JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

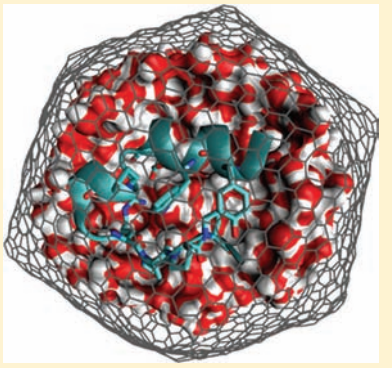
pubs.acs.org/JACS

Simulation Studies of Protein Folding/Unfolding Equilibrium under Polar and Nonpolar Confinement

Jianhui Tian[†] and Angel E. Garcia*

Department of Physics, Applied Physics and Astronomy and Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, New York 12180, United States

ABSTRACT: We study the equilibrium folding/unfolding thermodynamics of a small globular miniprotein, the Trp cage, that is confined to the interior of a 2 nm radius fullerene ball. The interactions of the fullerene surface are changed from nonpolar to polar to mimic the interior of the GroEL/ES chaperonin that assists proteins to fold in vivo. We find that nonpolar confinement stabilizes the folded state of the protein due to the effects of volume reduction that destabilize the unfolded state and also due to interactions with the fullerene surface. For the Trp cage, polar confinement has a net destabilizing effect that results from the stabilizing confinement and the competitive exclusion effect that keeps the protein away from the surface hydration shell and stronger interactions between charged side chains in the protein and the polar surface that compete against the formation of an ion pair that stabilizes the protein folded state. We show that confinement effects due to volume reduction can be overcome by sequence-specific interactions of the protein side chains with the encapsulating surface. This study shows that there is a complex balance among many competing effects that determine the mechanism of GroEL chaperonin in enhancing the folding rate of polypeptide inside its cavity.



INTRODUCTION

How proteins fold into their native structures in the cell remains to be an important unanswered question in biology. Numerous proteins fold spontaneously in vitro, while in the cell a large portion of newly synthesized proteins require the assistance of a molecular chaperone to reach their folded states efficiently .[1] As one specific example, the *Escherichia coli* chaperonin GroEL is the best characterized molecular chaperone that assists proteins to fold in vivo.[2] GroEL assists proteins to fold by two mechanisms: *cis* for small proteins and *trans* for large proteins. Both of the mechanisms include the nucleotide cycling and the cycles of polypeptides binding and releasing with respect to the cavity of GroEL.[3] However, whether GroEL plays an active role toward polypeptides folding or just a passive role by isolating polypeptides into the cavity during this process is yet to be clarified. Brinker et al.,[4] using an engineered GroEL chaperonin system, showed that confinement of unfolded proteins alone accelerates folding of proteins inside the cage without the ATP-dependent cycles of binding and releasing.

Many theoretical studies have been conducted to understand how confinement helps proteins to fold.[5−21] These studies, most of which use off-lattice models or Gō-like models, provide us much understanding of the problem.[5,7−15] However, these models neglected many details, such as solvent effects and binding of the protein to the confining surfaces, which might be critical when considering folding in vivo. Water under confinement behaves differently from that in the bulk both thermodynamically and kinetically,[22−25] and these properties cannot be captured and reproduced without explicit solvent models. A study by Lucent et al.,[20] using an explicit solvent model, showed that there is an "unfolding effect" and a protein is destabilized when confined by a purely repulsive potential together with solvent. This unfolding effect was unexpected. Lucent et al. hypothesized that confined solvent may create an energy landscape that is conducive to unfolding.

With the GroES binding, the GroEL cavity wall changes from a hydrophobic surface into a hydrophilic surface. Hydrophobic and hydrophilic surfaces influence the water dynamics in different ways and have been shown to have different effects on the folding of hydrophobic polymers.[26] Xu et al. also showed, by studying a system of methane molecules solvated by water under fullerene confinement, that polar confinement is more capable of aggregating hydrophobic molecules than nonpolar confinement.[27] The effects present on simple hydrophobic polymers interacting with surfaces are expected to be present and influence significantly protein folding under confinement. However, how different topologies and the polarity of confining volumes affect proteins to fold has not been studied so far.

Here, we employed a fullerene ball as the confinement medium with purely van der Waals interactions to describe interactions with the fullerene carbon atoms to mimic the GroEL cavity without GroES binding (i.e., hydrophobic confinement) and both electrostatic and van der Waals interactions with the fullerene carbon atoms to mimic the GroEL cavity with GroES binding (i.e., hydrophilic confinement). We study the folding of the Trp cage miniprotein under these confinement conditions, using all-atom, explicit solvent models to model how the GroEL

Received: June 18, 2011
Published: August 19, 2011

cavity environment may assist proteins to fold. The folding/unfolding equilibrium of the Trp cage miniprotein is studied, without any bias, using molecular dynamics simulations. The Trp cage miniprotein is a designed protein that shows equilibrium thermodynamic properties much like larger globular proteins.[28] Trp cage is composed of an α helix, a 3_{10} helix, and a polyproline II segment. The folding of the Trp cage has been widely studied experimentally.[29−35] The fast folding, on the microsecond time scale,[29] makes it amenable to molecular simulations. The Trp cage miniprotein has served as a model system in computational studies for force fields[36−40] and for testing modeling techniques[41−45] that are benchmarked against detailed structural,[28] thermodynamic,[46,47] kinetic,[29−32,48] and protein design[49] data. Multiple studies describing the folding of the Trp cage have successfully reproduced the structure of the folded state.[36,39−41,48,50−61]

In this study we show that the weak nonpolar confinement of the fullerene ball stabilizes the folded state, while a polar confinement decreases the folding stability. The trends in stabilization depend on confinement, on solvent-mediated interactions with the confining surface, and on the direct interactions of side chains with the surface that depend on the protein sequence. These findings point to a possible mechanism for GroEL-assisted folding that includes confinement, interactions with the confining surface, solvent, and the protein sequence.

