Inside the chaperonin toolbox: theoretical and computational models for chaperonin mechanism

IOP PUBLISHING PHYSICAL BIOLOGY

Phys. Biol. **6** (2009) 015003 (8pp) doi:10.1088/1478-3975/6/1/015003

Inside the chaperonin toolbox: theoretical and computational models for chaperonin mechanism

Del Lucent[1], **Jeremy England**[2] **and Vijay Pande**[1,3,4]

[1] Biophysics Program, Stanford University, Stanford, CA, USA
[2] Department of Physics, Stanford University, Stanford, CA, USA
[3] Departments of Chemistry & Structural Biology, Stanford, CA, USA

E-mail: pande@stanford.edu

Received 3 May 2008
Accepted for publication 21 July 2008
Published 10 February 2009
Online at stacks.iop.org/PhysBio/6/015003

Abstract

Despite their immense importance to cellular function, the precise mechanism by which chaperonins aid in the folding of other proteins remains unknown. Experimental evidence seems to imply that there is some diversity in how chaperonins interact with their substrates and this has led to a number of different models for chaperonin mechanism. Computational methods have the advantage of accessing temporal and spatial resolutions that are difficult for experimental techniques; therefore, these methods have been applied to this problem for some time. Here we review the relevant computational models for chaperonin function. We propose that these models need not be mutually exclusive and in fact can be thought of as a set of tools the chaperonin may use to aid in the folding of a diverse array of substrate proteins. We conclude with a discussion of the role of water in the chaperonin mechanism, a factor that until recently has been largely neglected by most computational studies of chaperonin function.

1. Introduction

Although much progress has been made in recent decades in the study of protein folding, it remains unclear how proteins succeed in folding in crowded, heterogeneous intracellular microenvironments that often differ dramatically from the favorable *in vitro* conditions under which most protein folding studies are conducted. A cell must be able to maintain the stability (and function) of its proteome at a variety of temperatures, and at varying conditions of ionic strength, pH and reduction potential. For this purpose, cells have evolved molecular chaperones, a class of molecules responsible for preventing protein aggregation, facilitating productive folding and targeting malfunctioning proteins for degradation.

Within the broad class of molecular chaperones are a special group known as the chaperonins. Chaperonins are large, barrel-shaped protein complexes that have the ability to bind and engulf unfolded/misfolded proteins [1–4]. Chaperonins are indispensable to cellular function, interacting with 10% of all cellular proteins [2, 4]. Thirteen of these proteins are essential ones that cannot fold properly without the aid of chaperonins [5]. Thus, chaperonins are required for cellular viability. Also, one can imagine that having a general protein folding machine may relieve some of the evolutionary strain of selection based on the stability of the folded protein and allow selection based on function.

Chaperonins are composed of a number of subunits. Although the exact number and sequence of the subunits differ from prokaryotes to archea to eukaryotes, the relative arrangement is the same. The subunits are arranged to form two rings stacked back to back. The central cavity formed by each ring constitutes an active site in which a substrate protein is bound, encapsulated and released via a complex multi-step cycle. Initially the chaperonin is in an open conformation where the central cavity is largely hydrophobic in nature. In this conformation, it binds an unfolded/misfolded substrate protein. Upon binding a molecule of ATP to each subunit, a large conformational change occurs in the

[4] Address for correspondence: James H Clark Center, 318 Campus Drive Stanford, CA 94305.

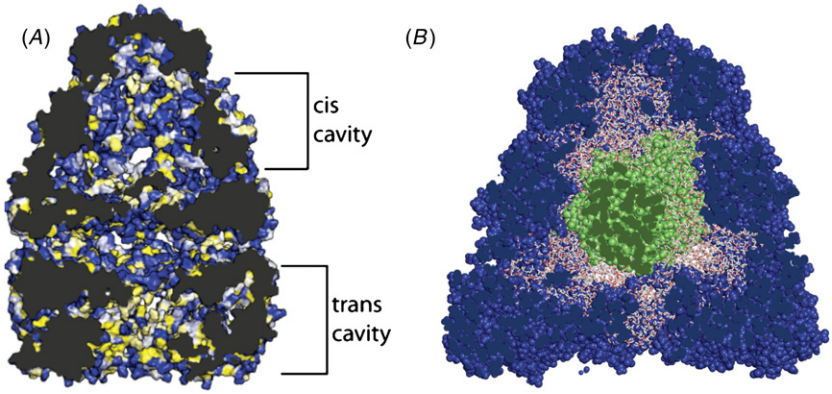

Figure 1. (*A*) Prokaryotic chaperonin GroEL sliced in half, with residues colored according to the Kyte–Doolittle hydrophobicity scale [59] with blue being most hydrophilic (−4.5) and yellow being most hydrophobic (4.5). This image comes from pdb code 1PCQ [60] with the unstructured c-terminal GGM repeats modeled in. (*B*) Cis chamber of GroEL (blue) encapsulating substrate protein MBP (1MPD [61]) in its native conformation (green). Highly confined TIP4P water can be seen surrounding the substrate protein.

chaperonin, expanding the cavity volume and changing the nature of the cavity surface from hydrophobic to hydrophilic (figure 1(*A*)). At this stage, the cavity is sealed with the substrate protein trapped inside. For the prokaryotic chaperonin GroEL, a co-protein (GroES) acts as a lid to seal the cavity. In the archaeal chaperonin (thermosome) and the eukaryotic chaperonin (TRiC), the lid is built into the chaperonin and a co-protein is not needed. The amount of time the protein remains encased in the cavity is limited by the hydrolysis of ATP. Once ATP has been hydrolyzed, binding of ATP and substrate protein in the opposing ring causes the closed ring to open, releasing ADP and the substrate protein. At this point, the protein may or may not be correctly folded.

