined by residues 217–225, which are very flexible (see Figure 3a), have not been aligned and thus are not shown in Figure 3b as high maxima. It is remarkable that most rigid regions of the thermophilic protein (local minima in the MSD profile) exhibit a higher mobility than the most rigid residues of the mesophilic homologue. Figure 4a clearly illustrates the broader distribution of MSD values in the mesophilic protein (blue curve) with respect to the thermophilic one (red curve). The inset of Figure 4a shows that very high MSD values are only adopted by the mesophilic protein.

The different dynamical heterogeneity of both proteins has consequences in the interpretation of neutron scattering data. In particular, elastic incoherent neutron scattering (EINS) provides a measure of the MSD of the sample by using the Gaussian approximation, which assumes that all atomic displacements are equal, ignoring dynamical heterogeneity.²⁹ Therefore, when comparing the dynamics of two proteins, their different dynamical heterogeneity will introduce different errors in the MSD obtained from the Gaussian approximation. Smith and co-workers³⁰ suggested a solution to this problem by modeling a Weibull distribution of atomic displacements when comparing the dynamics of a thermo-mesophilic pair of dihydrofolate reductases. By performing EINS experiments, the authors found a broader distribution of MSD mobilities in the thermophilic protein, in contrast to our results. To gain insight into this discrepancy, we have fitted a Weibull function to each distribution (see dashed lines in Figure 4a) and, subsequently, calculated the standard deviation (σ) associated with the fitted functions (see the Supporting Information for details in this calculation). This analysis reveals that the σ of the fit to the mesophile's distribution is now lower than that of the thermophile (0.24 vs 0.32 Å²), which is in contrast with the results obtained with the original data (0.75 vs 0.47 Å²). Note in Figure 4a the larger width of the fitted function of the thermophile. This apparent contradiction is reconciled by the fact that the tail of the Weibull function fitted to the mesophile's distribution does not include the most flexible residues, which contribute to increasing σ substantially. The inability of the Weibull function to include the most flexible residues therefore masks the heterogeneous character of the distribution. Thus, we can argue that our observation of a higher dynamical heterogeneity in the mesophilic protein is fully consistent with the results obtained by Smith and co-workers.³⁰

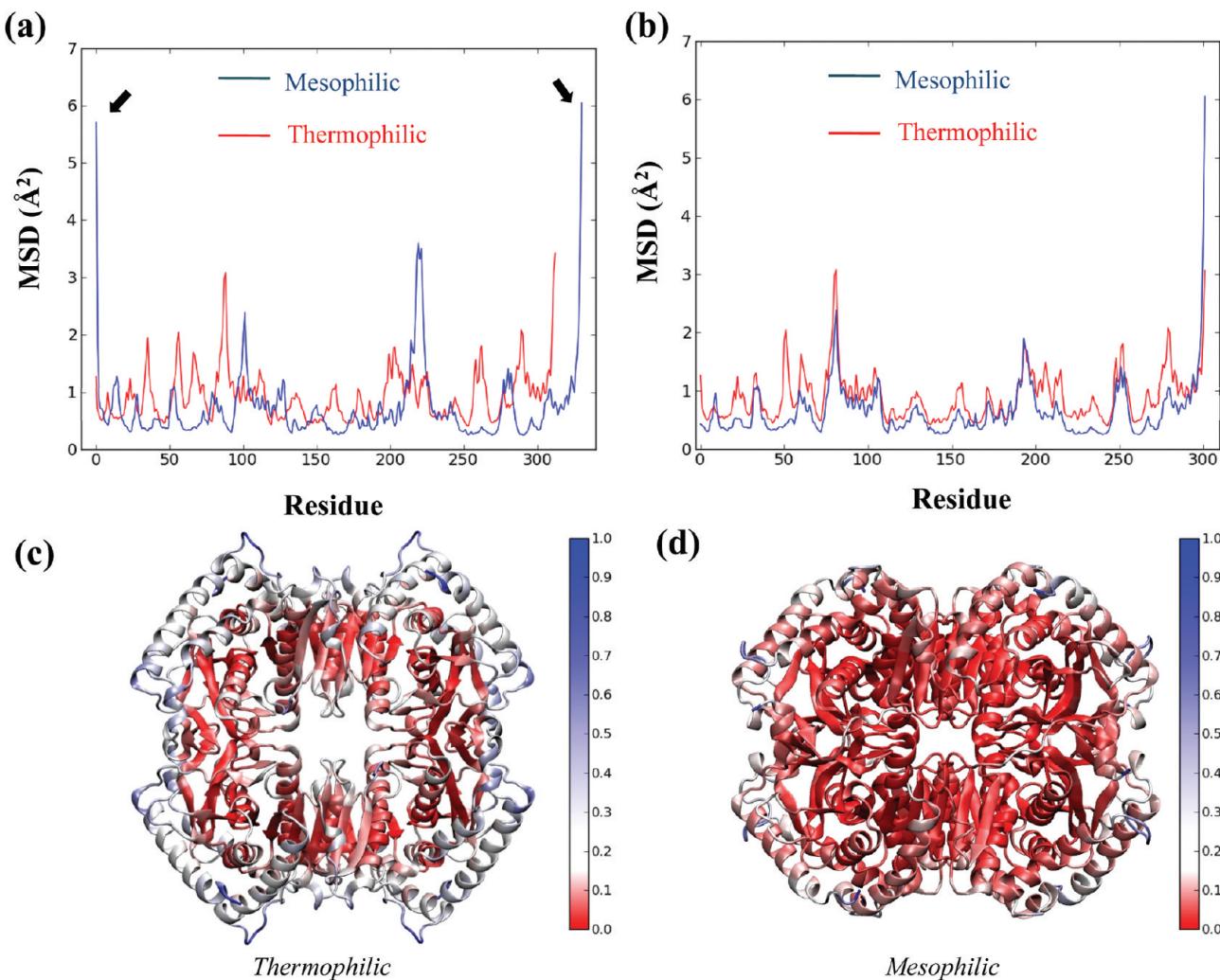


Figure 3. Residue mobilities at 280 K and on the 1 ns time scale. (a) MSD of each residue of the thermophilic and mesophilic proteins. (b) MSD of structurally aligned residues. (c, d) Ribbon diagrams of both proteins that are color-coded according to the MSD value of each residue (normalized by the highest MSD value among the two proteins). From red to blue, the MSD increases as indicated by color bars. The MSD of each residue has been computed as an average over the MSD of all backbone atoms of the same residue.

We have quantified this dispersion of MSD values with the standard deviation (σ) over all residues on different time scales and at different temperatures (see Figure 4b). The figure shows that σ increases with the time scale and temperature in both proteins, but this increase is more accentuated for the mesophilic protein. Figure 4b also shows that σ is higher for the mesophilic protein than for the thermophilic one in the range of time scales studied. In addition, σ increases with the temperature to a higher extent. We wonder whether the dispersion in residue mobilities can be regarded as a dynamic fingerprint distinguishing thermophilic proteins from mesophilic ones. If so, a MD simulation at a single temperature would find wide use in distinguishing proteins with different thermal adaptation.