METHODS

The Trp cage miniprotein structure is generated using the xleap program included in the AMBER 8.0 program.[62] The amino acid sequence is as follows: ace-NLYIQWLKDGGPSSGRPPPS-NMe. The N- and C-termini are capped with acetyl and methyl groups, and the constructed model peptide consists of 313 atoms. There are two positive side chains (Lys 8 and Arg 16) and one negative side chain (Asp 9). The net charge of the protein is a positive one. The peptide is solvated by TIP3P waters[63] in a cubic box and collapsed to a compact, unfolded structure under isobaric and isothermic conditions (1 atm and 300 K). One chloride ion is added to neutralize the system. This unfolded configuration is then used to model the folding/unfolding equilibrium of the protein in the inner core of the confining fullerene. We use a 2160 carbon atom fullerene ball as the confinement, whose radius of 2.0 nm is comparable to the radius of the GroEL cavity.[64] The van der Waals (vdW) parameters of the CA carbon atom in the AMBERff94 force field are used for the fullerene carbon atoms. However, the ε values are reduced by 50% to weaken the vdW interaction between the confinement wall and peptide following a method similar to that used by Waghe et al.[65] The reduced strengths of the interactions have been shown to be appropriate to model carbon nanotube structures in water.[66] The bond and angle interactions of the same kind of carbon atom in the AMBERff94 force field are used.[67] We also studied the protein folding under confinement for a system where the protein−fullerene interactions were taken as those of the CA atom in the AMBERff94 force field. This calculation represents the protein under strong nonpolar confinement. This system did not show any folding events within 80 ns/replica in the replica exchange molecular dynamics (REMD) calculations. Two kinds of confinement are described here: nonpolar confinement and polar confinement. For the nonpolar confinement, all carbon atoms have no partial charge, and there are only van der Waals interactions between the confinement wall and solvent or peptide. For the polar confinement, randomly generated charges are distributed on each of the carbon atoms. Charge values are random and uniformly distributed between −0.15e and +0.15e. The net charge on the fullerene is zero. These charges make the confinement wall hydrophilic. The number of water molecules inside the fullerene is chosen to make the water density close to the bulk water density in the center of the fullerene ball for both systems. In GroEL the net charge of the protein exposed to the inner cavity is negative, which should be balanced by ions present in the solution. This charge distribution is a simplification of the GroEL and its ionic environment inside the cavity.

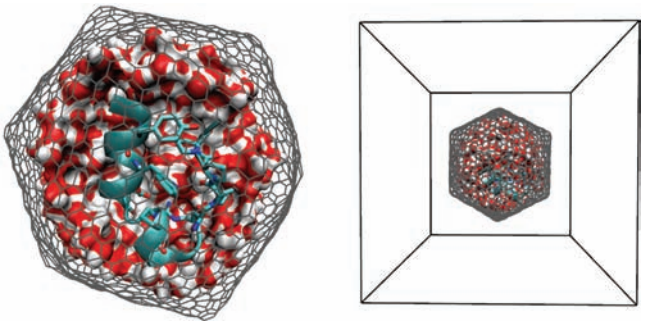


Figure 1. Trp cage under nonpolar fullerene confinement. The Trp cage, in cyan, is shown in cartoon presentation, waters are shown in white and red surfaces, and fullerene is in gray. The fullerene is located at the center of the box, and the remainder of the cubic box is empty.

The system under (strong and weak) nonpolar confinement contains 718 TIP3P waters, 1 chloride ion, and the Trp cage miniprotein inside a fullerene ball. The total number of atoms in the system is 4628. AMBERff94[67] force field parameters are used for the Trp cage miniprotein. Constant temperature and volume REMD simulations are conducted to sample the conformational space.[68] REMD is an enhanced sampling technique based on the parallel tempering Monte Carlo method,[68−71] where copies of identical systems are simulated at different temperatures. Periodically state exchange between replicas is attempted, and the acceptance rule for each move between states i and j, dictated by a Boltzmann distribution, is $P_{acc} = \min\{1, \exp[(\beta_i - \beta_j)(U(\vec{r}_i^N) - U(\vec{r}_j^N))]\}$, where $\beta = 1/k_B T$ and $U(\vec{r}_j^N)$ represents the configurational energy of the system in state i. Together with the exchange, the particle momenta are scaled by $(T_i/T_j)^{1/2}$, such that the kinetic energy terms in the Boltzmann factor cancel out.[68] REMD sampling can also be described in terms of umbrella sampling.[71] The temperature spacing between replicas is chosen to ensure sufficient overlapping of the energy distribution between neighboring replicas such that exchange attempts are, on average, accepted with a 25% probability.[72] The potential energy distributions of the system at 20 different temperatures are simulated at constant volume and constant temperature to set up the temperature distribution for the replicas. A total of 46 replicas are used to cover the temperature range from 275 to 695 K. The temperature distribution is as follows: 275, 280, 285, 290, 296, 302, 308, 314, 320, 326, 332, 339, 346, 353, 360, 367, 374, 382, 390, 398, 406, 414, 423, 432, 441, 451, 460, 470, 480, 491, 502, 513, 524, 535, 547, 559, 572, 584, 597, 610, 624, 638, 652, 666, 681, and 695 K. Gromacs 4[73] is used to perform the simulations. Exchanges are attempted every 2000 integration steps (4 ps). The system under strong nonpolar confinement did not show any folding and is not described any further.

A 7.5 nm × 7.5 nm × 7.5 nm cubic box is used, with the Trp cage−fullerene−water components located at the center of the box. The Nose−Hoover thermostat is used for the temperature coupling with a coupling time constant τ_T = 1.0 ps.[74,75] The protein, solvent, and fullerene atoms are coupled separately to thermostats with the same coupling parameters. van der Waals interactions are treated using a 0.9 nm cutoff distance. The electrostatic interactions are treated by smooth particle mesh Ewald summation.[76] All bond interactions involving hydrogen atoms are constrained using SETTLE[77] and SHAKE[78] to allow a 2 fs integration time step. During the simulations, the carbon atoms of the fullerene ball are restrained to their initial positions through

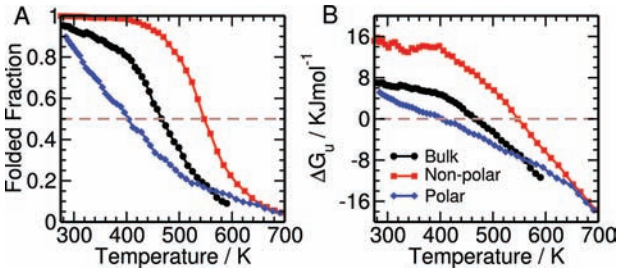


Figure 2. Temperature dependence of the folding/unfolding equilibrium for the three systems. (A) shows the fraction of states folded at different temperatures. (B) shows the free energy of unfolding at different temperatures.

a harmonic potential with a force constant of 1000 kJ/(mol nm^2). Each of the 46 replicas is simulated for 160 ns, for a total sampling time of 7.36 μs. Snapshots of the conformation are saved every 2 ps for analysis. Figure 1 shows the system of the Trp cage under nonpolar confinement.

