Accelerated folding in the weak hydrophobic environment of a chaperonin cavity: Creation of an alternate fast folding pathway

A. I. Jewett*, A. Baumketner**, and J.-E. Shea***

*A Department of Chemistry and Biochemistry, University of California, Santa Barbara, CA 93106, and †Institute for Condensed Matter Physics, 1 Svientsitsky Street, Lviv 79011, Ukraine

Edited by Peter G. Wolynes, University of California at San Diego, La Jolla, CA, and approved July 19, 2004 (received for review February 1, 2004)

Recent experiments suggest that the folding of certain proteins can take place entirely within a chaperonin-like cavity. These substrate proteins experience folding rate enhancements without undergoing multiple rounds of ATP-induced binding and release from the chaperonin. Rather, they undergo only a single binding event, followed by sequestration into the chaperonin cage. The present work uses molecular dynamics simulations to investigate the folding of a highly frustrated protein within this chaperonin cavity. The chaperonin interior is modeled by a sphere with a lining of tunable degree of hydrophobicity. We demonstrate that a moderately hydrophobic environment, similar to the interior of the GroEL cavity upon complexation with ATP and GroES, is sufficient to accelerate the folding of a frustrated protein by more than an order of magnitude. Our simulations support a mechanism by which the moderately hydrophobic chaperonin environment provides an alternate pathway to the native state through a transiently bound intermediate state.

Chaperones are large biomolecules that assist the folding of proteins under cellular conditions that are nonpermissive for folding. They are involved in the folding of ~20–30% of the proteins in Eubacteria, Eukarya, and Archaea (1). Most chaperones seem to contain a high percentage of nonpolar residues on their surface, which allow them to bind nonspecifically to exposed hydrophobic groups on the surfaces of a diverse set of misfolded protein substrates. The precise mechanism by which chaperones increase folding rates and yields is not understood and is further complicated by the fact that the mechanism for a given chaperone may very well be substrate-specific.

The complexity of chaperonin mechanisms is apparent in the case of the GroES/EL complex, one of the best-studied chaperonin systems. GroEL is composed of 14 subunits, 57 kDa each, arranged into two stacked back-to-back rings. Misfolded protein intermediates are believed to bind to a concentrated group of exposed hydrophobic residues near the opening of one of the GroEL cavities (the apical domain). Substrate binding is followed by binding of ATP and the co-chaperone GroES, which seals the opening to form a hollow container large enough to enclose the protein substrate. The process of binding to ATP and GroES also triggers a conformational change in GroEL that buries many of its hydrophobic residues (2–4). The protein remains sequestered within the cavity until hydrolysis of ATP triggers the release of GroES and the enclosed substrate protein. If the protein remains misfolded, it will quickly bind to another chaperone, and the process will repeat itself until the protein either folds to its biologically active native state or aggregates.

Experiments have yielded conflicting results and have led to the development of both “passive” and “active” theories for chaperonin-mediated folding. The passive mechanism suggests that the role of the chaperonin is merely to remove the misfolded protein from the crowded cellular environment (conducive to aggregation) and to place it in a cage where it can fold uncontaminated (5–8). The active mechanisms are of two sorts. The iterative annealing model (IAM) stipulates that the folding of frustrated proteins is accelerated through multiple rounds of substrate binding and release from GroEL. According to this model, conformational changes occurring in GroEL due to GroES binding disrupts any incorrect contacts that may be present in the substrate and frees it from kinetic traps (9–15). Multiple cycles of binding/release give a misfolded protein more chances to escape from kinetic traps and fold correctly, leading to an overall acceleration in folding rates. A prediction of the IAM is that the chaperonin is unable to accelerate the folding of unfurtrasted proteins with first order kinetics because, by definition, these proteins are not trapped in long-lived (relative to the folding time) intermediate states (13).