How Correlated Is the Dynamics on Different Time Scales and at Different Temperatures? Given that some experimental techniques only have access to short time scales, such as neutron scattering, one wonders whether the conformational dynamics on short time scales (ps) correlate with that on longer time scales (ns or beyond), which involve different types of motion. We have already seen that the overall flexibility increases with time scale and temperature. Do all

residues increase their mobilities with time scale and temperature at the same rate? How do residue mobilities correlate between different time scales and temperatures? In Figure 5, we have analyzed these issues. Figures 5a and S1 show that the shape of the MSD profile is well conserved among different time scales. Figure 5b illustrates this good correlation by comparing the MSDs between two different time scales differing by 3 orders of magnitude (20 ps and 10 ns). Such a smooth increase in flexibility with time scale, however, hides the fact that the increase in flexibility is not the same for all residues. For example, at 20 ps, regions a and b (see arrows in Figure 5a) have similar flexibilities, whereas on much longer time scales (10 ns) region b becomes far more flexible than a. This can be clearly seen when normalizing the flexibilities on different time scales (see Figure S2). Note that the thermophilic protein has a more conserved pattern than the mesophilic one. This implies that in the latter the rates of MSD increase with time scale are more residue-dependent or *heterogeneous* throughout the structure. This can be viewed as an alternative manifestation of the higher dynamical heterogeneity of the mesophilic protein. It is thus expected that these more heterogeneous rates of MSD increase in the mesophilic

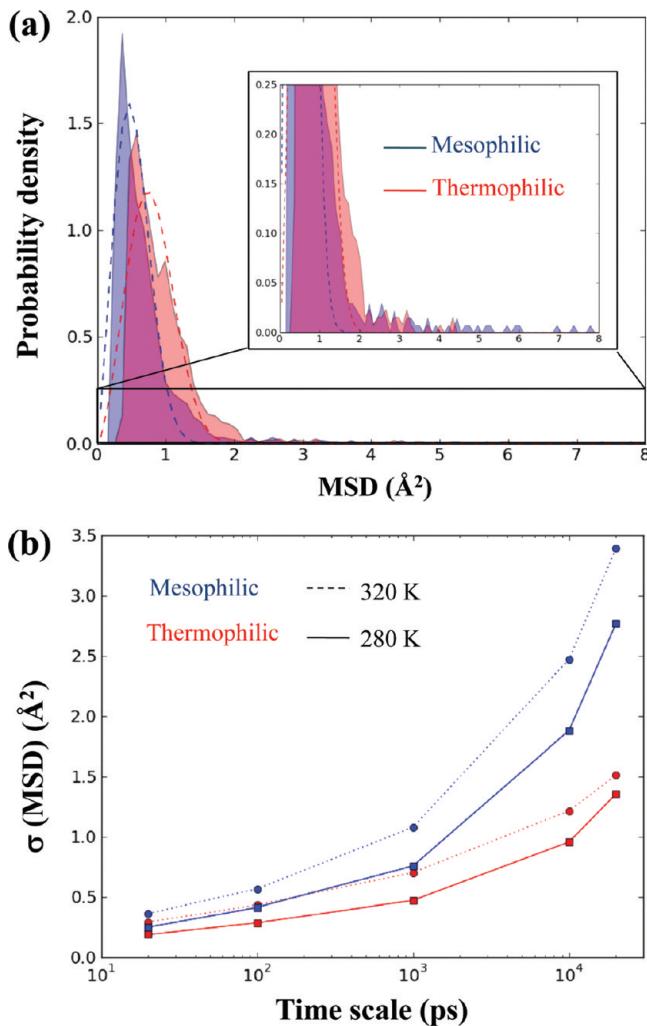


Figure 4. Dynamical heterogeneity of the thermophilic (red) and mesophilic (blue) proteins. (a) Normalized histograms of residue mobility at 1 ns and 280 K. The area common to both proteins is colored in violet to highlight the differences in their distributions. The inset shows the population of the highest MSD values. Dashed lines correspond to Weibull fits. (b) Standard deviation (σ) of MSD residues on different time scales and at different temperatures (280 K, solid lines; 320 K, dotted lines).

protein result in distributions of residues flexibilities with a steeper increase of σ upon raising the time scale (Figure 4b). The conserved pattern of the thermophile can be explained by its higher collectivity of protein motions, as will be described below. Therefore, although the general trend is that flexible regions remain flexible and fixed regions remain fixed, the relative flexibilities among residues on different time scales do change. This has also been reported from NMR experiments⁷ that compared two different time scales. Again, this emphasizes the crux of this paper: that talking about dynamics without specifying a time scale can result in contradictory conclusions.

Figure 5b also illustrates how the MSD of the mesophilic protein (in blue) at 320 K and a time scale of 10 ns exhibits a sharper increase, with respect to 20 ps, than the thermophilic homologue. The sharper increase in MSD of the most flexible residues of the mesophilic protein compensates for the lower mobility of its more rigid residues, resulting in an overall flexibility fairly similar to that of the thermophilic protein. This is a complementary perspective of the observation from Figure

1 that the overall MSD of the mesophilic protein at this temperature increases more rapidly with the time scale to the extent that at ~ 5 ns it reaches the MSD values of the thermophilic one. The corresponding representations at the other temperatures are shown in Figure S3.

In Figure 2, we have given support to the thermostability mechanism of thermophilic proteins suggested by Zaccaï and co-workers,¹¹ which is based on a reduced temperature dependence of the overall flexibility. To gain insight into the origin of our results, here we aim to deconvolute the contribution of residues to this overall behavior of the protein. With regard to the MSD variation with temperature, we have compared the residue flexibilities between different temperatures on a time scale of 10 ns. We note in Figures 5C and S4 that all residues of both proteins, except for residues 50–58 and the C-terminal tail of the thermophilic protein (see arrows in Figure 5C), tend to increase the MSD with the temperature. Figure 5d shows that this rate of increase is fairly similar among all residues and that the slope of this linear trend is higher for the mesophilic protein, in line with Figure 2, which provides evidence of its larger $d\text{MSD}/dT$. This figure shows that all residues contribute to increasing $d\text{MSD}/dT$. To better illustrate the temperature dependence of the MSD of each residue, we have represented this quantity in Figure 6 as the slope of a linear fit of MSD (at 10 ns) vs temperature. Now, we turn our attention to the factors that determine differences in the $d\text{MSD}/dT$ of residues in the two proteins.

Structural Differences between the Mesophilic and Thermophilic Homologues. Figure 6b shows the $d\text{MSD}/dT$ values shown in Figure 6a, but only for structurally aligned residues. Figure 6b shows that the $d\text{MSD}/dT$ values of the mesophilic protein are, in general, above those of the thermophilic homologue. As noted above, even some residues of the thermophilic protein (50–58 and the C-terminal tail) undergo a decrease in MSD with the temperature, in contrast to the mesophilic protein, the residues of which all have a positive $d\text{MSD}/dT$. However, it is worth emphasizing that some regions display very close similarity between the two proteins. To gain insight into the origin of these similarities and differences in the $d\text{MSD}/dT$ of both proteins, it is instructive to evaluate different contributions to the average $d\text{MSD}/dT$, as shown in Table 1. Herein, those residues that do not form helices or β sheets are termed coils, including loops, turns, or random coils.

The average of $d\text{MSD}/dT$ over structurally aligned residues, $\langle d\text{MSD}/dT \rangle_{\text{align}}$, is still higher for the mesophilic protein (0.0130 vs $0.0082 \text{ Å}^2 \text{ K}^{-1}$). However, it is interesting to note that the difference between both proteins is enlarged when comparing the average over unaligned residues, $\langle d\text{MSD}/dT \rangle_{\text{not align}}$ (0.0248 vs $0.0069 \text{ Å}^2 \text{ K}^{-1}$). This separation of contributions reveals that the flexibility of structural features unique to the mesophilic protein is much more sensitive to changes in temperature. In particular, a 72.4% of not aligned residues of the mesophilic protein form coils. In line with this, we show in Table 1 that the $d\text{MSD}/dT$ of all coil regions, $\langle d\text{MSD}/dT \rangle_{(L)}$, of the mesophilic protein is notably larger than that of helices and β sheets, $\langle d\text{MSD}/dT \rangle_{(H/E)}$ (0.0202 vs $0.0114 \text{ Å}^2 \text{ K}^{-1}$). In contrast to this, the $d\text{MSD}/dT$ of the thermophilic protein proves to be practically insensitive to the secondary structure. Therefore, an important structural reason underlies the higher $d\text{MSD}/dT$ of the mesophilic protein. Indeed, it is widely accepted that thermophilic proteins tend to exhibit shortened surface loops to lower the entropy change of