The system under polar confinement contains 826 waters, 1 chloride, a Trp cage, and a fullerene ball. The total number of atoms in the system is 4952. Similar procedures are used to set up the temperature distribution of the system. A total of 46 replicas are used to cover the temperature from 285 to 691 K, and the distribution is as follows: 285, 290, 295, 300, 306, 311, 317, 323, 329, 335, 341, 347, 354, 360, 367, 374, 381, 389, 396, 404, 412, 420, 428, 437, 445, 454, 464, 473, 483, 493, 503, 513, 524, 535, 546, 558, 570, 582, 595, 608, 621, 634, 648, 662, 677, and 691 K. The same box size and MD simulation parameters are used for this system. Each of the 46 replicas is also simulated for 160 ns.

The results of the simulations for the protein under weak nonpolar and polar confinement are compared with the results for the Trp cage in bulk water obtained by Paschek et al.[53,55] and Canchi et al.[60] using the AMBERff94 force field.[67] In bulk water, the Trp cage shows an unfolding temperature of 450 K. This high stability is due to the large stabilization of α helices of the AMBER94 force field.[79] Previous calculations of the Trp cage with AMBERff94 showed that REMD ensembles produced stationary averages in simulations extending to 60 ns per replica.[53,55]

RESULTS

The Trp cage ensemble of configurations are collected at different temperatures from the last 100 ns trajectories for each of the replicas under the nonpolar and polar confinement, and they are compared with the configurations collected from the last 100 ns trajectories for the Trp cage in bulk water studied by Canchi et al.[60] Figure 2A shows the fraction of folded states as a function of temperature for the three systems. At all temperatures, a larger fraction of states are folded under nonpolar confinement, when compared with the system in bulk water. For the system under polar confinement, a smaller fraction of states are folded at low temperatures and a larger fraction of states are folded at very high temperatures, when compared with the system in bulk water. The fraction folded under nonpolar confinement did not change much until 400 K. The fraction folded under polar confinement decreases linearly with increasing temperature. The stability profiles as a function of temperature for the Trp cage under confinement are different from the profile obtained for bulk water.

Figure 2B shows the Gibbs free energy of unfolding at different temperatures, which is calculated from the knowledge of the fraction folded using the relation $\Delta G_u = -RT \ln[(1 - x_{folded})/x_{folded}]$, assuming pressure changes upon folding/unfolding are small and that the system follows two-state thermodynamics.

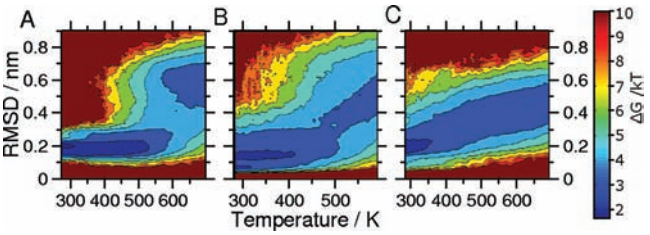


Figure 3. Free energy landscape of the Trp cage as a function of the C_α rmsd and temperature in the three systems. The contour plots are in units of k_BT, and the difference between neighboring contour lines is 1 k_BT. (A) System under nonpolar confinement. (B) System in bulk water. (C) System under polar confinement.

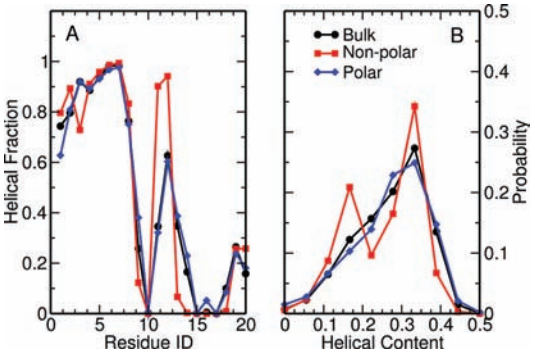


Figure 4. (A) Helical fraction for each of the residues in different systems. The residue is considered to belong to the helical basin if (ϕ, ψ) is at ($-60 \pm 30°$, $-60 \pm 35°$). (B) Probability of the helical content of the entire Trp cage miniprotein in different systems.

Here x_{folded} is the fraction of states folded, T is the temperature for each replica, and R is the molar gas constant. Nonpolar confinement stabilizes the Trp cage and increases the melting temperature from 470 K in the bulk to 550 K, and the free energy of unfolding is ΔG_u = 15 kJ/mol at 285 K compared with ΔG_u = 7 kJ/mol at 284 K in bulk water. Polar confinement destabilizes the protein and decreases the melting temperature to 400 K, and the free energy of unfolding is ΔG_u = 5 kJ/mol at 285 K. In Figure 3, we compare the free energy landscape of the Trp cage for the three systems as a function of the C_α rmsd and temperature. From the free energy landscapes, we see that the overall range of rmsd values sampled is narrower than for the bulk, indicating that the sampled configurational space is reduced in both systems under confinement. A detailed comparison shows that the folded basin (rmsd < 0.3 nm) of the nonpolar confinement expands to 530 K and is well separated from the unfolded basin at high temperatures. In contrast, the folded basin of the polar confinement extends to 350 K, and there is no free energy barrier separating the folded and unfolded basins at higher temperatures. These results clearly show the stabilization under nonpolar confinement and destabilization under polar confinement of the folded states.

We now analyze the secondary structural features of the Trp cage miniprotein in bulk water and nonpolar and polar confinement at the lowest temperature sampled (284, 285, and 285 K, respectively). The backbone dihedral angles (ϕ, ψ) are calculated for all 20 amino acid residues. A residue is considered to belong to the helical basin if $-90° < \phi < -30°$ and $-95° < \psi < -25°$. In Figure 4A, we showed the fraction of helical content for each of

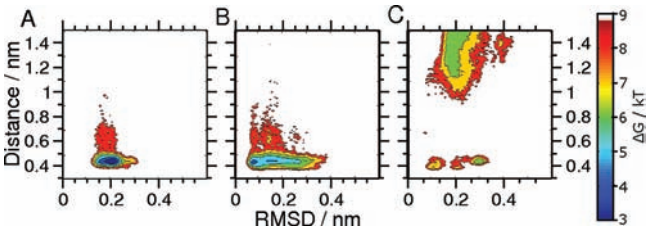


Figure 5. Free energy landscape of the Trp cage as a function of the C_α rmsd and ion pair distance between Arg 9 C_γ and Asp 16 C_ζ for the three systems. The contour plots are in units of k_BT, and the difference between neighboring contour lines is 1 k_BT. (A) System under nonpolar confinement at 285 K. (B) System in bulk water at 284 K. (C) System under polar confinement at 285 K.