The exact mechanism by which chaperonins facilitate folding of substrate proteins is unknown. There have been three major models proposed for the chaperonin mechanism, but it would seem plausible that all are correct to some extent. The first mechanism is the Anfinsen cage model; where the chaperonin cavity is thought to do nothing other than prevent aggregation (i.e. allow the substrate protein to fold in infinite dilution) [6]. The second is the iterative annealing model, where the chaperonin aids folding of kinetically trapped proteins by acting as an unfoldase, allowing proteins to be quickly removed from trapped conformations and have additional attempts at productive folding [7]. Finally, it has been proposed that the chaperonin acts as a foldase, where the microenvironment felt by the confined protein induces folding to take place more rapidly than it would in the cytosol [8].

For each of the above three mechanisms there is both a theoretical framework and experimental/simulation evidence for support. When considering that the chaperonin interacts with a great variety of substrates, it is not surprising that the chaperonin may have a number of 'tools' in its folding toolbox; some used for some proteins and some used for others. In support of this idea, none of the existing models are necessarily mutually exclusive.

The problem with confirming the above models is that they require a detailed knowledge of what is happening on the molecular level. With current experimental methodologies, it is difficult to spectroscopically observe the substrate protein while it is interacting with the chaperonin. For this purpose, simulation techniques can be employed to gain insight on timescales and spatial scales that are difficult for most experiments. The best use of simulation methodology is to complement experimental observations by providing access to details that cannot be readily visualized in experiments. In any case, both simulations and experiments will greatly benefit from theoretical models that can explain current results and make predictions for future work. Here we review both the theoretical models and the accompanying simulations that have attempted to shed some light on the mechanisms of chaperonin-assisted folding.

2. Folding by unfolding: iterative annealing model

The hydrophobic surface presented by the open chaperonin has been shown to bind exposed hydrophobic residues of non-native peptides [9, 10]. Numerous studies of chaperonin-assisted folding have shown that this leads to an unfolding of the bound protein [1]. How could this action lead to productive folding? It was an attempt to answer this question that brought about the first theoretical studies of chaperonin-mediated folding.

Gulukota and Wolynes published the first theoretical treatment of the role of chaperonins in protein folding in 1994 [11]. In their model, they proposed that when a substrate protein is bound by the chaperonin its dynamics take place on a new energy surface uncorrelated with the regular landscape on which the protein would normally fold. When released, the substrate protein would be at a different location on the original energy surface (and thus could make a new attempt at productive folding). They showed that if chaperonins were biased such as to preferentially bind non-native conformations (which presumably would be more likely to suffer from kinetic trapping and aggregation) a net increase in the yield of the folded protein would result from interacting with the chaperonin. They referred to their model as kinetic proofreading.

Todd *et al* refined this idea two years later [12]. Rather than assuming the substrate protein explores a different energy surface when bound to the chaperonin, they based their argument on a balance of rates. Assume that a protein can fold productively by a fast folding pathway with rate k_1 and misfold/aggregate with rate k_2. Once aggregated, the protein is kinetically trapped (i.e. reaches the native state with rate k_3, which is very slow). If $k_2 \gg k_1$, most proteins will misfold and the overall folding rate will be determined by k_3. If however, a chaperonin can bind and unfold kinetically trapped states with rate k_H (which is faster than k_3), then this new rate will determine the overall rate of folding. Both the models proposed by Gulokota and Wolynes, as well as Todd *et al* take advantage of the experimental observation that repeated iterations of the chaperonin reaction cycle are frequently needed to produce a reasonable yield of folded substrate protein. Thus, this class of models has been dubbed iterative annealing mechanisms.

The earliest simulations of chaperonin-assisted folding (much like the earliest simulations of protein folding *in vitro*) involved the usage of lattice models. These models treated the amino acids of a protein as nodes on a cubic lattice. The various conformational states of the protein are sampled using Monte Carlo techniques. Random sequences of amino acids are enumerated and foldable sequences are identified as those that reach the native state (global energy minimum) in a reasonable number of Monte Carlo steps. Individual interaction energies can come from a simple coarse-grained force field classifying residues into two or three classes (e.g. polar/hydrophobic) [13] or the energies can be drawn from some knowledge-based potential [14].

The iterative annealing mechanism was first examined in silico by numerous groups using this type of methodology. Chan and Dill, using a simple HP interaction potential and a 2D lattice model, were able to show that kinetically trapped conformations can be rescued by sequestration onto a hydrophobic surface (which represents the chaperonin in its open conformation) [15]. They go on to show that if the rates of binding and release from the chaperonin are sufficiently fast, the overall yield of folded protein is increased.