The second active model proposes that encapsulation in the chaperonin cavity can lead to increased folding rates by eliminating local energy minima that would otherwise serve as kinetic traps for the protein. This “smoothing of the energy landscape” model was suggested by Hartl (16) to explain chaperonin-mediated folding-rate increases that could not be attributed to multiple rounds of binding and release of the substrate protein in accordance with the IAM described above. In particular, Hartl noted that RuBisCo, a stringent GroEL substrate folds faster when trapped inside SR-1 (a mutant of GroEL that cannot unbind from GroES) than it does in a dilute bulk environment (16). It has been proposed that this scenario can be reconciled within a noncycling limit of the IAM (10, 12, 15), in which a single binding event [n = 1 (15)] would be sufficient to lead to accelerated folding events. Other experiments that show assisted folding in the absence of cycling include the work by Ewalt et al. (17), Martin and Hartl (18), and Mayhew et al. (19), who found that rhodanese (another stringent GroEL/ES substrate) does not unbind from GroEL during each cycle of ATP binding and hydrolysis. In addition, it has been observed that folding of hen lysozyme can be promoted by GroEL alone, in the absence of GroES and ATP (20). Similar results have been demonstrated for barnase (21) and for rhodanese (22).

Work in our group and in others has examined folding in a confined environment from a theoretical perspective, through simulations of proteins in a purely hydrophilic cage (23–25). This cage served as a crude model of the interior of a GroEL chaperonin complexed with ATP and GroES. Using this model, we demonstrated that encapsulation could lead to folding rate accelerations only for proteins with very modest amounts of frustration. Proteins with large amounts of frustration (i.e., the very proteins that are prone to misfolding and likely to require the help of a chaperonin to fold) were, however, found to show a decrease in folding rate at physiologically relevant temperatures. Hence, encapsulation of a misfolded frustrated protein in a purely hydrophilic cage does not lead to folding acceleration.
and can serve the role only of temporarily preventing the aggregation of this protein with other misfolded species present in the crowded cellular milieu.

In this article, we consider confinement in a more realistic chaperonin cage, one with a moderately hydrophobic rather than a purely hydrophilic lining. Indeed, visual inspection of the interior of the GroEL cavity upon complexation with the substrate and GroES reveals that the lining of the cavity is in fact moderately hydrophobic (12). We consider that a single binding event (IAM, \( n = 1 \)) has occurred and study the subsequent folding of the protein enclosed within the chaperonin cavity. This \( n = 1 \) step corresponds to the binding of the substrate protein to the hydrophobic apical domain of the GroEL molecule. Subsequent binding of ATP and GroES leads to a dramatic conformational change of the chaperonin in which the size of the chaperonin cavity increases and the hydrophobicity of its lining changes. We find that a cage with a modestly hydrophobic lining leads to a folding behavior opposite to that observed in the purely hydrophilic cage. Indeed, we observe an increase in folding rates for frustrated proteins and a decrease in folding rates for unfrustrated proteins. Our results support a mechanism by which folding is accelerated through the population of a new bound intermediate state, which provides an alternate pathway to the native state. The mechanism resembles in some respects the standard iterative annealing mechanism, but folding occurs inside the cavity and involves stochastic (rather than ATP-driven) cycles of binding and unbinding from the chaperonin.

Methods and Model

Protein Model. To investigate the folding kinetics of a protein in the vicinity of a chaperone by using simulations, the protein must be modeled in a sufficiently simple manner to allow hundreds of folding events to occur in a computationally reasonable time frame. Lattice protein models are a natural choice and were used in the seminal work by Betancourt and Thirumalai (12) to investigate the interaction of a protein with a chaperonin. In the present work, we use an off-lattice, rather than a lattice, representation of the protein and the chaperonin. This choice is motivated by a desire to describe in a more realistic manner the presentation of the protein and the chaperonin. This choice is motivated by a desire to describe in a more realistic manner the presentation of the protein and the chaperonin. In particular, we use an off-lattice four-stranded \( \alpha/\beta \) sandwich protein model developed by Honeycutt and Thirumalai (26). Peptide bond lengths and bond angles are enforced by using harmonic potentials. \( \sigma \) represents the equilibrium distance between bonded peptides. The parameter \( e_{\text{h}} \) represents the strength of a hydrophobic interaction (\( \sim 3 \times 10^{-3} \) atomic units). For the remainder of this article, all energies (and temperatures) will be reported in units of \( e_{\text{h}} \), and all distances will be expressed in units of the bond-length \( \sigma \).