the residues. Under nonpolar confinement, residue Tyr 3 shows lower helical content, compared with the systems in the bulk and under polar confinement, while Gly 11 and Pro 12 have higher helical content. Further analysis shows that the Tyr 3 side chain prefers to interact with the fullerene inner surface, and this interaction affects the backbone dihedral angles, thus reducing its helical content. The backbones of Gly 11 and Pro 12 are highly dehydrated, and dehydration enhanced the formation of helices.[79] The residues for the protein under polar confinement have helical fractions very similar to those in bulk water, and only small differences are observed for residues Asn 1, Asp 9, Ser 14, and Arg 16. Figure 4B shows the probability distribution of the helical content for the Trp cage in the three systems. For this plot, a configuration is considered helical when three consecutive residues are in the helical basin, as defined by (ϕ, ψ) angles in the $(-60 \pm 30°, -60 \pm 35°)$ region. Every further neighboring helical residue increases the helical content by one, so that the (hypothetical) maximum helical content of the Trp cage is 18. The helical content in Figure 4B is the number of helical residues divided by 18. The protein under nonpolar confinement has significantly larger helical content probabilities at 0.16 and 0.34 than in bulk water. However, it has lower helical content probabilities at 0.22 and 0.28. The low helical content probability at 0.22 and 0.28 is due to the decreased helical fraction of Tyr 3. The helical content of protein under polar confinement is practically the same as in bulk water.

We further analyze the structural features of the states in the three systems by calculating two-dimensional free energy landscapes of the Trp cage as a function of the C_α rmsd and the ion pair distance, d_{pair}, between Arg 9 C_γ and Asp 16 C_ζ. The salt bridge formation (d_{pair} < 0.5 nm) between these two residues is known to stabilize the Trp cage folded state.[80] Figure 5 shows the energy landscapes for the system under nonpolar confinement at 285 K, in bulk water at 284 K, and under polar confinement at 285 K. Under nonpolar confinement, Trp cage shows a well-defined folded basin in which the ion pair is formed. In bulk water, the folded basin is larger than the basin under nonpolar confinement and includes states with larger rmsd's that form the ion pair. Under polar confinement, there are folded basins with two minima (centered at rmsd's near 0.1 and 0.3 nm and d_{pair} < 0.5 nm) in which the ion pair is formed. There is another folded basin with rmsd near 0.2 nm, where the ion pair is not formed (d_{pair} > 0.8 nm). The ion pair distance distributions for the systems under nonpolar confinement and in bulk water have sharp peaks at 0.45 nm and a low-population tail that extends to 0.8 nm. Under polar confinement, the ion pair distance

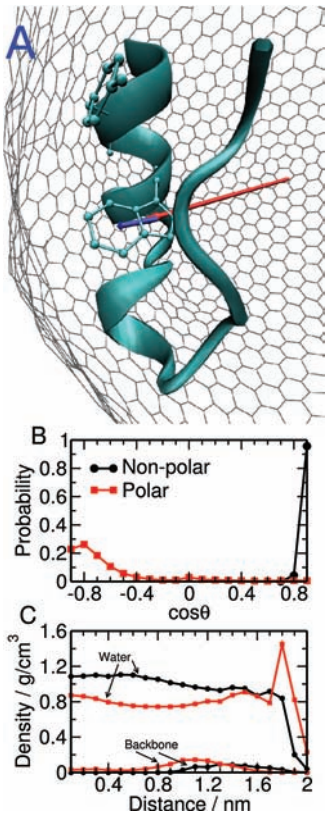


Figure 6. (A) Definition of the orientation angle θ. We define θ as the angle between a vector which lies on the indole plane of Trp 6 and points from the center of the pyrrole ring (blue arrow) to the center of the benzene ring in Tyr 3 and a vector from the center of the fullerene pointing out to the tryptophan side chain (red arrow). (B) Probability distribution of cos(θ) of the two systems under confinement to describe the orientation of the Trp cage. The black line is for nonpolar confinement, and the red line is for polar confinement. (C) Density of water (upper part) and the Trp cage backbone (C_α, C, N) (lower part) along the radial direction of the fullerene ball. The distance is measured from the center of the fullerene. The color scheme is the same as in (B).

distribution is divided into two well-populated regions, one located at 0.45 nm and the other between 1.0 and 1.5 nm, indicating that there is a significant population of the folded protein in which the ion pair is not formed. An analysis of the interactions between side chains and the fullerene surface shows that all the charged side chains (Lys 8, Arg 9, Asp 16) bind to the (partially) charged fullerene atoms present under polar confinement, thus decreasing the probability of forming the ion pair. The Trp cage can maintain the folded state with and without the ion pair.[81]

Stabilization/Destabilization under Confinement. To understand the origin of the stabilization effect under nonpolar confinement and the destabilization effect under polar confinement, we analyze the orientation of the Trp cage protein secondary structural elements relative to the fullerene surface. Residues Tyr 3, Trp 6, Leu 7, Gly 11, Pro 12, Pro 18, and Pro 19 define the hydrophobic core of the Trp cage. Residues Leu 2 to Lys 8 form an α helix, and residues Pro 17 to Pro 19 form a polyproline II helix that packs against the α helix in the folded state. The axes of the α helix and the polyproline II helix are used to define a plane. One side of this plane is mostly hydrophobic, containing Tyr 3, Trp 6, Leu 7, Pro 12, and Pro 18 side chains; the other side is mostly hydrophilic, containing Gln 5, Asp 9, and Arg

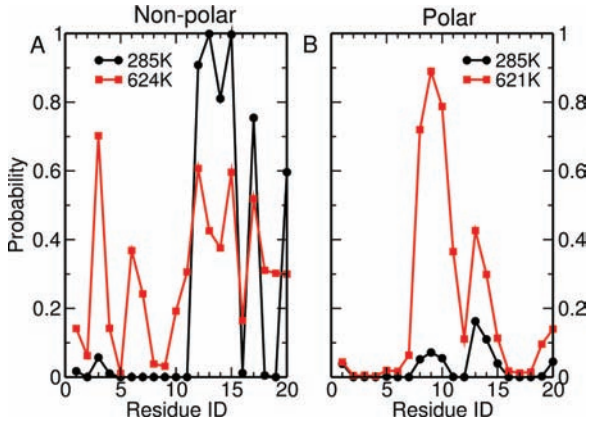


Figure 7. Probability of having C_α atoms within 0.45 nm of the fullerene carbon atoms in ensembles dominated by the folded (285 K) and unfolded (624 K) protein configurations. (A) is for nonpolar confinement. (B) is for polar confinement.