Similarly Sfatos *et al* showed, using a 3D lattice model and designed sequences, that well-designed sequences (sequences that have a very stable, compact native fold) experience a moderate increase in the yield of the native protein while poorly designed sequences (where the native fold is slightly less compact and stable) experienced a large increase in yield [16]. This discrimination in enhancement was hypothesized to have evolutionary significance: if a protein that carries out an important function is less energetically stable than average, a chaperonin might make it nevertheless possible for changes in the protein's sequence to be selected based partly on function rather than purely on stability.

In the case of Chan *et al*, the chaperonin is represented as a fixed hydrophobic surface, while in Sfatos *et al* it is not explicitly represented (the activity of the chaperonin was represented by generating an unfolded sequence in a random conformation after a set number of Monte Carlo steps). These studies did not address the question of how the affinity of the chaperonin for the substrate will affect folding yields. Betancourt and Thirumalai later published a study in which the 'chaperonin' was a cubic box with tunable surface hydrophobicity [17]. They found that there is an optimal level of affinity for the misfolded substrate protein at which folding yield is maximized. The explanation of this finding is intuitive: the chaperonin must bind misfolded conformations tightly enough to induce unfolding, but loosely enough to allow refolding on a reasonable timescale.

These pioneering studies were the first attempts to use simulation methods to understand chaperonin biology and helped to bolster the iterative annealing hypothesis. The advantage of these simple models is that they allow complete enumeration of phase space, such that thermodynamic quantities can be calculated exactly. Nonetheless, there are significant disadvantages in the use of simplified lattice models to represent real proteins. Unfortunately, the models are only caricatures of real proteins and do not contain secondary structure. They cannot predict detailed structural features of proteins (the features that give proteins their function). Additionally because lattice models use simplified move sets rather than integrating real equations of motion, they cannot predict kinetic properties. Many of the above-mentioned studies have attempted to alleviate this by fitting Monte Carlo steps to existing kinetic models. These sorts of fits are completely dependent on the underlying kinetic model and may not reflect the barriers in the free-energy surface. The rates predicted by lattice models may be useful to guide further study but are likely not quantitatively useful.

In an attempt to study the same phenomena using more details potentially relevant to chaperonin function, Jewett, Baumketner and Shea simulated the folding of a small alpha/beta sandwich peptide confined to a cavity of tunable hydrophobicity with a coarse-grained off-lattice model and Langevin dynamics [18, 19]. This model allows the adoption of secondary structure by permitting the backbone to move freely (as opposed to being confined to points on a lattice) and uses spheres designated as hydrophobic, hydrophilic or neutral in place of amino acids. The force field governing the movement of this model resembles a molecular mechanics force field in that it has potential energy terms for bond stretching, bending, backbone torsional motion and pairwise interactions.

Using this model Shea and coworkers found that when the confining sphere is moderately hydrophobic (attractive to hydrophobic residues), misfolded conformations would transiently bind and unbind the surface. This binding and release breaks incorrect contacts between residues and effectively unfolds the protein. What differentiates this from other iterative annealing models is that the cycling of binding and release is not mediated by a timed change in the confining surface from hydrophilic to hydrophobic, but rather keeps a moderate hydrophobic nature at all times. As this is inconsistent with what is known about the structure of the chapeornin cavity (it cycles between hydrophobic and hydrophilic states on a timescale determined by ATP hydrolysis) they have argued that this mechanism could be used by other chaperones (not chaperonins) and even chemical

chaperones as a mechanism of enhancing protein folding. They have expanded on this in subsequent work by modeling the chaperone as a moderately hydrophobic sphere that is added as a co-solute to the simulation [20].

Cheung and Thirumalai also have applied more sophisticated simulation methodologies to iterative annealing. They used a more complex coarse-grained model (this time for a WW domain peptide) where each amino acid is represented by two beads (an alpha carbon and a side chain) and the Hamiltonian contains terms for hydrogen bonding, bond bending and stretching, backbone torsional rotation, as well as a statistical potential for governing side-chain pairwise interactions [21]. The chaperonin is modeled as a confining spherical potential with tunable hydrophobicity (a scalable attraction between the confining wall and hydrophobic residues similar to the method used by Shea and coworkers) but the potential is allowed to vary as a function of time in order to simulate multiple iterations of the chaperonin reaction cycle. Using this model, Cheung and Thirumalai construct a phase diagram of states realized inside their model of the chaperonin cavity. With this analysis, they report that at temperatures below the folding temperature, varying the hydrophobicity will cause the protein to populate a molten globule state, the native state, and a surface absorbed state similar to that seen by Shea and coworkers. By combining this with the dynamic cycling of their model chaperonin, Cheung *et al* find that there is both an optimal hydrophobicity and an optimal cycling rate for enhancing the population of the native state.

Although these studies represent a seminal contribution to the study of chaperonin-assisted folding, they all seem to suggest essentially the same mechanistic concept: kinetically trapped substrate proteins can be rescued by moderately avid binding to the hydrophobic walls of the open chaperonin. There is both an optimal hydrophobicity and an optimal cycle time, which will maximize folding yield. If either parameter is far from the optimum, the protein will spend too much time either bound to the chaperonin or will not have enough time to unfold. Since none of the existing simulations has been able to recapitulate folding rates of actual proteins, nor have these studies been able to show that these optimal affinities and loading rates correspond to experimentally observed rates of chaperonin cycling, there is clearly room for future work to make a more direct connection to experiment in these areas.