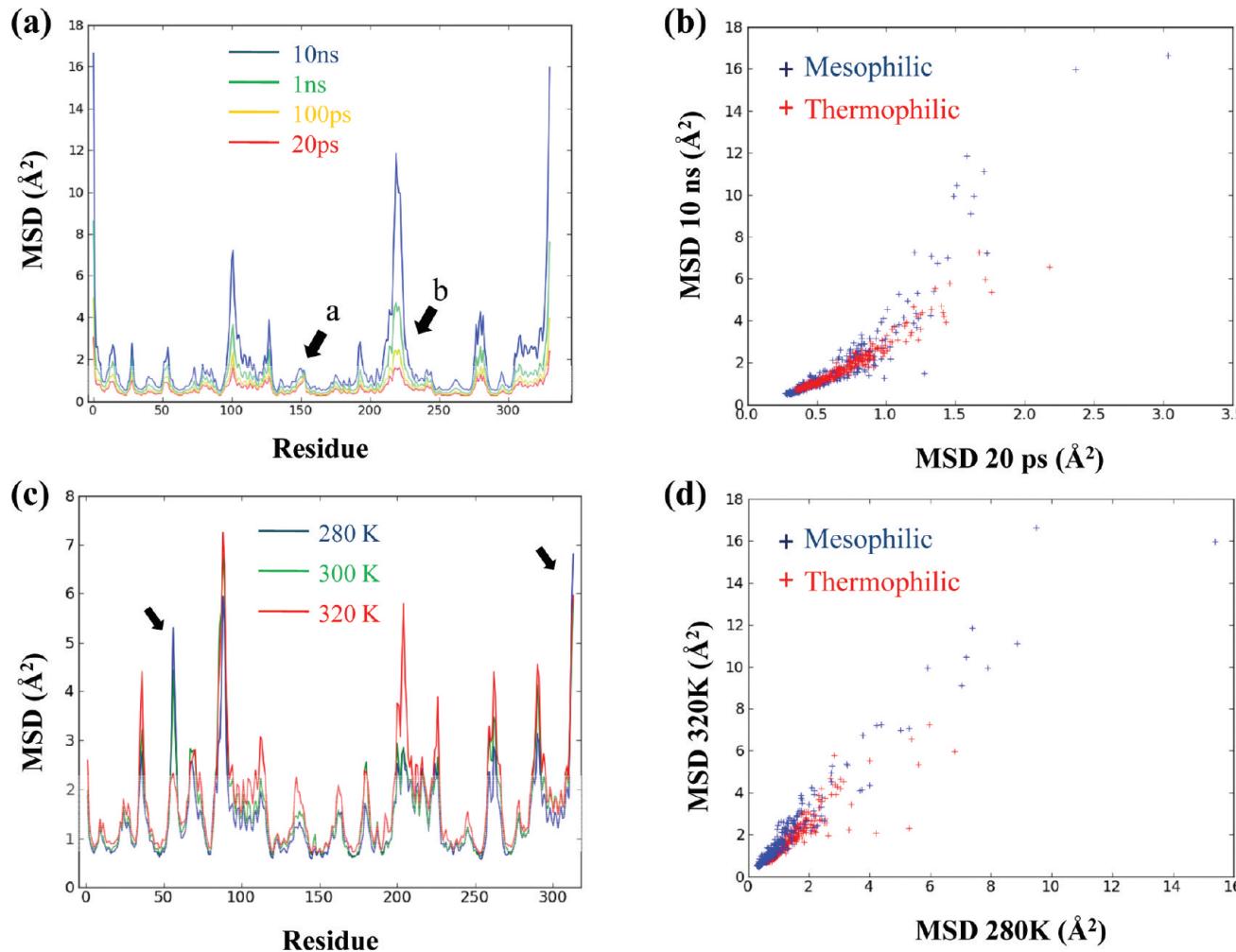


Figure 5. Flexibility on different time scales and temperatures. (a) MSD of the mesophilic protein at 320 K and time scales: 20 ps (red), 100 ps (yellow), 1 ns (green), and 10 ns (blue). (b) Correlation between the MSDs at 20 ps and 10 ns of the thermophilic (red crosses) and mesophilic (blue crosses) proteins at 320 K. (c) MSD (10 ns) of the thermophilic protein at 280 (blue), 300 (green), and 320 K (red). (d) Correlation between the MSDs (10 ns) at 280 and 320 K for the thermophilic (red crosses) and mesophilic (blue crosses) proteins.

unfolding,³¹ which provides higher stability. Here, we show that this structural feature has implications on the sensitivity of protein flexibility to changes in the temperature.

A related structural feature that has been associated with enhanced thermostability is an improved packing.^{32,33} To further explore this idea, we have quantified the packing density of both proteins computing an index that averages the number of contacts at each residue (d_i), as introduced by Chennubhotla and Bahar (see Methods). First, we have averaged the density of all residues ($1/N \sum_{i=1}^N d_i$) in the crystal structures of both proteins. It turns out that the degree of packing is higher in the thermophilic protein (6.8 vs 6.3). To characterize the distribution of packing across the protein structure, Figure 7 displays the density averaged over residues at different shells of the protein structure (dashed lines). We have divided the protein structure into five shells (s_i , $i = 1-5$) as a function of distance to the center of mass with a 10 Å width. For instance, s_1 refers to residues at a distance ranging from 0 to 10 Å from the center of mass. It is worth reminding the reader that in the first 10 Å shell of the thermophilic protein there are no residues due to the large central hole (see Figure 3c).

Figure 7 illustrates that the crystal structure of the thermophilic protein is more compact than that of the

mesophilic homologue near the surface. This is consistent with the larger content of residues forming coils, which have limited intraprotein contacts, at the surface of the mesophilic protein. Nevertheless, such differences in density between both proteins disappear (6.13 ± 0.03 and 6.16 ± 0.04) when considering several snapshots of the MD simulations. From these results, we cannot discern a relationship between packing and the different flexibility observed for the two proteins. Interestingly, the density in both cases drops from the values obtained for the crystal structures as a result of the relaxation of the structure when immersed into the solvent. This indicates that protein packing is sensitive to the environment, and thus an analysis only focused on the X-ray structure is insufficient. However we could not find a definitive explanation for why the difference in densities disappears.

Salt Bridge Distribution and Solvent Accessibility. It is important to remark that the flexibility of structurally analogous regions is still significantly more temperature-dependent in the mesophilic protein. Thus, there must be other factors that determine the $d\text{MSD}/dT$ in these regions. We now examine how the larger proportion of charged residues generally observed in thermophilic proteins can contribute to lowering $d\text{MSD}/dT$ by forming salt bridge interactions. Previously, we

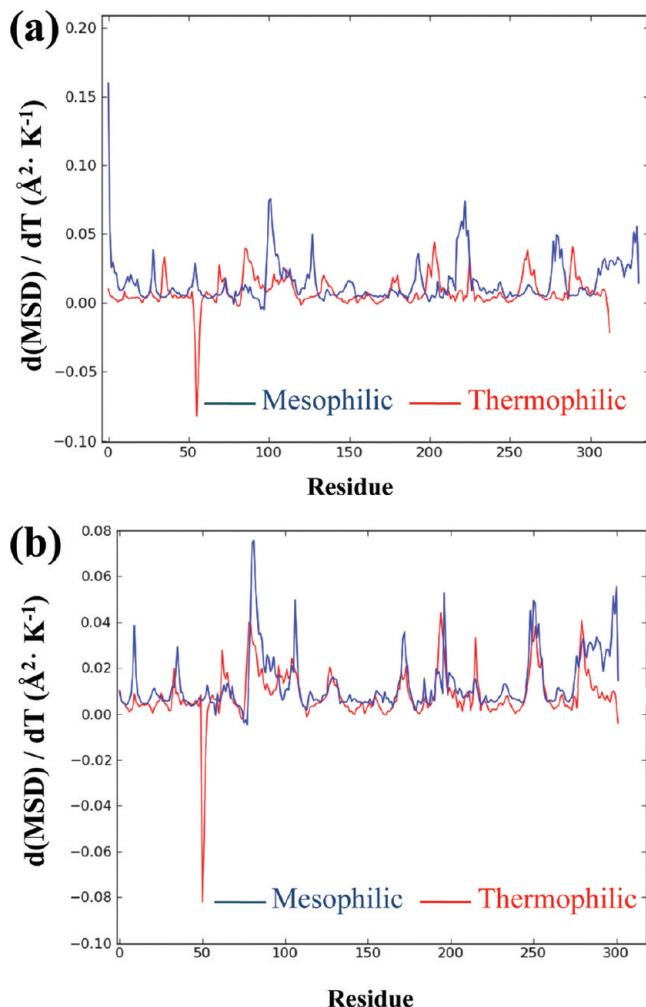


Figure 6. (a) Representation of the $d(\text{MSD})/dT$ values computed for the backbone atoms of each residue of the thermophilic (red) and mesophilic (blue) proteins on the 10 ns time scales. (b) Same representation as panel a, but only for structurally aligned residues.