16 side chains. We define angle θ as the angle between a vector $\vec{a}$, which lies on the indole plane of Trp 6 and points from the center of the pyrrole ring to the center of the benzene ring in Tyr 3, and a vector $\vec{b}$ from the center of the fullerene pointing out to the tryptophan side chain. The orientation of the protein inside the fullerene ball, described by this angle θ, is shown in Figure 6A. If $\cos(\theta)$ is positive, the hydrophobic side is facing the fullerene inner surface; if $\cos(\theta)$ is negative, the hydrophilic side is facing the fullerene inner surface. In Figure 6B, we showed the distribution of $\cos(\theta)$ of the Trp cage under the two confinement conditions. Under nonpolar confinement, $\cos(\theta)$ is distributed above 0.7 and peaked at 0.9, which indicates the hydrophobic side is facing the fullerene inner surface and the indole plane is almost perpendicular to the surface. Under polar confinement, $\cos(\theta)$ has a much wider distribution and is mainly negative, indicating that the hydrophilic side is facing the fullerene inner surface. We also studied the location of the Trp cage within the cavity. Figure 6C shows the density of water and the Trp cage backbone (C_α, C, N) along the radial direction of the fullerene ball, with the origin at the center of the ball. Under nonpolar confinement, the Trp cage backbone density is restricted to positions from 1.0 to 1.9 nm to the center of the fullerene and the density of water in the center of fullerene is larger than 1.0 g/cm^3. Under polar confinement, the Trp cage backbone density is widely distributed from 0 to 1.6 nm and the density of water at the fullerene surface has a peak of 1.5 g/cm^3. An enriched layer of water molecules near the polar confinement inner surface exists, and the Trp cage does not bind to the fullerene inner surface as it did under nonpolar confinement. This enriched layer of water molecules is due to the strengthening of water−fullerene interactions and is responsible for what Athawale et al. have termed "competitive expulsion".[82] Competitive expulsion results from a water-mediated fullerene−protein repulsion that excludes protein atoms from penetrating the charged fullerene hydration shell, since the water interactions are stronger. Charged protein side chains will compete against water and penetrate the fullerene hydration shell. We observe that the Lys 8, Asp 9, and Arg 16 charged side chains preferentially bind to the charged fullerene atoms and diffuse on the surface by hopping. The location of the Trp cage inside the nonpolar and polar confinement volumes agrees well with the binding of a polypeptide to the hydrophobic GroEL cavity surface and the release of the polypeptide from the GroEL hydrophilic cavity surface.[3]

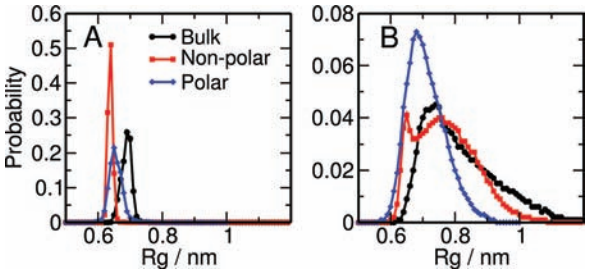


Figure 8. Probability distribution for the radius of gyration of the C_α atoms in the ensembles sampled by the system under nonpolar, polar, and bulk environments at (A) low (284−285 K) and (B) high (>600 K) temperatures, corresponding to the distributions dominated by the folded and unfolded states, respectively.

We now analyze the interaction of the Trp cage with the fullerene inner surface by calculating the probability of having the C_α atoms within 0.45 nm of the fullerene carbon atoms for ensembles dominated by the folded (285 K) and unfolded (624 and 621 K) states under nonpolar and polar confinement, respectively. In Figure 7 we show that for the folded ensemble under nonpolar confinement, the residues Pro 12, Ser 13, Ser 14, Gly 15, Pro 17, and Ser 20 have high probability to be in contact with the fullerene inner surface. The near 100% contact probability of Ser 13 and Gly 15 again verifies that the protein is bound to the fullerene inner surface with a preferred orientation. The folded ensemble under polar confinement does not show much contact with the fullerene inner surface, which agrees with the density distribution of the Trp cage under polar confinement shown in Figure 6C. In the unfolded ensemble under nonpolar confinement, Ser 13 and Gly 15 have a lower probability of binding to the fullerene inner surface. Also, more residues are in contact with the fullerene inner surface under nonpolar confinement. This indicates that, in the unfolded ensemble, the structures are more extended but are still close to the fullerene inner surface. For the unfolded ensemble under polar confinement, a large probability is observed for some residues (8−10 and 13) to be in contact with the fullerene inner surface, but overall most of the residues are not in contact with the inner surface. This indicates that, in the unfolded ensemble, the protein has a preferred orientation, but is not close to the fullerene inner surface.

Figure 8 shows the radius of gyration of Trp cage C_α atoms for the folded ensembles and unfolded ensembles, sampled by the system in bulk, polar, and nonpolar environments. In the folded ensembles (at low temperature) we obtained a distribution peaked at smaller R_g under nonpolar and polar confinement, compared with the distribution for the system in bulk water. This suggests the presence of an enhanced hydrophobic effect that keeps the protein more compact. A similar enhanced hydrophobic effect under confinement has been reported previously by Vaitheeswaran et al.[83] and Xu et al.[27] when studying methane molecules under confinement and by Ghosh et al.[84] in ionic solutions. As the hydrophobic effect is an essential driving force for protein folding, it definitely contributes to the decrease of the free energy of folding under confinement. In the unfolded ensemble we obtained distributions for R_g in the bulk that reach larger values of R_g than for the systems under confinement. This is the expected confinement effect that reduces extended conformations in the folded state, thus reducing the free energy of the unfolded state relative to the unfolded state in the bulk.

In Figure 9, we compare the backbone hydration of the Trp cage folded ensemble in the three systems. The Trp cage

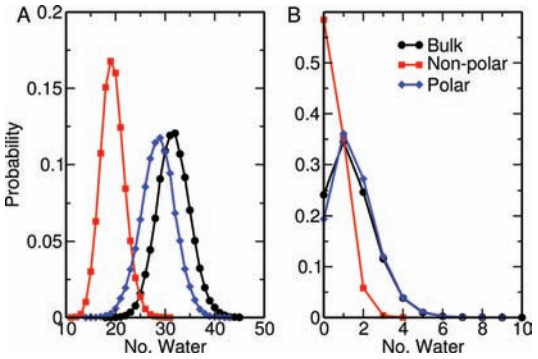


Figure 9. Comparison of the solvation of the Trp cage in the three systems. (A) Distribution of solvation water numbers within 0.35 nm of the backbone (N)H or (C)O. (B) Distribution of water numbers within 0.35 nm of the Trp 6 side chain heavy atoms. The black line with circles is for the bulk, the red line with squares is for nonpolar confinement, and the blue line with tilted squares is for polar confinement.

backbone is highly dehydrated under nonpolar confinement and slightly dehydrated under polar confinement, when compared with that in bulk water. However, the Trp 6 side chain experiences a hydration environment under polar confinement very similar to that in bulk water and has much less solvation water under nonpolar confinement. The differences in the backbone and Trp 6 side chain solvation are related to the location and orientation of the Trp cage under confinement. We have already shown that the hydrophobic side of the Trp cage faces the fullerene wall and is in close contact with the inner surface under nonpolar confinement. This location and orientation cause the dehydration of the hydrophobic side and the Trp 6 side chain. Under polar confinement, the hydrophilic side of the Trp cage faces the fullerene inner surface and locates near the fullerene inner surface, while the Trp 6 side chain is exposed to water in the center of the confining volume. This explains why, under polar confinement, the Trp cage backbone is slightly dehydrated and the Trp 6 side chain has hydration very similar to that in bulk water.