Tehver and Thirumalai have recently created an analytical kinetic model, which is predictive if parameterized from experimental data (i.e. if one knows the folding rate, misfolding rate and affinity for the chaperonin one can calculate the native state yield) [22]. Tehver and Thirumalai show in the steady-state solution of their model that a hypothetical mutation, which maximizes substrate turnover through the chaperonin will also maximize folding yields. Jewett and Shea have also recently proposed a simple kinetic model considering the effects of a mutation that increased chaperonin cycling [23]. Their results agree with that of Tehver and Thirumalai in that such a mutation can increase the yield of folded protein under typical experimental conditions (where the concentration of protein is relatively low). They also consider that under conditions of very high substrate concentration, such a mutation can be detrimental, by increasing the amount of time unfolded protein remains in aggregation favorable conditions (i.e. outside of the chaperonin cavity). Nonetheless, both of these studies show the value of simple analytical models when making predictions about biophysical mechanisms.

3. Stimulated folding: landscape modulation

Although the iterative annealing hypothesis represents an important contribution to our understanding of chaperonin mechanism, it alone cannot explain all of chaperonin-assisted folding. The transient binding of misfolded proteins to hydrophobic surfaces in an ATP dependent manner is a characteristic that is not unique to chaperonins (for example Hsp70, a small monomeric molecular chaperone, can bind and unfold misfolded proteins in an ATP dependant manner) [2]. If iterative annealing were the whole story, why would nature have evolved such a large and complicated multimeric complex with a large sealed cavity? Furthermore, why is it that a number of essential proteins absolutely require chaperonins to fold [5] and yet silencing of other chaperones is less destructive by comparison [24]?

Perhaps the simplest mechanism for chaperonin action is the Anfinsen cage model or the 'infinite dilution model'. According to this model, the only purpose of the closed chaperonin cavity is to prevent aggregation by encasing the unfolded protein in a chamber that prevents association with other unfolded proteins. Although this mechanism may very well play a role for all proteins, many substrates cannot fold without the aid of chaperonins, even under conditions of extreme dilution [8, 25]. In addition, the Anfinsen cage model cannot explain why some substrates experience a large increase in the folding rate from a single round of chaperonin enclosure [8, 26]. This foldase activity cannot be explained by an infinite dilution model. Instead, the closed chaperonin cavity must somehow create a microenvironment that is conducive to folding. The various explanations for this are what we refer to as landscape modulation models.

The simplest idea for landscape modulation involves the entropic effect of confining a polymer in an inert cavity. Zhou and Dill proposed an analytical model of this effect in 2001 [27]. They modeled an unfolded protein as an ideal polymer and calculated a distribution of end-to-end distances by solving a classical diffusion equation (where time is replaced by polymer length). This distribution was then integrated over various closed volumes to calculate a partition function for the confined unfolded polymer. Approximating the partition function of the folded state as the volume of a sphere with a radius equal to the native protein's radius of gyration, Zhou and Dill were able to estimate the stabilization of confining a protein of a given length to a cavity of a given size and shape. The key result of this treatment is the folded state will be stabilized because of exclusion of expanded states in the unfolded ensemble that cannot fit within the confining volume. The stabilization is at a maximum when the size of the confining volume is only slightly larger than the native state. Larger volumes provide only a small stabilization while

smaller volumes destabilize the native state (as it can no longer fit in the confining volume).

Klimov *et al* were the first to explore this model through simulation. They simulated the folding of a small beta hairpin, using an off-lattice Go model (a model where the native interactions are weighted above the nonnative interactions in the Hamiltonian) confined to a spherical cavity by a repulsive potential [28]. They found that at elevated temperatures, significant structure persisted among denatured conformations and overall, the protein was resistant to denaturation. They also note that there is an optimal level of confinement for maximizing the folding rate (which is qualitatively consistent with Zhou and Dill's model).

A year later Takagi *et al* used a similar methodology for a more exhaustive examination of the confined stabilization model [29]. They confined five proteins (with diverse topologies) to various sized cylindrical cavities. Their results show that the folding temperature increases as confinement increases. They also claim that for some proteins, confinement destabilizes the denatured state so much that folding apparently proceeds in a downhill fashion. For all of the proteins they studied there seems to be a level of confinement that is optimal for folding (a result similar to Klimov *et al* and qualitatively consistent with Zhou and Dill's model). Rathore *et al* extended this work by using a simulated tempering method to achieve very well converged thermodynamic properties as a function of temperature for some of the same proteins [30]. They argue that the entropic stabilization reported by Takagi *et al* is not universal but is compensatory to an enthalpic destabilization induced by confinement. Thus, the amount by which polymeric confinement can assist folding is based not only on the size of the protein and the confining volume, but also on the topology of the protein as well.

In order to understand the process of folding from a kinetic/mechanistic point of view, it is important to understand the nature of the transition state ensemble. Cheung and Thirumalai characterized the effect of crowding and confinement on an off-lattice representation of a WW domain protein [31]. They found that confinement did not significantly alter the transition state when a Go model was used. They did find however, that the overall shape of the transition state becomes more ellipsoid-like in shape (calculated from an average inertia tensor). Cheung later followed up on this work showing that if the confinement volume was changed to match the shape of the oblate transition state, folding rates could be increased [32].