Table 1. Contributions to the Temperature Dependence of MSD on the 10 ns Time Scale

	thermophilic	mesophilic
$\langle d(\text{MSD})/dT \rangle^a$	0.0081	0.0147
$\langle d(\text{MSD})/dT \rangle_{\text{align}}^b$	0.0082	0.0130
$\langle d(\text{MSD})/dT \rangle_{\text{not align}}^c$	0.0069	0.0248
$\langle d(\text{MSD})/dT \rangle_{(L)}^d$	0.0082	0.0202
$\langle d(\text{MSD})/dT \rangle_{(H/E)}^e$	0.0081	0.0114
% of aligned residues forming coils	34.8	39.4
% of aligned residues forming helices/sheets	65.2	60.6
% of not aligned residues forming coils	63.6	72.4
% of not aligned residues forming helices/sheets	36.4	27.6

^aAverage over all residues. ^bAverage over structurally aligned residues. ^cAverage over not structurally aligned residues. ^dAverage over residues forming coils (L). ^eAverage over residues forming helices (H) or sheets (E).

have shown in Figure 6b that a striking difference in the thermal behavior of both proteins lies in residues 50–58 and the C-terminal tail of the thermophilic protein, which undergo a decrease in MSD with the temperature. To rationalize the origin of this amazing behavior, we have examined all salt bridges formed by charged residues of these two regions. In

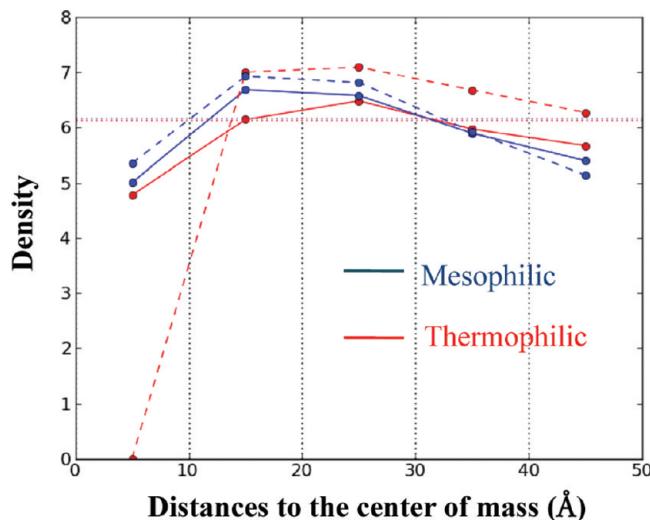


Figure 7. Density of residues at different shells of the crystal structures (dashed lines) and MD snapshots (solid lines) of the thermophilic (red) and mesophilic (blue) proteins. For each protein, we have averaged the density over 15 snapshots (five separated by 50 ns at each temperature). Horizontal dashed lines represent the average (over all residues) of the data series from MD snapshots. The last shell corresponds to distances greater than 45 Å.

particular, we determined the probability density of the distance of all salt bridges formed between a given residue and oppositely charged residues within 20 Å. Figure 8 shows the corresponding probability densities for residues R56 and K311, which is located at the C-terminal tail. It is evident from the figure that both residues exhibit an increase in the probability of forming salt bridges at short distances upon raising the temperature. In particular, the probability of finding a short salt bridge (with distance ≤ 4.5 Å), or *occupancy*, is found to increase 25% and 23% in R56 and K311, respectively, when raising the temperature from 280 to 320 K. It is noteworthy that, in the case of residue R56, increasing the temperature from 280 to 320 K entails the formation of new salt bridges with residue R56, since at low temperatures no short salt bridges are formed, as shown in Figure 8a (blue curve). Therefore, either the formation of new salt bridges or the strengthening of already present salt bridges with temperature is likely to constrain the motion of the specific residue and that of neighboring residues. Taking into account that these effects are induced by increases in temperature, the negative $d(\text{MSD})/dT$ observed in Figure 6 for these regions is consistent with this idea. Indeed, such tightening of salt bridges upon raising the temperature has already been pointed out by others as a mechanism for thermal stability.^{34–37} As the temperature increase, the dielectric constant of water decreases, leading to a reduction of the desolvation penalty of salt bridge formation that ultimately strengthens this interaction and the robustness of the structure. Our present observation is fully consistent with recent NMR studies by Vinther et al. on another thermostable protein.³⁷ They revealed that, on the pico- to nanosecond time scale, the backbone flexibility in different areas of the protein decreases with the temperature as a result of salt bridge tightening, which ultimately contributes to increasing protein stability.

In Figure 9a, we show that the content in charged residues follows a similar trend as that of packing at different shells of the crystal structure. The main difference in the content of

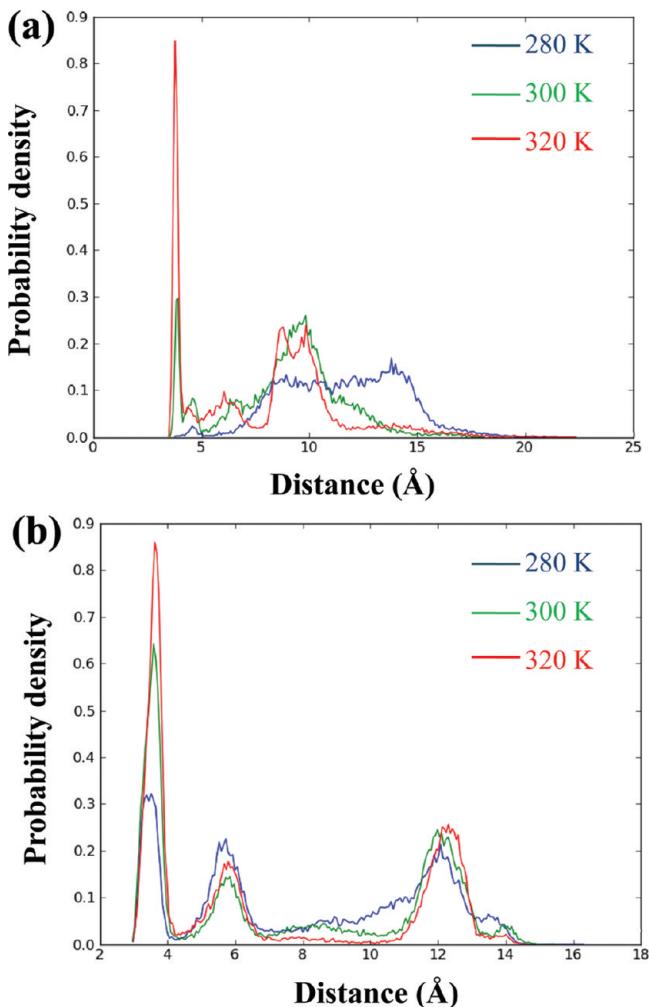


Figure 8. Probability distributions of all salt bridges integrating residues R56 (panel a) and K311 (panel b) of the thermophilic protein at 280 (blue), 300 (green), and 320 K (red). The occupancy of the salt bridge is the integral of these distributions from 0 to 4.5 Å. The salt bridge distance is defined between the CZ atom of Arg (and NZ atom of Lys) and the CG atom of Asp (and CD atom of Glu).

charged residues lies again near the surface, as commonly accepted.³ While the mesophilic protein presents a rather constant proportion of charged residues throughout the structure, the thermophilic one steadily increases this proportion as the distance to the surface is reduced. For instance, charged residues at the surface of the thermophilic protein represent 55%, whereas it is only 32% in the mesophilic homologue. Regarding the content on charged residues, it is interesting to note that the thermophilic protein presents another signature of enhanced thermostability, that is, the higher arginine-to-lysine ratio^{1,38,39} relative to the mesophilic homologue (0.66 vs 0.42).