To enforce steric hindrance and to simulate the effect of attraction between two hydrophobic monomers, a Lennard-Jones potential was chosen. The forces between nonbonded monomers are sequence-dependent. Details are provided in supporting information, which is published on the PNAS web site.

To examine how the mechanism of chaperonin-assisted folding depends on the degree of frustration of the substrate protein, we considered two variants of the protein model described above, one frustrated and one unfrustrated. The frustrated protein can fold directly to the lowest energy native state in which the hairpin is oriented toward the back of the plane, or visit a misfolded intermediate state that occurs when the \( \beta \)-hairpin forms incorrectly early on during the folding process. (See Fig. 1.)

Thermodynamics of the frustrated protein were obtained by the weighted histogram analysis method (29). The collapse temperature \( T_c \) (as determined by the maxima of the specific heat) was found to be \( 0.4025 e_{\text{h}}/k_B \) in the bulk, and \( 0.42 e_{\text{h}}/k_B \) when the protein was confined to a purely repulsive sphere of radius \( R = 6 \sigma \). Details of the confining potential are presented in Chaperonin Model.

The folding temperature was estimated to be around \( T_f = 0.25 e_{\text{h}}/k_B \). This temperature was obtained by considering the temperature at which the free energy displayed two equal free-energy minima in the nonnative and folded states. The temperature at which the protein folded most quickly, \( T_m \), was measured to be \( T_m = 0.385 e_{\text{h}}/k_B \) in the bulk, and \( T_m = 0.41 e_{\text{h}}/k_B \) under confinement. We note that \( T_m > T_f \), a signature of frustrated proteins (23).

The unfrustrated model has strong torsion angle forces in the turn regions that favor the correct \( \beta \)-hairpin orientation and destabilize the misfolded state. It is a fast folder with single exponential folding kinetics and nearly identical folding and collapse temperatures that are sharply defined (\( T_c = 0.385 e_{\text{h}}/k_B \), \( T_f = 0.39 e_{\text{h}}/k_B \), \( T_l = 0.41 e_{\text{h}}/k_B \) when confined within a hydrophilic sphere of radius \( R = 6 \sigma \)). For the unfrustrated model, the temperature of optimal folding was \( T_m = 0.33 e_{\text{h}}/k_B < T_c \). The native state of this model has the same fold as the frustrated protein, as well as the same orientation of its \( \beta \)-hairpin.

Chaperonin Model. Our earlier work (23) modeled the chaperonin by a hydrophilic sphere. Here, we consider a more realistic depiction of the chaperonin interior in which the cage has a moderately hydrophobic interior. To study the effect of the degree of hydrophobicity of the interior of the cage on folding rates, we introduce a potential for the wall with a tunable degree of hydrophobicity \( h \).
The wall is modeled by monomers with a surface density of 1/σ², each one exerting a Lennard-Jones force on each monomer in the substrate of the form 4ε[(σ/a)² - h(σ/a)⁴] for hydrophobic peptides, and 4ε₀[(σ/a)² - h(σ/a)⁴] for hydrophilic or neutral peptides. The parameter h represents the degree of hydrophobicity of the interior surface of the chaperonin cage. A wall with a purely hydrophobic lining has an h value of 1 whereas a purely hydrophilic lining has an h value of 0. Further details are given in supporting information.

The radius of the sphere was chosen to be R = 6a, large enough to accommodate the native state, yet small enough so that confinement effects were significant. An earlier study in our group (23) demonstrated that confinement in a hydrophilic container slows the folding of our protein model by discouraging the complete opening of the a-hairpin. As before, the protein displays bi-exponential folding kinetics when confined within a weakly hydrophobic sphere (h = 0.45). As h is increased from 0 to 0.45, the average folding time decreases as the lifetime of the slow track decreases. However, when the protein is contained within a strongly hydrophobic sphere (h ≥ 0.5), average folding times increase, and the folding kinetics become single-exponential as escape from the bound state becomes the rate-limiting step for folding. The mean folding time is smallest at the intermediate value h = 0.475.