These simple polymer entropy models have received much attention, as they tend to agree (at least qualitatively) with some experimental results. In particular, Eggers and Valentine have demonstrated increased thermostability of proteins encapsulated in silica nanopores [33]. Also more recently Tang *et al* have shown that mutations to the bacterial chaperonin GroEL which are believed to effectively reduce the cavity volume have led to enhanced folding rates of some substrates [26]. In these experiments, the cavity volume was modulated by changing the length of unstructured GGM repeats found at the c-terminus of each GroEL subunit (which are believed to protrude into the confining cavity). It is unknown how these motifs interact with the substrate protein however, and it may be the case that they perturb the system in ways other than reducing the effective confining volume. For example, Farr *et al* have recently suggested that the mutations used in Tang *et al* change folding rates by affecting ATP hydrolysis rather than confinement volume [34].

Despite some qualitative agreement with experiment, all of the above listed simulations make use of a biasing potential known as a Go potential. There are some important caveats regarding the use of a Go potential that should be briefly mentioned. For the use of a Go model to be effective, it is assumed that non-native contacts are not very important for the underlying dynamics of the system. If this assumption is not tested, it is not known *a priori* whether or not the Go potential is obscuring important aspects of the folding dynamics of the system. Many simulations using coarse-grained models have found that confinement assists collapse, but not folding when a Go potential is not present [18, 19, 31]. Thus, all of the models that use Go potentials are at their essence saying that if polymer entropy and collapse are the dominant aspects of protein folding (and not enthalpic or water-mediated effects), then confinement to an inert nanopore will stabilize the native state and enhance folding. It should be noted that Zhou and Dill's model only considers these entropic effects and there is nothing to say that for many systems, effects other than polymer entropy would be important.

4. Water: the forgotten degrees of freedom?

All of the work discussed so far neglects to mention any possible role that aqueous solvent may have on chaperonin-mediated protein folding. It is understandable that most simulation methodologies neglect the effects of explicit solvent in order to eliminate the number of degrees of freedom in the problem and thus make it more computationally tractable. Still, when choosing a model, one must have significant reason to remove degrees of freedom; otherwise there is a large risk of missing the important details of the process under study.

It is well known that a major driving force in protein folding is the hydrophobic effect; the thermodynamic drive for water to minimize its free energy by minimizing the amount of exposed solute surface area that cannot participate in hydrogen bonding [35]. In most cases, this effect can be treated in a mean field manner, but there are many examples in protein folding wherein explicit inclusion of solvent degrees of freedom is essential for a complete understanding of the underlying dynamics [36–39]. Additionally, water is a polar molecule and thus experiences forces compelling it to orient itself in the presence of an external electric field. Subtle effects such as these can have large consequences when the length scales involved are small, as they are in the case of nanoscopic confinement [40–42].

The first study that included explicit water when studying folding in a confined environment was performed by Sorin and Pande [43]. In their work, they performed molecular dynamics simulations of the folding of an isolated alpha helix (Fs peptide) confined to a carbon nanotube with explicit solvent (TIP3P water). Their results showed that the folded helix

was destabilized with increasing confinement; the opposite result one would expect when applying a polymer entropy argument (and the opposite effect from that observed in a similar simulation without explicit solvent [44]). Sorin and Pande attributed this effect to the decreased translational and rotational entropy of the nanoconfined solvent. Their model indicates that in a situation where solvent is confined on the nanoscale, the free-energy cost of forming a water–protein hydrogen bond (compared to forming a protein–protein hydrogen bond) is diminished.

Lucent *et al* continued along this line of thinking with a subsequent study wherein the folding of the small globular protein villin confined to an inert spherical boundary was studied with molecular dynamics [45]. Lucent *et al* performed two sets of simulations: one where the protein was confined but the explicit solvent did not feel the confining potential, and another set wherein both the protein and the solvent felt the confining potential. They found that when only the protein was confined, the probability of folding (P_{fold}) increased. Examination of the protein conformations revealed that many of the expanded conformations found in the bulk were no longer present (a result consistent with polymer entropy models). Upon confining the solvent as well as the protein, the opposite effect was observed: the probability of folding decreased. The existence of a collapsed, misfolded state that was pushed to the solvent/confining potential boundary was observed. This implies that when solvent is confined, the free-energy penalty of disturbing the solvent hydrogen bond network outweighs the stabilization experienced by polymeric confinement.

These studies indicate the importance of explicit solvent on confined folding. However, none of these simulations can account for the experimental observation that mutations that increase the hydrophilicity of the chaperonin cavity can increase its foldase activity. Weissman and colleagues have shown via directed evolution that mutating certain residues in the chaperonin cavity from hydrophobic to polar leads to increased folding rates for green fluorescent protein [46]. Additionally Tang *et al* have recently demonstrated that mutations which reduce the number of polar residues in the chaperonin cavity cause a decrease in the folding rate of a mutant of maltose binding protein [26]. Similar results have also been shown for rhodanese and malate dehydrogenase when hydrophilic residues in the unstructured C-terminus of each GroEL subunit are mutated to hydrophobic residues [47].