A clear manifestation of salt bridge tightening with temperature at the surface of the thermophilic protein is the thermal behavior of the radius of gyration (R_g), which is sensitive to fluctuations of the surface. Figure 9b shows the time evolution of R_g at different temperatures. While the radius of gyration of the mesophilic protein displays modest variations with the temperature, the thermophilic protein shows a significant decrease of R_g upon increasing the temperature. In particular, the central hole of the thermophilic protein

undergoes a contraction that entails a significant decrease in the solvent accessible surface (SAS) of residues within 20 Å from the center of mass (second shell), as shown in Figure S5 (panel a). In contrast, the SAS of the mesophilic protein (within this 20 Å shell) is much less sensitive to temperature variations (see panel b in Figure S5). This implies that upon increasing temperature, the thermophilic protein has a tendency to disrupt protein–water interactions to enhance intraprotein interactions, as opposed to the mesophilic homologue. Despite such reduction in the SAS of the thermophilic protein, it is worth noting that it still remains largely above that of the mesophilic protein. The larger SAS arises from a wider central hole that is filled with water. Indeed, it is widely accepted that thermophilic proteins tend to have a more accentuated hydrophilic character,⁴⁰ which is intimately linked to the larger proportion of charged residues at the surface, as shown above.

Globular proteins have solid-like interiors and liquid-like surfaces, and this can be quantified by the Lindemann coefficient (Δ_L).⁴¹ The larger SAS of the thermophilic protein should imply that a larger proportion of residues are liquid-like, resulting in a more liquid-like overall protein. We have computed Lindemann coefficients (from a normal-mode analysis using FlexServ⁴²) of the crystal structures of the two proteins and also of several snapshots of the MD simulations (see Table S1). The Δ_L values obtained for the crystal structures (and six snapshots of the MD simulations) of the thermophilic and mesophilic proteins are 0.131 (0.138 ± 0.002) and 0.118 (0.121 ± 0.001), respectively. This supports the higher liquid-like character of the thermophilic protein. In addition, it is interesting to note that the relaxation of the protein structure upon running the simulation entails an increase in the liquid-like behavior of both proteins, this increase being more important in the more thermostable protein. It is worth reminding the reader that the Lindemann coefficient is calculated from a normal-mode analysis of the structure, and not from the SAS. Therefore, the more solvent-exposed structure of the thermophilic protein results in both this higher liquid-like behavior and a larger SAS.

Figure 9c shows that the thermophilic protein systematically experiences salt bridge tightening in each shell but that the major tightening occurs in the second shell. In particular, we note that salt bridges involving R56, which is located within this shell and has been highlighted in Figure 8a, are those that contribute the most to tightening all salt bridges within this shell. It is this shell which builds the surface of the central hole and thus the aforementioned shrinkage of the hole just arises from the salt bridges tightened in this region. On the contrary, Figure 9d shows that the salt bridge tightening with temperature undergone by the mesophilic protein is less systematic in the different shells. In some cases, an increase in temperature does not necessarily involve an increase in the occupancy of salt bridges. The larger proportion of charged residues (see Figure 9a) underscores the importance of the more efficient tightening of salt bridges upon increasing the temperature throughout the thermophilic structure. Although we have identified two regions with a negative $dMSD/dT$, this does not exclude the possibility that the motion of other regions is constrained due to an increase in temperature. We suggest that salt bridge tightening in other protein regions would partially counterbalance the natural increase in mobility with temperature.

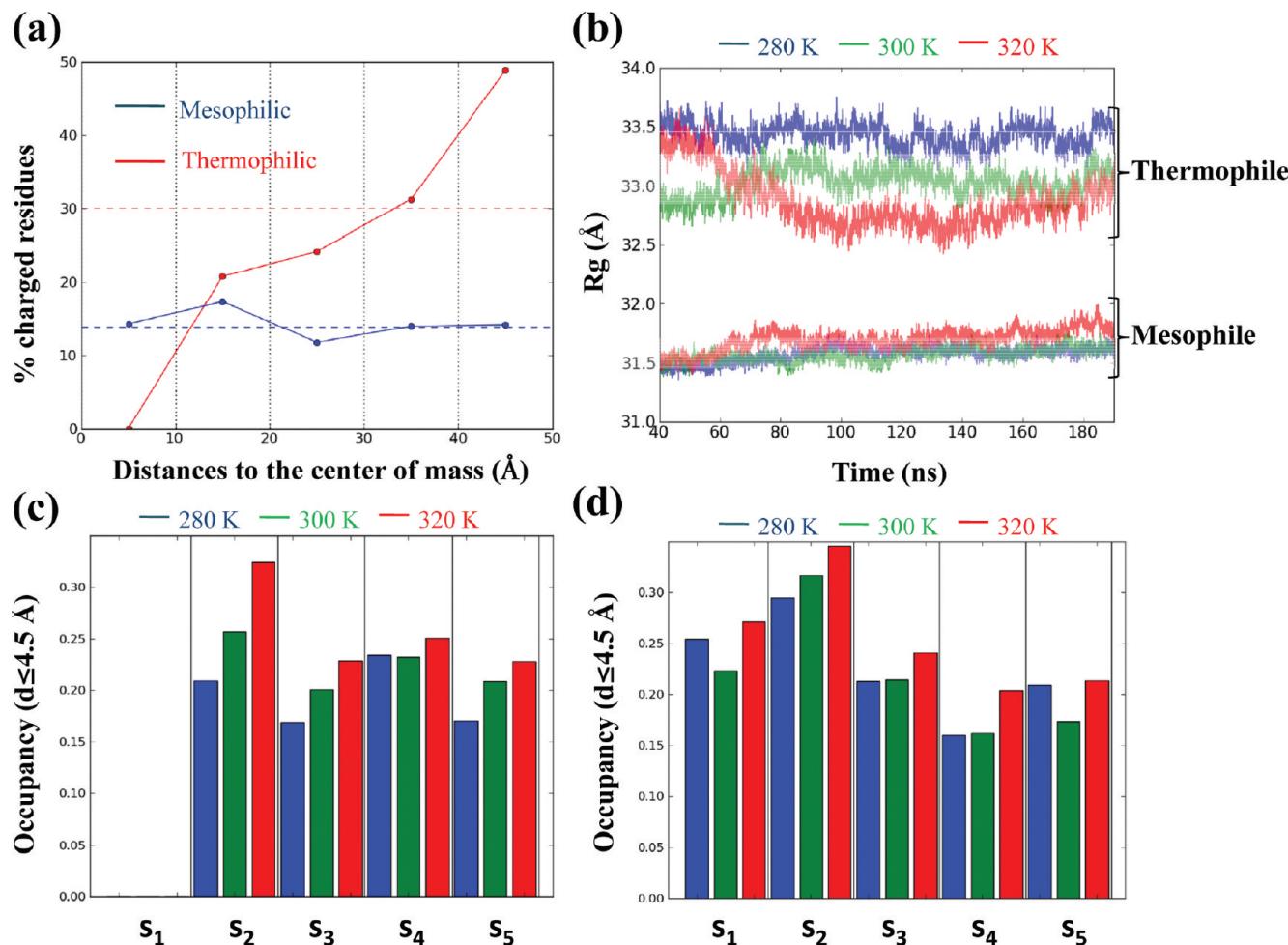


Figure 9. (a) Proportion of charged residues at different shells of the protein structure. Vertical dotted lines denote the range of distances defining a shell. Horizontal dashed lines represent the average of each data series over the five shells considered. The last shell corresponds to distances greater than 45 Å. (b) Radius of gyration as a function of time at the three temperatures studied (280 K, blue; 300 K, green; 320 K, red). (c) Average occupancy of all salt bridges of the thermophilic protein within each shell s_i at the three temperatures (same color code as in panel b). The protein structure is divided into five shells of 10 Å width as in Figure 7. (d) Same as panel c for the mesophilic protein.