The yield of folded protein with time is given in Fig. 3 for a variety of hydrophobicities. The protein has biexponential kinetics both in the bulk and when confined to a purely hydrophilic sphere (h = 0). In the bulk (green curve), the protein folds by means of a “fast” mechanism 43% of the time, with a rate of k_F = (1,284τ⁻¹). The remaining 57% of the time, the protein folds by means of a “slow” mechanism (k_S = (4,099,873τ⁻¹). The overall mean first passage time in the bulk is 2,449,000τ. During numerous folding experiments, we examined the protein when it was clearly in the slow track. (i.e., when it had not folded in under t = 500,000τ). We found that it contained incorrectly formed a-hairpins, resembling the misfolded structure shown in Fig. 1. Confinement within a hydrophilic (purely repulsive) sphere of radius 6a, (h = 0, R = 6a) does not greatly affect the fast-folding channel, but slows down the slow-folding channel by a factor of almost two (k_S = (7,502,064τ⁻¹), leading to a new longer mean first-passage time of 4,515,000τ (black curve). Confinement in the hydrophilic cage seems to slow the folding of the frustrated protein by discouraging the complete opening of the a-hairpin, hence preventing escape from the misfolded state.

**Free Energy Surfaces for Folding: Rate Acceleration Through the Population of a New Bound Intermediate State.** We can gain further insight into the effects of hydrophobic confinement on folding by considering how confinement modifies the folding free-energy surfaces.

The free-energy surfaces for the frustrated protein at the folding temperature T_f in the bulk and in a weakly hydrophobic cage (h = 0.475) are given in Fig. 4 as a function of the RMSD (RMSD) with respect to the misfolded state (RMSD_M) and the RMSD with respect to the native state (RMSD_N). The free energy surface for the bulk case (Fig. 4a) shows two main basins at (RMSD_M = 0.5, RMSD_N = 1.3) and (RMSD_M = 1.3, RMSD_N = 0.25) corresponding to the compact misfolded and...
natural state, respectively. These minima are separated by a barrier. The free-energy surface for the confined protein in a weakly hydrophobic cage (Fig. 4b), on the other hand, shows three main basins at (RMSD$_M$ = 0.5, RMSD$_N$ = 1.3), (RMSD$_M$ = 1.3, RMSD$_N$ = 0.25), and (RMSD$_M$ = 2.25, RMSD$_N$ = 2.2), corresponding to the compact misfolded state, native state, and a new bound state, respectively. Representative conformations of each ensemble are shown in Fig. 4b. The free-energy surfaces clearly reveal that a new route to the folded state is provided by the presence of the hydrophobic wall through the population of a new intermediate bound state. In order for a conformation to go from the compact misfolded state to the native state, the protein must rearrange its $\beta$-sheet, which is incorrectly formed in the misfolded state. This process is facilitated by the presence of the moderately hydrophobic chaperonin wall. The implications of this new intermediate state for folding rates are discussed in Discussion and Conclusions.

As the hydrophobicity of the wall ($h$) is increased from 0 to ≈0.45, the lifetime of the slow track decreases (see the black and blue curves from Fig. 3b). This trend is due to the fact that the rate of binding to the chaperonin wall increases as $h$ increases and that transient binding to the chaperonin wall depopulates the misfolded state (12). At large hydrophobicities ($h > 0.475$), on the other hand, folding rates decrease with increasing $h$ and the protein starts to display single-exponential folding kinetics. This result is because, at large $h$, the protein immediately binds to the chaperonin wall before it has a chance to collapse to either the native or misfolded states. For large $h$, escape from the bound state becomes the rate-limiting step.