One can imagine that within the tight confines of the closed chaperonin cavity, the confined water is exposed to a peculiar environment (figure 1(*B*)). On the one hand, there is an unfolded protein with exposed hydrophobic moieties. On the other hand, the closed cavity constitutes an extended polar surface. The confined solvent has to balance between two different modes of organization. Around the hydrophobic solute, the solvent will have either depleted density or reduced orientational entropy in order to maintain its hydrogen bond network and minimize its free energy. At the same time, around the hydrophilic surface, the solvent molecules will feel an electrostrictive force, increasing their density and decreasing their orientational entropy as they attempt to orient their dipoles in the field generated by the surface. A number of molecular dynamics simulations have shown that water in a peculiar environment such as this (nanoconfinement between a hydrophobic and hydrophilic surface) can mediate a repulsive force between the two surfaces [48–51]. This sort of phenomena have been observed in recent computational work by Weixin *et al* in which the confinement to a polar nanopore enhances aggregation of small nonpolar solutes [52].

Recently, England and Pande have proposed an analytical model for chaperonin-mediated foldase activity based on these principles [53]. They present a Landau-like continuum model for the free energy of bulk solvent based on spatially varying fluctuations in local density and hydrogen bond network quality. They add to this an additional free-energy term that captures interactions between the confined solvent and the confining surface. When both the substrate protein and the confining surface are hydrophobic, there is a water-mediated attraction consistent with the hydrophobic effect. When the confining surface is hydrophilic, there is a net repulsion between the surfaces. This force constitutes a greater thermodynamic drive to bury hydrophobic residues than that which would be felt in bulk. Thus, according to this model, the hydrophilic surface mediates a local enhancement of the hydrophobic effect. Although an acceleration of folding through this effect is yet to be verified by simulations, England, Lucent and Pande have recently performed a molecular dynamics study that shows that the local density of water inside all-atom chaperonin models correlates well with experimental folding rates for the GroEL mutants used in Tang *et al*. This both supports the model for a water-mediated foldase activity and elegantly demonstrates the utility of simulation in making subtle observations that would be difficult to make experimentally, yet complement experimental observables.

5. Summary and outlook

The results of the various studies described above indicate that chaperonins are complex machines that seem to make use of a variety of methods to maintain the cellular proteome. Among these are the iterative annealing models where the chaperonin acts as an unfoldase and the landscape modulation models where the chaperonin acts as a foldase. All of the existing work seems to promote the idea that these models are not mutually exclusive. Still, much needs to be done to fully understand the range of applicability of these mechanisms to different classes of substrates. Most of the reviewed work (both experiments and simulations) has been based on our knowledge of the prokaryotic chaperonin GroEL, but we should not assume complete generality of these findings. There are marked differences between prokaryotic, archaeal, and eukaryotic chaperonins, and investigation of the significance of these differences is essential for a complete understanding of chaperonin mechanism, as well as protein folding in general. Jacob *et al* have recently attempted to address this issue using simple lattice models [56]. We feel that studies that break into this sort of new ground are essential in achieving a complete understanding of chaperonin biology.

It should be noted that whether or not a specific model has limited applicability to chaperonin mechanism, it might still have great conceptual utility in other fields. One simple example is that of recent groups who are using the general principles of iterative annealing, and confinement induced stabilization to increase the effectiveness of structural refinement methods [57, 58]. One can easily imagine applications well beyond this, where engineered chaperonins or chaperonin-like molecules could be used in various health and industrial applications.

Finally, we would like to point out that all of this work serves to demonstrate the utility of simulation techniques in understanding real biological problems. Nevertheless, in order to continue to contribute to the understanding of complicated phenomena such as chaperonin-mediated protein folding we must be very critical of our models. Of course, we cannot always use detailed models that demand computational resources that are not available. Furthermore, while the most complex model is not always the best model to use in all circumstances, the choice of computational method employed in any given study is often dictated by computational expense involved, thus creating limits to which hypotheses can be addressed by computational methods. For example, choosing to ignore solvent degrees of freedom is a wise choice if it is merited; however, recent studies are revealing that solvent may play an important role and those using simpler models bear the burden of hypothesizing what the effects of leaving out such details may be. Nonetheless, the development and evaluation of the various models of chaperonin-assisted folding have progressed hand in hand with the development of better computational methods for studying protein folding. It is our expectation that as computer technology and simulation methodology continue to advance so will our ability to evaluate and refine our models of chaperonin mechanism, allowing one to not only make quantitative predictions that can be verified by experiment but then to use this validation to make observations that are difficult or impossible to achieve experimentally.

Acknowledgments

Both DL and JE contributed equally to this work. We would like to thank the NIH Nanomedicine Center (PN1 EY016525–02) for financial support. JE would like to thank the Annie and John Hertz Foundation for financial support.