It is pertinent to note that the radius of gyration undergoes slow fluctuations, which is indicative of a slowly convergent property. This makes monitoring this property at long time scales necessary to allow meaningful interpretation. For instance, from the 40–50 ns interval, it would seem that the radius of gyration is converged (see Figure S6), whereas the full 200 ns run reveals that the fluctuating R_g is unlikely to be converged, and thus, much longer trajectories are required.

In this work, we provide evidence of two factors determining the thermal behavior of MSD: structure and salt bridge interactions. Although both factors play a primary role on the picosecond to nanosecond time scale here explored, this does not rule out the contribution of other factors underlying the difference in $d\text{MSD}/dT$ between both proteins. Future work examining longer time scales is expected to provide new insight into their different sensitivity of protein motions to temperature changes.

Principal Component Analysis. We have carried out a principal component analysis of the trajectories at each temperature to give insight into differences in collective motions between both proteins. We have projected the trajectories obtained at each temperature into the first two PCs, which describe a larger amount of variance. Figure 10

represents the contour plots of the probability density of projections. It is evident from the figure that the conformational fluctuations of the two proteins are distributed in a strikingly different fashion along the first two PCs. At a given temperature, the thermophilic protein fluctuates to a higher extent than the mesophilic one, when moving either within the same minimum or between minima of the conformational space defined by the first two PCs. Of course, the larger fluctuations observed in the thermophilic protein system are consistent with the higher flexibility on short time scales, as pointed out above. The idea that the largest conformational transitions occur through *tight* pathways in the mesophilic protein is appealing, while in the thermophilic homologue *broader* routes are allowed. This ultimately provides insight into a significant difference in topology of the energy landscape of both types of proteins.

Taking into account such a difference in the dispersion of conformational fluctuations of the two systems along the first two PCs (Figure 11), we examined how this might be linked to the difference in dynamical heterogeneity highlighted above. We scrutinized this possibility with the collectivity²⁷ of each principal component (see Methods) to give a measure of how

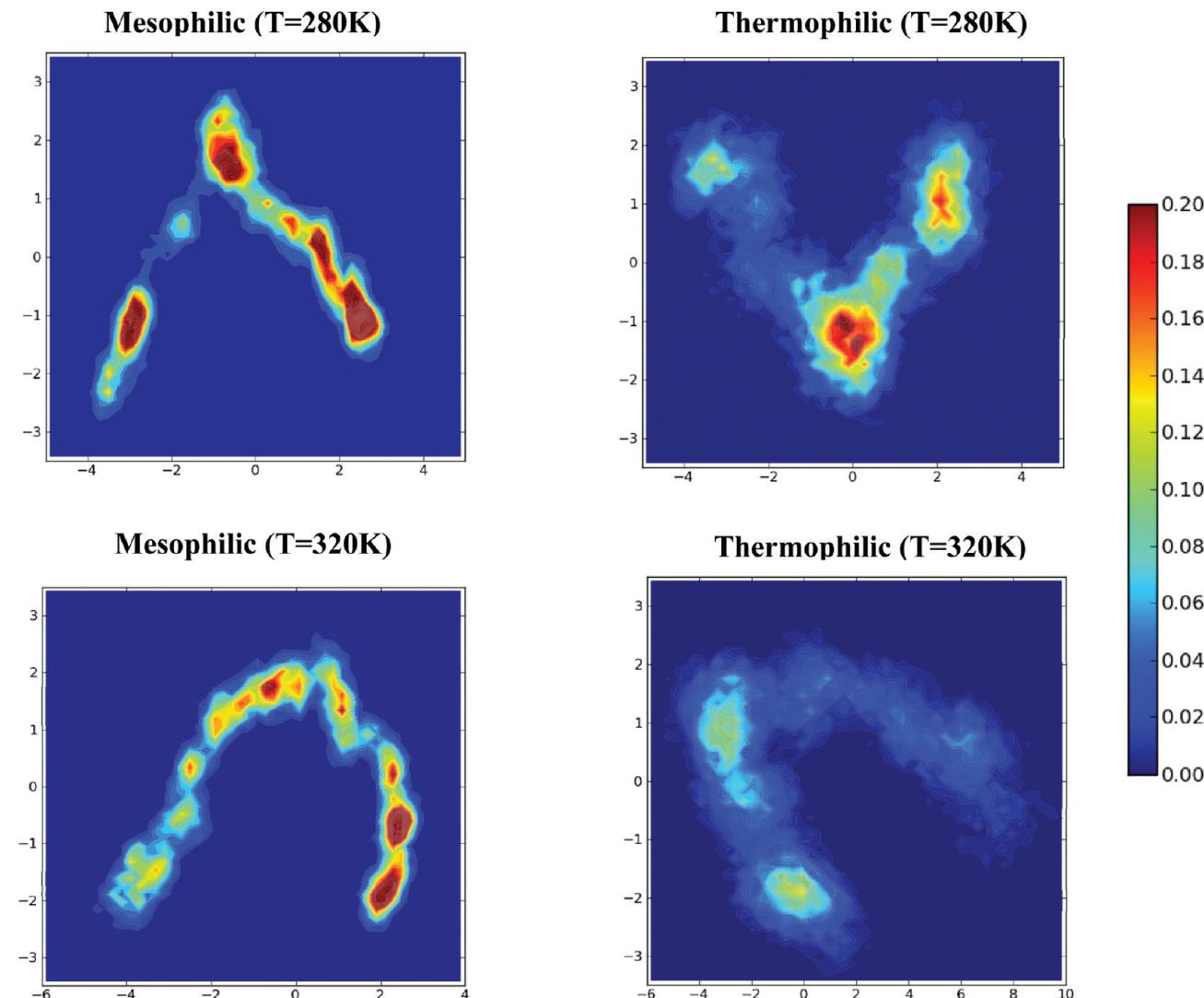


Figure 10. Contour plots of the projection of trajectories into the first two principal components at 280 and 320 K.

many residues contribute to a given mode and thus characterize the dynamical heterogeneity associated with each mode.

Figure 11 displays the collectivity indices of the first 100 principal components at 280 (panel a) and 320 K (panel b). Interestingly, in the first PCs, the collectivity of the mesophilic protein is notably lower than that of the thermophilic counterpart, the difference in collectivity of the first two PCs addressed above being remarkable. These results indicate that the main conformational fluctuations of the mesophilic protein have a more local character than in the thermophilic one, which presents more cooperative motions. Put more succinctly, the higher dynamical heterogeneity of the mesophilic protein suggested above is clearly manifested here by the low collectivity of the first principal components. Given the marked difference in the cooperative nature of motions of the two proteins, we wonder whether this might be related to the different topology of their energy landscapes, as illustrated in Figure 10. Does the broader energy landscape of the thermophilic protein describe more collective motions? There is an apparent correlation when considering the first two PCs, but to explore further this idea we examined the variance of the projection of trajectories into each of the first 100 principal components (see the Methods for details on the calculation of

this variance). Figure 11 (panels c and d) shows the variance of projections of each PC for the two proteins. It reveals that the variance in the first PCs is systematically higher in the thermophilic protein. Interestingly, the modes in which the thermophilic protein has a higher collectivity than the mesophilic homologue coincide well with those in which the variance is also higher. This strongly suggests that the more cooperative motions of the thermophilic protein emerge from broader energy landscapes. It is of interest to assess whether this feature of the energy landscape can be regarded as a hallmark distinguishing thermophilic proteins from their mesophilic counterparts in general. For a more detailed characterization of the landscape of both proteins, we leave as future work the use of clustering analysis methods in conjunction with Markov models for studying the kinetics associated with the conformational transitions of each protein.