DISCUSSION

We have described the folding/unfolding equilibrium of the Trp cage miniprotein under polar and nonpolar confining environments by molecular dynamics simulations. The simulations showed that a significant stabilization effect is obtained for the Trp cage under weak nonpolar confinement, while a destabilization is obtained under polar confinement. We also obtained complete destabilization of the folded state under strong nonpolar interactions. This shows that there are at least three effects determining the stability of proteins under confinement: an entropic effect due to volume reduction (confinement), a direct interaction of the protein with the surface, and solvent-mediated effects, such as the well-known hydrophobic effect and competitive expulsion effect. In the hydrophobic effect, solvent-mediated interactions bring nonpolar groups in the protein and the fullerene surface together and increase the number of water−water interactions. In the competitive expulsion effect, the interactions of water with the confining surface exclude the protein atoms from penetrating the surface solvation shell. This last effect is present when the confining surface is polar or charged. For the entropic effect, Zhou et al.[85] studied proteins under confinement using polymer theory and stated that proteins will be stabilized as the reduced confinement volume excludes many of the unfolded configurations, thus reducing the entropy of the unfolded states while leaving the folded configurations unchanged. We observed a reduction of the overall sampled configurational space of the Trp cage under both nonpolar and polar confinement, when compared with the system in bulk water in Figure 3. We can calculate the change of the free energy of folding for the Trp cage under confinement by using the random-flight Gaussian chain model:[85]

$$\Delta\Delta G_f/k_BT = -2\pi^2 Nb^2/3d^2 - 3\ln(1-2a_N/d) + \ln(6/\pi^2)$$

where d is the diameter of the fullerene, N is the number of residues, b is the effective bond length, and a_N is the radius of gyration of the native structure. Taking d = 3.7 nm, N = 20, b = 0.4 nm, and a_N = 0.64 nm, we get $\Delta\Delta G_f = -0.8\,k_BT = -1.8$ kJ/mol at 285 K. This decrease in the free energy of folding only accounts partially for the total decrease in free energy obtained in the simulations, $\Delta\Delta G_f = -8.0$ kJ/mol, under nonpolar confinement, relative to the free energy of folding in bulk water. In addition to the entropic effect expected from polymer confinement, direct interactions of the protein and the confinement wall are present. Depending on the kind of confinement, the interactions vary. We also studied the folding of the Trp cage miniprotein under nonpolar confinement using the stronger van der Waals parameters of the sp^2 carbon atom of a pure aromatic in the AMBERff94 force field and did not observe a single folding event within 87 ns per replica in an REMD simulation. In these simulations, the Trp cage is always unfolded and bound to the fullerene wall, showing that, under this strong nonpolar confinement, the protein−fullerene wall interactions are stronger than the protein intramolecular interactions that drive folding. A similar effect has been observed by Jamadagni et al.[26] when simulating the folding of a hydrophobic chain in the presence of a hydrophobic surface. Minimalist models of protein folding in the presence of an interacting carbon nanotube have shown that, in the presence of strong interactions of a protein with a wall, the system will maximize the number of interactions with the surface and destabilize the folded state, which bury hydrophobic groups.[86] For the Trp cage we found that when the strengths of the interactions between the protein and fullerene wall are reduced to the smaller values that have been shown to be appropriate for modeling the vdW interactions in carbon nanostructures,[65,66] the folded state is stabilized. The strengths of the interactions between the protein and the fullerene wall are also small under polar confinement, but larger than under nonpolar confinement. Similar behavior has also previously been suggested by Betancourt et al., in studies of simple lattice models of proteins,[5] who observed that the folding rate is enhanced when the hydrophobic interaction between the confinement and the peptide is within an optimum range, while too weak or too strong interactions will retard the folding rates. Xu et al.[14] also found that the enhancement or retardation of protein folding under confinement depends on the interplay of two factors: the spatial confinement and the affinity strength to the confinement wall.

Under polar confinement, the decrease of the free energy of folding due to confinement is opposed by the net increase in free energy of folding. Under polar confinement, the protein prefers to remain largely unbound to the fullerene surface. This effect is related to solvation and electrostatic interactions of the protein

with the fullerene surface. As shown in Figure 7, under polar confinement, the protein core is detached from the surface. The unbinding of a polymer to a polar surface has been observed in the simulation of a nonpolar chain in the presence of a polar surface.[26] This effect has been explained in terms of the favorable interaction of water with the polar surface and the reduction of the attraction due to water-mediated interactions between water and the protein or polymer.[26] This effect has been termed the "competitive expulsion" effect.[82] However, this competitive expulsion effect is not sufficiently strong to overcome the electrostatic attraction between charged side chains (Lys, Arg, and Asp) and the polar fullerene surface. As a result, the charged side chains bind to the surface. In Figure 5, our calculations show a higher percentage of ion pair formation under nonpolar confinement and a lower percentage under polar confinement, when compared with the system in bulk water. The binding of the charged side chains to the fullerene atoms competes with the formation of the Asp−Arg ion pair between Asp 9 and Arg 16 that is known to contribute to the stability of the Trp cage folded structure. Studies in our group have found that the elimination of the ion pair by protonating the Asp 9 side chain decreases the free energy of folding by ∼3 kJ/mol.[81]

Among all the effects that we have discussed above, the entropic effect and the enhanced hydrophobic effect both decrease the free energy of folding for the Trp cage under nonpolar and polar confinement. The increase of ion pair formation under nonpolar confinement further decreases the free energy, resulting in a net stabilization effect as high as −8 kJ/mol. Overall, the interaction of the Trp cage with the fullerene wall under weak nonpolar and polar confinement favors protein folding. However, the disruption of the ion pair formation under polar confinement overcomes the stabilization from the two effects described before and results in a net destabilization of 1.7 kJ/mol.