References

[1] Fenton W A and Horwich A L 2003 Chaperonin-mediated protein folding: Fate of substrate polypeptide *Q. Rev. Biophys.* **36** 229–56
[2] Hartl F U and Hayer-Hartl M 2002 Molecular chaperones in the cytosol: From nascent chain to folded protein *Science* **295** 1852–8
[3] Lin Z and Rye H S 2006 Groel-mediated protein folding: Making the impossible, possible *Crit. Rev. Biochem. Mol. Biol.* **41** 211–39
[4] Spiess C, Meyer A S, Reissmann S and Frydman J 2004 Mechanism of the eukaryotic chaperonin: Protein folding in the chamber of secrets *Trends Cell Biol.* **14** 598–604
[5] Kerner M J *et al* 2005 Proteome-wide analysis of chaperonin-dependent protein folding in escherichia coli *Cell* **122** 209–20
[6] Ellis R J 1994 Molecular chaperones. Opening and closing the anfinsen cage *Curr. Biol.* **4** 633–5
[7] Weissman J S, Kashi Y, Fenton W A and Horwich A L 1994 Groel-mediated protein folding proceeds by multiple rounds of binding and release of nonnative forms *Cell* **78** 693–702
[8] Brinker A, Pfeifer G, Kerner M J, Naylor D J, Hartl F U and Hayer-Hartl M 2001 Dual function of protein confinement in chaperonin-assisted protein folding *Cell* **107** 223–33
[9] Farr G W, Furtak K, Rowland M B, Ranson N A, Saibil H R, Kirchhausen T and Horwich A L 2000 Multivalent binding of nonnative substrate proteins by the chaperonin groel *Cell* **100** 561–73
[10] Fenton W A, Kashi Y, Furtak K and Horwich A L 1994 Residues in chaperonin groel required for polypeptide binding and release *Nature* **371** 614–9
[11] Gulukota K and Wolynes P G 1994 Statistical mechanics of kinetic proofreading in protein folding *in vivo Proc. Natl. Acad. Sci. USA* **91** 9292–6
[12] Todd M J, Lorimer G H and Thirumalai D 1996 Chaperonin-facilitated protein folding: Optimization of rate and yield by an iterative annealing mechanism *Proc. Natl. Acad. Sci. USA* **93** 4030–5
[13] Dill K A, Bromberg S, Yue K, Fiebig K M, Yee D P, Thomas P D and Chan H S 1995 Principles of protein folding – a perspective from simple exact models *Protein Sci.* **4** 561–602
[14] Betancourt M R and Thirumalai D 1999 Pair potentials for protein folding: Choice of reference states and sensitivity of predicted native states to variations in the interaction schemes *Protein Sci.* **8** 361–9
[15] Chan H S and Dill K A 1996 A simple model of chaperonin-mediated protein folding *Proteins: Structure, Function, and Genetics* **24** 345–51
[16] Sfatos C D, Gutin A M, Abkevich V I and Shakhnovich E I 1996 Simulations of chaperone-assisted folding *Biochemistry* **35** 334–9
[17] Betancourt M R and Thirumalai D 1999 Exploring the kinetic requirements for enhancement of protein folding rates in the groel cavity *J. Mol. Biol.* **287** 627–44
[18] Baumketner A, Jewett A and Shea J E 2003 Effects of confinement in chaperonin assisted protein folding: Rate enhancement by decreasing the roughness of the folding energy landscape *J. Mol. Biol.* **332** 701–13
[19] Jewett A I, Baumketner A and Shea J E 2004 Accelerated folding in the weak hydrophobic environment of a chaperonin cavity: creation of an alternate fast folding pathway *Proc. Natl. Acad. Sci. USA* **101** 13192–7
[20] Jewett A I and Shea J E 2006 Folding on the chaperone: yield enhancement through loose binding *J. Mol. Biol.* **363** 945–57
[21] Cheung M S and Thirumalai D 2006 Nanopore-protein interactions dramatically alter stability and yield of the native state in restricted spaces *J. Mol. Biol.* **357** 632–43
[22] Tehver R and Thirumalai D 2008 Kinetic model for the coupling between allosteric transitions in groel and substrate protein folding and aggregation *J. Mol. Biol.* **377** 1279–95
[23] Jewett A I and Shea J E 2008 Do chaperonins boost protein yields by accelerating folding or preventing aggregation? *Biophys. J.* **94** 2987–93
[24] Genevaux P, Keppel F, Schwager F, Langendijk-Genevaux P S, Hartl F U and Georgopoulos C 2004 *In vivo* analysis of the overlapping functions of dnak and trigger factor *EMBO Rep.* **5** 195–200

[25] Lin Z and Rye H S 2004 Expansion and compression of a protein folding intermediate by groel *Mol. Cell* **16** 23–34

[26] Tang Y C, Chang H C, Roeben A, Wischnewski D, Wischnewski N, Kerner M J, Hartl F U and Hayer-Hartl M 2006 Structural features of the groel-groes nano-cage required for rapid folding of encapsulated protein *Cell* **125** 903–14

[27] Zhou H X and Dill K A 2001 Stabilization of proteins in confined spaces *Biochemistry* **40** 11289–93

[28] Klimov D K, Newfield D and Thirumalai D 2002 Simulations of beta-hairpin folding confined to spherical pores using distributed computing *Proc. Natl. Acad. Sci. USA* **99** 8019–24

[29] Takagi F, Koga N and Takada S 2003 How protein thermodynamics and folding mechanisms are altered by the chaperonin cage: molecular simulations *Proc. Natl. Acad. Sci. USA* **100** 11367–72

[30] Rathore N, Knotts T A t and de Pablo J J 2006 Confinement effects on the thermodynamics of protein folding: Monte Carlo simulations *Biophys. J.* **90** 1767–73

[31] Cheung M S, Klimov D and Thirumalai D 2005 Molecular crowding enhances native state stability and refolding rates of globular proteins *Proc. Natl. Acad. Sci. USA* **102** 4753–8

[32] Zhang S Q and Cheung M S 2007 Manipulating biopolymer dynamics by anisotropic nanoconfinement *Nano Lett.* **7** 3438–42