**Stochastic Cycling Inside the Moderately Hydrophobic Cavity.** Further inspection of the population of the bound intermediate state over time reveals that, at hydrophobicities around 0.475, the hydrophobic attraction is strong enough to quickly bind the substrate to the wall, but sufficiently weak to allow natural thermal excitations to free the substrate from the wall. The resulting cycle of binding and rebinding is similar in spirit to the ATP-driven GroEL/ES binding/rebinding chaperonin cycle hypothesized by Todd et al. (9) for the iterative annealing mechanism, with the significant difference that our cycling occurs within the cavity and is promoted by thermal fluctuations rather than ATP. Evidence of this mechanism can be found in Fig. 5, where the energy of interaction between the chaperone and the substrate [$U_{cap}(t)$] is plotted as a function of time in a static hydrophobic cage ($h = 0.475$) (green curve).

We note that, if the mechanism for acceleration of folding for a static weakly hydrophobic container truly involves “cycling,” one can predict that a static weakly hydrophobic container will not accelerate the folding of substrates with simple two-state, exponential folding kinetics. We tested the dynamics of folding of the unfrustrated model near its folding temperature while confined in a weakly hydrophobic container (Fig. 2h). As expected, folding of the unfrustrated model was not accelerated by the presence of a hydrophobic surface. Instead, folding rate decreased as the strength of attraction to the wall increased. As in the case of the frustrated model, folding effectively came to a halt when the hydrophobicity of the wall exceeded $h > 0.5$.

**Discussion and Conclusions.** Our results demonstrate that it is possible for a chaperonin molecule such as GroEL to actively accelerate protein folding inside a cavity in the absence of ATP-driven cycling, even under conditions where confinement alone hinders folding (see Fig. 3). The frustrated protein in the mildly hydrophobic cage spends the majority of its time in one of three states: misfolded, bound to the chaperone, and native. To fold, a misfolded protein has to spontaneously rearrange itself into a shape that is commensurate with folding, and this process may require breaking numerous incorrect contacts, requiring the protein to climb over an enormous energetic barrier. A chaperone with a weakly hydrophobic surface could circumvent that barrier by providing an alternate

**Fig. 5.** The energy of interaction between cage and substrate [$U_{cap}(t)$] plotted as a function of time at the folding temperature for a static hydrophilic cage ($h = 0$) (in yellow) and a weakly hydrophobic cage ($h = 0.485$) (in green). At $h = 0.485$, $U_{cap}(t)$ flips back and forth between $-0$ and $-14k_B$. This plot illustrates that a weakly hydrophobic surface can cause cycles of binding and release of substrate (in the absence of ATP).
observe that the protein can undergo several rounds of stochastic cycling within the chaperonin cage, “iterating” itself to the native state. This mechanism bears some resemblance to the iterative annealing mechanism, with the critical differences that a bound intermediate state (rather than an unfolded state) is generated and that the cycling occurs in the cage and is thermally rather than ATP driven.

A consequence of this mechanism is that only the folding of proteins with trapped, multieponential kinetics can be accelerated. An additional requirement is that the strength of the attraction to the surface must lie in a range where both binding and spontaneous unbinding can occur frequently. This conclusion is supported by experimental evidence that some substrates can unbind from chaperones without the aid of ATP. For example, GroEL can spontaneously dissociate from barnase with a rate constant near 0.4 s\(^{-1}\) (35) [faster than the rate of ATP hydrolysis (36)]. The rates of spontaneous binding and release must both be faster than the lifetime of well-populated trapped misfolded intermediates to significantly accelerate folding.

In their pioneering work using lattice models to investigate chaperonin-mediated protein folding, Betancourt and Thirumalai (12) performed a study of several possible means by which chaperonins could accelerate folding rates. Their studies revealed that hydrophobic confinement can lead to accelerated folding, although to a significantly lesser extent than suggested by our off-lattice simulations. Later work (15) suggested that folding within a chaperonin cage can be viewed as the “nongelling” limit of the IAM, in which only a single binding event (\(n = 1\)) occurs. Our simulations examine the folding of the protein subsequent to this \(n = 1\) step and reveal and quantify the role of the nature of the lining of the chaperonin cage in determining the degree of folding rate enhancement that can be achieved at physiological temperatures.