CONCLUSIONS

The present study focuses on differences in the dynamical properties of a thermophilic protein and its mesophilic homologue that can be associated with thermal stability. The analysis of molecular dynamics trajectories at three different temperatures reveals that the overall flexibility (MSD) of the

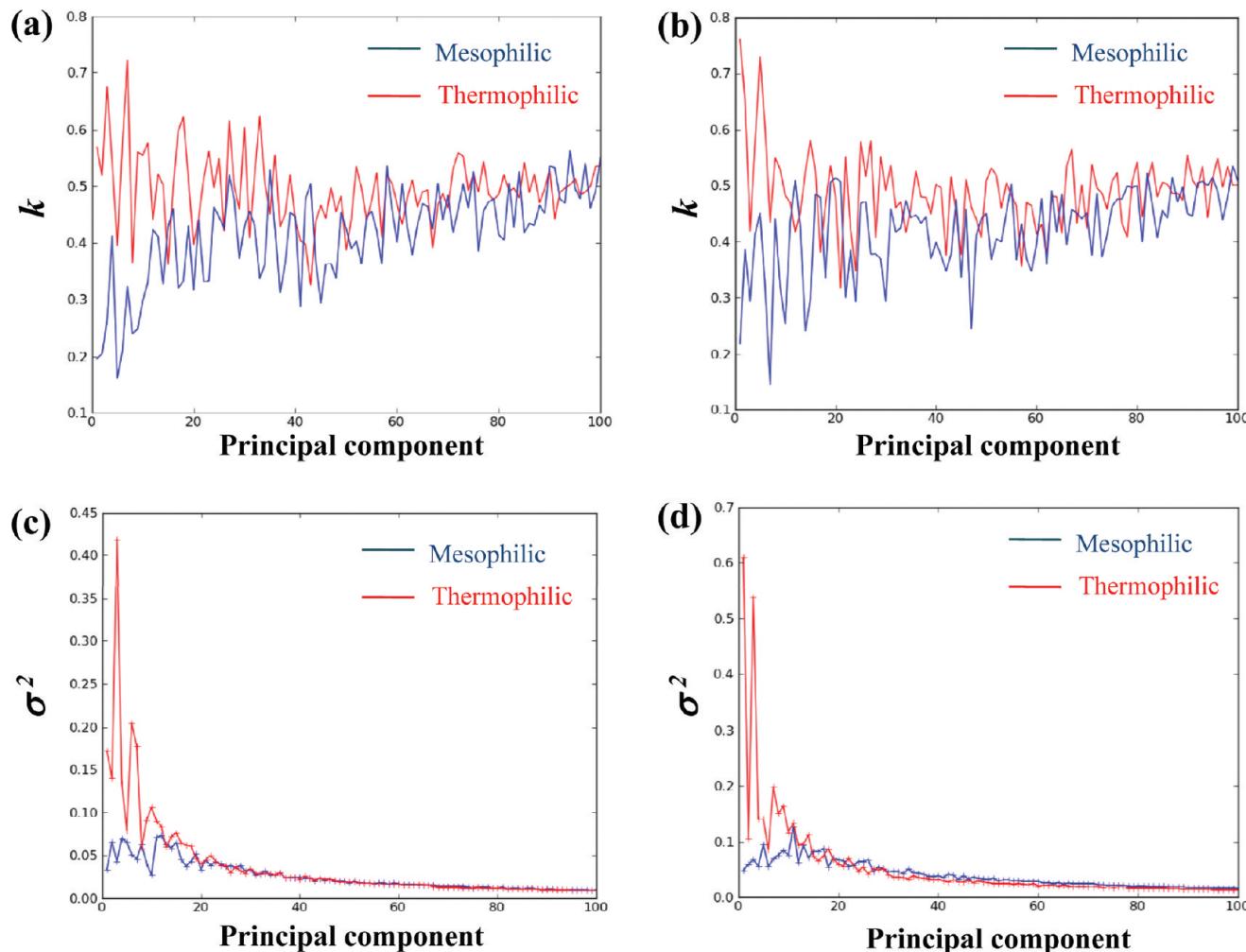


Figure 11. Collectivity of the first 100 principal components at (a) 280 K and (b) 320 K. Variance of the projection of trajectories into the first 100 PCs at (c) 280 K and (d) 320 K. This has been computed with eq 4. Data series for the mesophilic and thermophilic proteins are depicted in blue and red, respectively.

thermophilic protein is higher than that of the mesophilic homologue on the picosecond and few-nanoseconds time scales. However, the flexibility of the mesophilic protein exhibits a higher sensitivity to changes in the temperature. The lower temperature dependence of the MSD of the thermophilic protein lies in the lower content on coil residues (in nonaligned regions) and the larger content on charged residues which mediate salt bridge interactions that tend to strengthen with the temperature. Moreover, the distribution of mobilities among protein residues differs in the two proteins. The mesophilic protein exhibits a broader distribution, which implies a higher dynamical heterogeneity. In this case, loops and N and C termini are very mobile but the core of the protein is more rigid than in the thermophilic protein. Such a difference in the dynamical heterogeneity is reflected in the principal components describing the main conformational fluctuations of the proteins. The modes of the thermophilic protein are characterized by a higher collectivity entailing more cooperative motions, whereas the first modes of the mesophilic homologue tend to have a more local character. Overall, we suggest that the more cooperative motions of the thermophilic protein emerge from broader energy landscapes.

On balance, our findings from simulations on the lower $d\text{MSD}/dT$ of the thermophilic protein support the fundamen-

tal idea originally suggested by Zaccai and co-workers.¹¹ Given the high correlation of the flexibility between time scales differing by 3 orders of magnitude (from 20 ps to 10 ns), it is tempting to regard a low $d\text{MSD}/dT$ as a dynamic fingerprint of enhanced thermostability. Here, we have extended the time scales studied in the experiment, and as future work, we aim to study this contrasting thermal behavior on much longer time scales, which are more relevant to unfolding events. This will be done with coarse-grained models that will be calibrated by the all-atom simulations analyzed here. In addition to this, the present paper suggests two additional dynamic fingerprints distinguishing thermophilic and mesophilic proteins: first, the higher dynamical heterogeneity of the mesophilic protein and, second, the broader energy landscape of the thermophilic protein related to more cooperative motions. It should be emphasized that these predictions need further testing in other thermo-mesophilic pairs of proteins, including pairs having different oligomeric states. This would elucidate the extent to which these properties can be generalized and that they do not merely arise from structural differences that may not be related to thermostability, such as the larger central hole of the thermophilic protein.

■ ASSOCIATED CONTENT

§ Supporting Information

Details on the Weibull function; data on Lindemann coefficients; and plots showing different analyses of the MSD at different time scales, radii of gyration, and solvent accessible surfaces. This information is available free of charge via the Internet at <http://pubs.acs.org/>

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: ramon.crehuet@iqac.csic.es.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful to Miriam Reverter for her assistance in the analysis of trajectories. The authors thank the Galicia Supercomputing Center for computational resources. This work was supported by grants from the JAE programme of CSIC, a CSIC “intramural” project (200980I058), the Spanish MEC (CTQ2009-08223), and the Catalan AGAUR (2005SGR00111).