[33] Eggers D K and Valentine J S 2001 Molecular confinement influences protein structure and enhances thermal protein stability *Protein Sci.* **10** 250–61

[34] Farr G W, Fenton W A and Horwich A L 2007 Perturbed atpase activity and not 'Close confinement' Of substrate in the cis cavity affects rates of folding by tail-multiplied groel *Proc. Natl. Acad. Sci. USA* **104** 5342–7

[35] Chandler D 2005 Interfaces and the driving force of hydrophobic assembly *Nature* **437** 640–7

[36] Nymeyer H and Garcia A E 2003 Simulation of the folding equilibrium of {alpha}-helical peptides: A comparison of the generalized born approximation with explicit solvent *Proc. Natl. Acad. Sci.* **100** 13934–9

[37] Rhee Y M, Sorin E J, Jayachandran G, Lindahl E and Pande V S 2004 Simulations of the role of water in the protein-folding mechanism *Proc. Natl. Acad. Sci. USA* **101** 6456–61

[38] Ruhong Z 2003 Free energy landscape of protein folding in water: explicit vs. Implicit solvent *Proteins: Struct. Funct. Genet.* **53** 148–61

[39] Sorin E J, Rhee Y M, Shirts M R and Pande V S 2006 The solvation interface is a determining factor in peptide conformational preferences *J. Mol. Biol.* **356** 248–56

[40] Hummer G, Rasaiah J C and Noworyta J P 2001 Water conduction through the hydrophobic channel of a carbon nanotube *Nature* **414** 188–90

[41] Mashl R J, Joseph S, Aluru N R and Jakobsson E 2003 Anomalously immobilized water: A new water phase issue by confinement in nanotubes *Nano Lett.* **3** 589–592

[42] Singh S, Houston J, van Swol F and Brinker C J 2006 Superhydrophobicity: drying transition of confined water *Nature* **442** 526

[43] Sorin E J and Pande V S 2006 Nanotube confinement denatures protein helices *J. Am. Chem. Soc.* **128** 6316–7

[44] Ziv G, Haran G and Thirumalai D 2005 Ribosome exit tunnel can entropically stabilize alpha-helices *Proc. Natl. Acad. Sci. USA* **102** 18956–61

[45] Lucent D, Vishal V and Pande V S 2007 Protein folding under confinement: a role for solvent *Proc. Natl. Acad. Sci. USA* **104** 10430–4

[46] Wang J D, Herman C, Tipton K A, Gross C A and Weissman J S 2002 Directed evolution of substrate-optimized groel/s chaperonins *Cell* **111** 1027–39

[47] Machida K, Kono-Okada A, Hongo K, Mizobata T and Kawata Y 2008 Hydrophilic residues 526kndaad531 in the flexible c-terminal region of the chaperonin groel are critical for substrate protein folding within the central cavity *J. Biol. Chem.* **283** 6886–96

[48] Bulone D, Martorana V, San Biagio P L and Palma-Vittorelli M B 1997 Effects of electric charges on hydrophobic forces *Phys. Rev.* E **56** R4939

[49] Bulone D, Martorana V V, San Biagio P L and Palma-Vittorelli M B 2000 Effects of electric charges on hydrophobic forces: II *Phys. Rev.* E **62** 6799–809

[50] Dzubiella J and Hansen J P 2003 Reduction of the hydrophobic attraction between charged solutes in water *J. Chem. Phys.* **119** 12049–52

[51] Dzubiella J and Hansen J P 2004 Competition of hydrophobic and coulombic interactions between nanosized solutes *J. Chem. Phys.* **121** 5514–30

[52] Weixin X and Yuguang M 2008 Polar confinement modulates solvation behavior of methane molecules *J. Chem. Phys.* **128** 234506

[53] England J, Lucent D and Pande V 2008 Rattling the cage: computational models of chaperonin-mediated protein folding *Curr. Opin. Struct. Biol.* **18** 163–9

[54] England J and Pande V 2008 *Biophys. J.* in press

[55] England J, Lucent D and Pande V 2008 *J. Am. Chem. Soc.* in press

[56] Jacob E, Horovitz A and Unger R 2007 Different mechanistic requirements for prokaryotic and eukaryotic chaperonins: a lattice study *Bioinformatics* **23** i240–248

[57] Fan H and Mark A E 2006 Mimicking the action of groel in molecular dynamics simulations: application to the refinement of protein structures *Protein Sci.* **15** 441–8

[58] Furuta T, Fujitsuka Y, Chikenji G and Takada S 2008 In silico chaperonin-like cycle helps folding of proteins for structure prediction *Biophys. J.* **94** 2558–65

[59] Kyte J and Doolittle R F 1982 A simple method for displaying the hydropathic character of a protein *J. Mol. Biol.* **157** 105–32

[60] Chaudhry C, Farr G W, Todd M J, Rye H S, Brunger A T, Adams P D, Horwich A L and Sigler P B 2003 Role of the gamma-phosphate of atp in triggering protein folding by groel-groes: Function, structure and energetics *Embo J.* **22** 4877–87

[61] Shilton B H, Shuman H A and Mowbray S L 1996 Crystal structures and solution conformations of a dominant-negative mutant of escherichia coli maltose-binding protein *J. Mol. Biol.* **264** 364–76