CONCLUSION

Using replica exchange molecular dynamics simulations, we have studied Trp cage folding/unfolding equilibrium under nonpolar and polar confinement and compared it with Trp cage folding in bulk water. Our results show that the net effect of protein folding stability under confinement has contributions from many factors. These include, but may not be limited to, the entropic stabilization effect, an enhanced hydrophobic effect, and the competitive expulsion of protein atoms from the polar fullerene surface. Although some of these effects should apply to proteins in general, sequence- and structure-specific features of the protein are also important and can overcome the others. These observations are in agreement with studies of proteins under confinement showing that stabilizing excluded volume effects can be changed by soft interactions.[87−89]

The small Trp cage miniprotein has many of the elements of globular proteins and upon confinement exhibits different behaviors, depending on the confining environment. It is found that the system under nonpolar confinement stabilizes the folded state of the Trp cage protein, while in the system under polar confinement the competition between ion pair formation that stabilizes the folded structure and interactions between charged side chains with the polar fullerene surface lowers the overall stability of the folded protein. The Trp cage inside these confinement volumes has a preferred orientation and is in contact with the fullerene wall under nonpolar confinement, but not under polar confinement. Analysis of the free energy landscape of the protein reveals that, under polar confinement, the Trp cage shows a barrierless folding behavior, which is not expected. However, the description of the energy landscape as a function of structural order parameters does not describe the kinetics of folding. Predictably, under nonpolar confinement, nonpolar groups interacted preferentially with the surface. Under polar confinement, the protein is located away from the surface, except for charged side chains. The structure of the folded state of Trp cage is sufficiently plastic to respond to the surface interactions while maintaining the protein core, secondary structure, and fold. This study shows that there is a complex balance among many competing effects that may determine the mechanism of GroEL chaperonin in enhancing the folding rate of polypeptide inside its cavity.

AUTHOR INFORMATION

Corresponding Author

angel@rpi.edu

Present Addresses

[†]Theoretical Biology and Biophysics, Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545.

ACKNOWLEDGMENT

We thank S. Vaitheeswaran and D. R. Canchi for comments and suggestions. This work was supported by grants from the National Science Foundation, NSEC DMR-0642573 and MCB-0543769.

REFERENCES

(1) Hartl, F. U. *Nature* **1996**, *381*, 571–580.
(2) Sigler, P. B.; Xu, Z. H.; Rye, H. S.; Burston, S. G.; Fenton, W. A.; Horwich, A. L. *Annu. Rev. Biochem.* **1998**, *67*, 581–608.
(3) Chaudhuri, T. K.; Verrna, V. K.; Maheshwari, A. *Prog. Biophys. Mol. Biol.* **2009**, *99*, 42–50.
(4) Brinker, A.; Pfeifer, G.; Kerner, M. J.; Naylor, D. J.; Hartl, F. U.; Hayer-Hartl, M. *Cell* **2001**, *107*, 223–233.
(5) Betancourt, M. R.; Thirumalai, D. *J. Mol. Biol.* **1999**, *287*, 627–644.
(6) Baumketner, A.; Jewett, A.; Shea, J. *J. Mol. Biol.* **2003**, *332*, 701–713.
(7) Cheung, M. S.; Thirumalai, D. *J. Mol. Biol.* **2006**, *357*, 632–643.
(8) Cheung, M. S.; Thirumalai, D. *J. Phys. Chem. B* **2007**, *111*, 8250–8257.
(9) Jewett, A. I.; Baumketner, A.; Shea, J. E. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 13192–13197.
(10) Rathore, N.; Knotts, T. A.; de Pablo, J. J. *Biophys. J.* **2006**, *90*, 1767–1773.
(11) Ziv, G.; Haran, G.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 18956–18961.
(12) Takagi, F.; Koga, N.; Takada, S. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 11367–11372.
(13) Friedel, M.; Sheeler, D. J.; Shea, J. E. *J. Chem. Phys.* **2003**, *118*, 8106–8113.
(14) Xu, W. X.; Wang, J.; Wang, W. *Proteins: Struct., Funct., Bioinf.* **2005**, *61*, 777–794.
(15) Ojeda, P.; Garcia, M. E.; Londono, A.; Chen, N. Y. *Biophys. J.* **2009**, *96*, 1076–1082.
(16) Stan, G.; Lorimer, G. H.; Thirumalai, D.; Brooks, B. R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 8803–8808.
(17) Stan, G.; Brooks, B. R.; Lorimer, G. H.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 4433–4438.
(18) Stan, G.; Brooks, B. R.; Thirumalai, D. *J. Mol. Biol.* **2005**, *350*, 817–829.
(19) Stan, G.; Thirumalai, D.; Lorimer, G. H.; Brooks, B. R. *Biophys. Chem.* **2003**, *100*, 453–467.