■ REFERENCES

- (1) Vieille, C.; Zeikus, G. J. Hyperthermophilic enzymes: Sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.* **2001**, *65*, 1.
- (2) Razvi, A.; Scholtz, J. M. Lessons in stability from thermophilic proteins. *Protein Sci.* **2006**, *15*, 1569.
- (3) Kumar, S.; Nussinov, R. How do thermophilic proteins deal with heat? *Cell. Mol. Life Sci.* **2001**, *58*, 1216.
- (4) Wrba, A.; Schweiger, A.; Schultes, V.; Jaenicke, R.; Zavodszky, P. Extremely thermostable D-glyceraldehyde-3-phosphate dehydrogenase from the Eubacterium Thermotoga-Maritima. *Biochemistry* **1990**, *29*, 7584.
- (5) Lazaridis, T.; Lee, I.; Karplus, M. Dynamics and unfolding pathways of a hyperthermophilic and a mesophilic rubredoxin. *Protein Sci.* **1997**, *6*, 2589.
- (6) Zavodszky, P.; Kardos, J.; Svignor, A.; Petsko, G. A. Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 7406.
- (7) Butterwick, J. A.; Loria, J. P.; Astrof, N. S.; Kroenke, C. D.; Cole, R.; Rance, M.; Palmer, A. G. Multiple time scale backbone dynamics of homologous thermophilic and mesophilic ribonuclease HI enzymes. *J. Mol. Biol.* **2004**, *339*, 855.
- (8) Wolf-Watz, M.; Thai, V.; Henzler-Wildman, K.; Hadjipavlou, G.; Eisenmesser, E. Z.; Kern, D. Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nat. Struct. Mol. Biol.* **2004**, *11*, 945.
- (9) Fitter, J.; Heberle, J. Structural equilibrium fluctuations in mesophilic and thermophilic alpha-amylase. *Biophys. J.* **2000**, *79*, 1629.
- (10) Hernandez, G.; Jenney, F. E.; Adams, M. W. W.; LeMaster, D. M. Millisecond time scale conformational flexibility in a hyperthermophile protein at ambient temperature. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 3166.
- (11) Tehei, M.; Madern, D.; Franzetti, B.; Zaccai, G. Neutron scattering reveals the dynamic basis of protein adaptation to extreme temperature. *J. Biol. Chem.* **2005**, *280*, 40974.
- (12) Grottesi, A.; Ceruso, M. A.; Colosimo, A.; Di Nola, A. Molecular dynamics study of a hyperthermophilic and a mesophilic rubredoxin. *Proteins: Struct., Funct., Genet.* **2002**, *46*, 287.
- (13) Wintrode, P. L.; Zhang, D. Q.; Vaidehi, N.; Arnold, F. H.; Goddard, W. A. Protein dynamics in a family of laboratory evolved thermophilic enzymes. *J. Mol. Biol.* **2003**, *327*, 745.
- (14) Colombo, G.; Merz, K. M. Stability and activity of mesophilic subtilisin E and its thermophilic homolog: Insights from molecular dynamics simulations. *J. Am. Chem. Soc.* **1999**, *121*, 6895.
- (15) Stone, M. J. NMR relaxation studies of the role of conformational entropy in protein stability and ligand binding. *Acc. Chem. Res.* **2001**, *34*, 379.
- (16) Marcos, E.; Mestres, P.; Crehuet, R. Crowding Induces Differences in the Diffusion of Thermophilic and Mesophilic Proteins: A New Look at Neutron Scattering Results. *Biophys. J.* **2011**, *101*, 2782.
- (17) Lee, B. I.; Chang, C.; Cho, S. J.; Eom, S. H.; Kim, K. K.; Yu, Y. G.; Suh, S. W. Crystal structure of the MJ0490 gene product of the hyperthermophilic archaeabacterium Methanococcus jannaschii, a novel member of the lactate/malate family of dehydrogenases. *J. Mol. Biol.* **2001**, *307*, 1351.
- (18) Dunn, C. R.; Wilks, H. M.; Halsall, D. J.; Atkinson, T.; Clarke, A. R.; Muirhead, H.; Holbrook, J. J. Design and synthesis of new enzymes based on the lactate-dehydrogenase framework. *Philos. Trans. R. Soc. London, Ser. B* **1991**, *332*, 177.
- (19) Hess, B.; Kutzner, C.; Van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435.
- (20) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225.
- (21) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926.
- (22) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald - An N-log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089.
- (23) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463.
- (24) Berendsen, H. J. C.; Postma, J. P. M.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684.
- (25) Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. Essential dynamics of proteins. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 412.
- (26) Chennubhotla, C.; Bahar, I. Markov propagation of allosteric effects in biomolecular systems: application to GroEL-GroES. *Mol. Syst. Biol.* **2006**, *2*, 36.
- (27) Bruschweiler, R. Collective protein dynamics and nuclear-spin relaxation. *J. Chem. Phys.* **1995**, *102*, 3396.
- (28) Holm, L.; Kaariainen, S.; Rosenstrom, P.; Schenkel, A. Searching protein structure databases with DaliLite v.3. *Bioinformatics* **2008**, *24*, 2780.
- (29) Hayward, J. A.; Smith, J. C. Temperature dependence of protein dynamics: Computer simulation analysis of neutron scattering properties. *Biophys. J.* **2002**, *82*, 1216.
- (30) Meinholt, L.; Clement, D.; Tehei, M.; Daniel, R.; Finney, J. L.; Smith, J. C. Protein dynamics and stability: The distribution of atomic fluctuations in thermophilic and mesophilic dihydrofolate reductase derived using elastic incoherent neutron scattering. *Biophys. J.* **2008**, *94*, 4812.
- (31) Thompson, M. J.; Eisenberg, D. Transproteomic evidence of a loop-deletion mechanism for enhancing protein thermostability. *J. Mol. Biol.* **1999**, *290*, 595.
- (32) Chan, M. K.; Mukund, S.; Kletzin, A.; Adams, M. W. W.; Rees, D. C. Structure of a hyperthermophilic tungstopterin enzyme, aldehyde ferredoxin oxidoreductase. *Science* **1995**, *267*, 1463.
- (33) Britton, K. L.; Yip, K. S. P.; Sedelnikova, S. E.; Stillman, T. J.; Adams, M. W. W.; Ma, K.; Maeder, D. L.; Robb, F. T.; Tolliday, N.; Vetrani, C.; Rice, D. W.; Baker, P. J. Structure determination of the glutamate dehydrogenase from the hyperthermophile Thermococcus litoralis and its comparison with that from Pyrococcus furiosus. *J. Mol. Biol.* **1999**, *293*, 1121.

- (34) de Bakker, P. I. W.; Hunenberger, P. H.; McCammon, J. A. Molecular dynamics simulations of the hyperthermophilic protein Sac7d from *Sulfolobus acidocaldarius*: Contribution of salt bridges to thermostability. *J. Mol. Biol.* **1999**, *285*, 1811.
- (35) Thomas, A. S.; Elcock, A. H. Molecular simulations suggest protein salt bridges are uniquely suited to life at high temperatures. *J. Am. Chem. Soc.* **2004**, *126*, 2208.
- (36) Danciulessu, C.; Ladenstein, R.; Nilsson, L. Dynamic arrangement of ion pairs and individual contributions to the thermal stability of the cofactor-binding domain of glutamate dehydrogenase from *Thermotoga maritima*. *Biochemistry* **2007**, *46*, 8537.
- (37) Vinther, J. M.; Kristensen, S. M.; Led, J. J. Enhanced Stability of a Protein with Increasing Temperature. *J. Am. Chem. Soc.* **2011**, *133*, 271.
- (38) Mrabet, N. T.; Vandenbroeck, A.; Vandenbrande, I.; Stanssens, P.; Larocque, Y.; Lambeir, A. M.; Matthijssens, G.; Jenkins, J.; Chiadmi, M.; Vantilbeurgh, H.; Rey, F.; Janin, J.; Quax, W. J.; Lasters, I.; Demaeeyer, M.; Wodak, S. J. Arginine residues as stabilizing elements in proteins. *Biochemistry* **1992**, *31*, 2239.
- (39) Siddiqui, K. S.; Poljak, A.; Guilhaus, M.; De Francisci, D.; Curmi, P. M. G.; Feller, G.; D'Amico, S.; Gerday, C.; Uversky, V. N.; Cavicchioli, R. Role of lysine versus arginine in enzyme cold-adaptation: Modifying lysine to homo-arginine stabilizes the cold-adapted alpha-amylase from *Pseudoalteromonas haloplanktis*. *Proteins: Struct., Funct., Bioinf.* **2006**, *64*, 486.
- (40) Melchionna, S.; Sinibaldi, R.; Briganti, G. Explanation of the stability of thermophilic proteins based on unique micromorphology. *Biophys. J.* **2006**, *90*, 4204.
- (41) Zhou, Y. Q.; Vitkup, D.; Karplus, M. Native proteins are surface-molten solids: Application of the Lindemann criterion for the solid versus liquid state. *J. Mol. Biol.* **1999**, *285*, 1371.
- (42) Camps, J.; Carrillo, O.; Emperador, A.; Orellana, L.; Hospital, A.; Rueda, M.; Cicin-Sain, D.; D'Abramo, M.; Gelpi, J. L.; Orozco, M. FlexServ: an integrated tool for the analysis of protein flexibility. *Bioinformatics* **2009**, *25*, 1709.