In what follows, we discuss the implications of the IAM \(n = 1\) mechanism and contrast rate accelerations arising from (i) the \(n = 1\) limit of the IAM (i.e., a protein folding after a single round of denaturation), (ii) the effect of confinement after the \(n = 1\) IAM step (i.e., a protein folding in a purely hydrophilic cage after a single round of denaturation), and (iii) the effect of confinement in a moderately hydrophilic cage after the \(n = 1\) IAM step.

The IAM \(n = 1\) step involves the tight binding of the protein to the highly hydrophobic apical region of GroEL. Subsequent binding of ATP and GroES leads to a dramatic rearrangement of the apical domain, including a change in hydrophobicity, and the shape and size of the cavity. In this step, the chaperonin acts as an unfoldase, mechanically unfolding the substrate protein. After this initial step, once the protein finds itself encapsulated, it encounters a different hydrophobic environment than the one presented at the apical domain. When folding occurs inside the cage, the chaperonin no longer plays the role of a simple unfoldase.

We first turn to the case of IAM \(n = 1\) in the absence of subsequent confinement. Consider \(\phi\) to be the fraction of proteins folding after the fast tract and \(1 - \phi\) the fraction of proteins folding according to the slow tract. The unfoldase activity of the \(n = 1\) ATP-driven step will affect only the molecules on the slow tract. In the best case scenario, all of the slow-tract proteins will be unfolded and allowed to repartition through the fast and slow tract to the native state. The partitioning will remain the same, with \(\phi\) molecules folding quickly and \((1 - \phi)\) molecules folding slowly. If we consider the folding time before the \(n = 1\) step to be \(t_f\), then the maximal rate benefit from the \(n = 1\) step will be \((1 - \phi)t_f\,\phi\), with an overall folding time reduction of \((1 - \phi)\). In our model, this single unfolding event would correspond to a decrease in folding time by a factor of 0.57.
Let us now examine the effect of confinement after this \( n = 1 \) step. This effect can be isolated by considering folding in a purely hydrophilic cage. We have shown in our earlier work (23) that proteins can experience a rate enhancement only in a purely hydrophilic environment at temperature above the temperature of fastest folding time \( T_m \). For frustrated proteins, this temperature lies above the folding temperature \( T_f \) because such proteins experience significant trapping in misfolded conformations at temperatures where the folded state is significantly populated. For our particular model protein, confinement in a hydrophilic cage at \( T_f \) increases folding times by a factor of two, essentially negating the rate gain from the frustration effect. This process leads to a folding rate acceleration of an order of magnitude greater than achieved by the IAM \( n = 1 \) step alone. The \( n = 1 \) IAM step is necessary to create a lining of the chaperonin cage conducive to the formation of a bound state for stochastic cycling but does not (at least in our model) lead to as significant a rate enhancement as thermally induced cycling within the chaperonin walls through a bound intermediate state.

The picture observed here may have broader applications than simply to the GroEL chaperonin. In particular, it may help explain how the folding of certain proteins can be accelerated by other types of non-ring chaperones that do not depend on ATP, for example, the small heat shock proteins (sHSP) (12, 37). Our mechanism is compatible with the observation that sHSP chaperones promote protein folding in the absence of ATP and can spontaneously associate and dissociate from their substrates (38). In addition, we speculate that, under certain conditions, some frustrated proteins can fold faster in a crowded cellular environment than in the bulk. In the cellular milieu, the protein may experience many weak attractions with other molecules that compete with the formation of incorrect intra-protein interactions. Provided that the protein does not aggregate with itself, these weak interactions could play the role of a weakly hydrophobic chaperonin. Correctly folded proteins would not be very susceptible to unfolding by this mechanism because they would presumably have fewer exposed hydrophobic residues available to be tempted by a modest hydrophobic environment surrounding them.

This work was supported by the David and Lucile Packard Foundation and the National Science Foundation through Career Award 0133504.