(20) Lucent, D.; Vishal, V.; Pande, V. S. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 10430–10434.
(21) Songha, A. K. .; Keyes, T. *J. Phys. Chem. B* **2010**, *114*, 16908–16917.
(22) Mashl, R. J.; Joseph, S.; Aluru, N. R.; Jakobsson, E. *Nano Lett.* **2003**, *3*, 589–592.
(23) Levinger, N. E. *Science* **2002**, *298*, 1722–1723.
(24) Fenn, E. E.; Wong, D. B.; Fayer, M. D. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 15243–15248.
(25) Faeder, J.; Ladanyi, B. M. *J. Phys. Chem. B* **2000**, *104*, 1033–1046.
(26) Jamadagni, S. N.; Godawat, R.; Dordick, J. S.; Garde, S. *J. Phys. Chem. B* **2009**, *113*, 4093–4101.
(27) Xu, W. X.; Mu, Y. G. *J. Chem. Phys.* **2008**, *128*, 234506.
(28) Neidigh, J.; Fesinmeyer, R.; Andersen, N. *Nat. Struct. Biol.* **2002**, *9*, 425–430.
(29) Qiu, L.; Pabit, S.; Roitberg, A.; Hagen, S. *J. Am. Chem. Soc.* **2002**, *124*, 12952–12953.
(30) Qiu, L. L.; Hagen, S. J. *Chem. Phys.* **2004**, *307*, 243–249.
(31) Neuweiler, H.; Doose, S.; Sauer, M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 16650–16655.
(32) Ahmed, Z.; Beta, I.; Mikhonin, A.; Asher, S. *J. Am. Chem. Soc.* **2005**, *127*, 10943–10950.
(33) Barua, B.; Lin, J. C.; Williams, V. D.; Kummler, P.; Neidigh, J. W.; Andersen, N. H. *Protein Eng. Des. Sel.* **2008**, *21*, 171–185.
(34) Mok, K. H.; Kuhn, L. T.; Goez, M.; Day, I. J.; Lin, J. C.; Andersen, N. H.; Hore, P. J. *Nature* **2007**, *447*, 106–109.
(35) Doose, S.; Neuweiler, H.; Sauer, M. *ChemPhysChem* **2009**, *10*, 1389–1398.
(36) Simmerling, C.; Strockbine, B.; Roitberg, A. *J. Am. Chem. Soc.* **2002**, *124*, 11258–11259.
(37) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712–725.
(38) Naduthambi, D.; Zondlo, N. J. *J. Am. Chem. Soc.* **2006**, *128*, 12430–12431.
(39) Zhan, L.; Chen, J. Z. Y.; Liu, W.-K. *Proteins: Struct., Funct., Bioinf.* **2007**, *66*, 436–443.
(40) Day, R.; Paschek, D.; Garcia, A. E. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 1889–1899.
(41) Pitera, J.; Swope, W. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 7587–7592.
(42) Juraszek, J.; Bolhuis, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 15859–15864.
(43) Juraszek, J.; Bolhuis, P. G. *Biophys. J.* **2008**, *95*, 4246–4257.
(44) Zhang, C.; Ma, J. *J. Chem. Phys.* **2010**, *132*, 244101.
(45) Zheng, W.; Gallicchio, E.; Deng, N.; Andrec, M.; Levy, R. M. *J. Phys. Chem. B* **2011**, *115*, 1512–1523.
(46) Streicher, W. W.; Makhatadze, G. I. *Biochemistry* **2007**, *46*, 2876–2880.
(47) Wafer, L. N. R.; Streicher, W. W.; Makhatadze, G. I. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 1376–1381.
(48) Nikiforovich, G.; Andersen, N.; Fesinmeyer, R.; Frieden, C. *Proteins: Struct., Funct., Genet.* **2003**, *52*, 292–302.
(49) Bunagan, M.; Yang, X.; Saven, J.; Gai, F. *J. Phys. Chem. B* **2006**, *110*, 3759–3763.
(50) Snow, C.; Zagrovic, B.; Pande, V. *J. Am. Chem. Soc.* **2002**, *124*, 14548–14549.
(51) Chowdhury, S.; Lee, M.; Xiong, G.; Duan, Y. *J. Mol. Biol.* **2003**, *327*, 711–717.
(52) Zhou, R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13280–13285.
(53) Paschek, D.; Nymeyer, H.; Garcia, A. E. *J. Struct. Biol.* **2007**, *157*, 524–533.
(54) Patriksson, A.; Adams, C. M.; Kjeldsen, F.; Zubarev, R. A.; van der Spoel, D. *J. Phys. Chem. B* **2007**, *111*, 13147–13150.
(55) Paschek, D.; Hempel, S.; Garcia, A. E. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 17754–17759.
(56) Kannan, S.; Zacharias, M. *Proteins: Struct., Funct., Bioinf.* **2009**, *76*, 448–460.
(57) Marinelli, F.; Pietrucci, F.; Laio, A.; Piana, S. *Plos Comput. Biol.* **2009**, *5*, e1000452.
(58) Zhuravlev, P. I.; Materese, C. K.; Papoian, G. A. *J. Phys. Chem. B* **2009**, *113*, 8800–8812.
(59) Velez-Vega, C.; Borrero, E. E.; Escobedo, F. A. *J. Chem. Phys.* **2010**, *133*, 105103.
(60) Canchi, D. R.; Paschek, D.; Garcia, A. E. *J. Am. Chem. Soc.* **2010**, *132*, 2338–2344.
(61) Canchi, D. R.; Garcia, A. E. *Biophys. J.* **2011**, *100*, 1527–1533.
(62) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668–1688.
(63) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
(64) Horwich, A. L.; Farr, G. W.; Fenton, W. A. *Chem. Rev.* **2006**, *106*, 1917–1930.
(65) Waghe, A.; Rasaiah, J. C.; Hummer, G. *J. Chem. Phys.* **2002**, *117*, 10789–10795.
(66) Werder, T.; Walther, J.; Jaffe, R.; Halicioglu, T.; Koumoutsakos, P. *J. Phys. Chem. B* **2003**, *107*, 1345–1352.
(67) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gouls, I. R.; Merz, K. M.; Fergueson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
(68) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141–151.
(69) Hukushima, K.; Nemoto, K. *J. Phys. Soc. Jpn.* **1996**, *65*, 1604–1608.
(70) Hansmann, U. H. E.; Okamoto, Y. *Curr. Opin. Struct. Biol.* **1999**, *9*, 177–183.
(71) Nymeyer, H.; Gnanakaran, S.; Garcia, A. E., *Numerical Computer Methods*, Part D; Methods in Enzymology, Vol. 383; Academic Press, Inc.: San Diego, CA, 2004; pp 119+.
(72) Garcia, A.; Herce, H.; Paschek, D. *Annu. Rep. Comput. Chem.* **2006**, *2*, 83–95.
(73) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306–317.
(74) Nose, S. *Mol. Phys.* **1984**, *52*, 255–268.
(75) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695–1697.
(76) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
(77) Miyamoto, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 952–962.
(78) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.
(79) Garcia, A. E.; Sanbonmatsu, K. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2782–2787.
(80) Hudaky, P.; Straner, P.; Farkas, V.; Varadi, G.; Toth, G.; Perczel, A. *Biochemistry* **2008**, *47*, 1007–1016.
(81) Jimenez-Cruz, C.; A., C.; Makhatadze, G. I.; Garcia, A. E. *ChemPhysChem* **2011**, doi number: 10.1039/c1cp21193E.
(82) Athawale, M. V.; Jamadagni, S. N.; Garde, S. *J. Chem. Phys.* **2009**, *131*, 115102.
(83) Vaitheeswaran, S.; Thirumalai, D. *J. Am. Chem. Soc.* **2006**, *128*, 13490–13496.
(84) Ghosh, T.; Kalra, A.; Garde, S. *J. Phys. Chem. B* **2005**, *109*, 642–651.
(85) Zhou, H. X.; Dill, K. A. *Biochemistry* **2001**, *40*, 11289–11293.
(86) Vaitheeswaran, S.; Garcia, A. E. *J. Chem. Phys.* **2011**, *134*, 125101.
(87) Miklos, A. C.; Li, C.; Sharaf, N. G.; Pielak, G. J. *Biochemistry* **2010**, *49*, 6984–6991.
(88) Schlesinger, A.; Wang, Y.; Tadeo, X.; Millet, O.; Pielak, G. *J. Am. Chem. Soc.* **2011**, *133*, 8082–8085.
(89) Hofmann, H.; Hillger, F.; Pfeil, S. H. .; Hoffmann, A.; Streich, D.; Haenni, D.; Nettels, D.; Lipman, E. A. .; Schuler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 11793–11798.