Published in final edited form as:

Phys Chem Chem Phys. 2014 April 14; 16(14): 6358–6366. doi:10.1039/c3cp52872c.

Effect of interactions with the chaperonin cavity on protein folding and misfolding†

Anshul Sirur^a, Michael Knott^a, and Robert B. Best^{a,b,*}

^aCambridge University, Department of Chemistry, Lensfield Road Cambridge CB2 1EW, United Kingdom

^bLaboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520, United States

Abstract

Recent experimental and computational results have suggested that attractive interactions between a chaperonin and an enclosed substrate can have an important effect on the protein folding rate: it appears that folding may even be slower inside the cavity than under unconfined conditions, in contrast to what we would expect from excluded volume effects on the unfolded state. Here we examine systematically the dependence of the protein stability and folding rate on the strength of such attractive interactions between the chaperonin and substrate, by using molecular simulations of model protein systems in an idealised attractive cavity. Interestingly, we find a maximum in stability, and a rate which indeed slows down at high attraction strengths. We have developed a simple phenomenological model which can explain the variations in folding rate and stability due to differing effects on the free energies of the unfolded state, folded state, and transition state; changes in the diffusion coefficient along the folding coordinate are relatively small, at least for our simplified model. In order to investigate a possible role for these attractive interactions in folding, we have studied a recently developed model for misfolding in multidomain proteins. We find that, while encapsulation in repulsive cavities greatly increases the fraction of misfolded protein, sufficiently strong attractive protein-cavity interactions can strongly reduce the fraction of proteins reaching misfolded traps.

1 Introduction

Chaperonins are an important class of macromolecular machines which assist the cell in preventing protein misfolding and aggregation^{1,2}. The simplest mechanism for avoiding aggregation is simple sequestration in the chaperonin cavity³. Beyond this, it is clear that chaperonins such as GroEL may have additional effects on the folding itself^{4–6}. One facet is the role of the ATP-driven cycle of conformational changes involved in the capture and

†Electronic Supplementary Information (ESI) available: [Supporting Table giving model fit parameters and Supporting Figure giving $P(TP|Q)$]. See DOI: 10.1039/b000000x/

Fax: +1-301-496-0825; Tel: +1-301-496-5414; robertbe@helix.nih.gov.

Supporting information

Details on the analysis of the simulations as well as additional discussion. This material is available free of charge at <http://pubs.acs.org/>.

release of the protein⁷, which may allow for kinetic proofreading⁸. The exact nature of such conformational changes is still somewhat controversial. Nonetheless, simple physical considerations indicate that confinement within the chaperone must also affect folding, even without considering the role of ATP-driven changes. At the most elementary level, confinement in a cavity increases the free energy of the unfolded state more than that of the folded state, resulting in a lower stability and increased folding rate. A number of theoretical, experimental and numerical studies have been conducted to understand this effect, and these have generally confirmed expectations arising from the effect of confinement on the unfolded state, except in extreme confinement scenarios where the folding rate may again slow down^{9–12}.

In addition to the unavoidable consequences of the volume excluded by the chaperonin, more complex effects are possible by virtue of attractive interactions with the cavity wall^{13–16}. These have been shown to affect the folding rate and stability profoundly. Some recent studies have considered solvent effects on folding under confinement, either via explicit solvent simulation^{17–20} or theory²¹, although conclusions regarding the relative importance of solvent mediated and direct interactions with the walls differ.

Recently, we studied the folding of the protein Rhodanese in a single-ring GroEL/GroES system, using a coarse-grained model in which the substrate protein and entire chaperonin system were represented by one bead per residue¹⁶. Using a transferable interaction potential between substrate and chaperonin, we found that the effect of such interactions was more than sufficient to offset the effect of excluded volume, in agreement with experiments which found that confinement in the same single-ring GroEL variant either did not affect, or slowed down, folding²². Although this result was interesting, the size of the system did not permit us to perform a more systematic study or to develop models to describe the phenomenology.

In this work, we study the effect of confinement, within model cavities, on the folding of three small proteins, chosen from various structural classes: prb_{7–53} (all- α), NTL9 (α/β), and protein G (α/β), using G-type folding models motivated by the energy landscape theory for folding^{23,24}. The chaperonin cavity is modelled here as a sphere, either with a repulsive surface, or with variable short range attractive interactions with the protein. The model is sufficiently simple for us to scan a large region of parameter space. We find an interesting non-monotonic dependence of the protein stability and folding rates on the attraction strength, with weak attractions stabilising the folded state and stronger interactions destabilising it. We have developed a simple phenomenological model describing the effect of an attractive cavity on the folded and unfolded states, which explains the dependence of folding rates on the interaction strength. Lastly, we ask whether there may be any benefit to simple attractive interactions with the cavity wall. We show, in the context of multidomain protein folding, that sufficiently strong interactions can, in principle, help to prevent misfolding due to domain swapping.

2 Methods

The intramolecular interactions for each peptide were derived from the experimental structures (1div²⁵, 1prb²⁶, 1pgb²⁷) using the G β model of Karanicolas and Brooks²⁸. Only residues 1–39 of 1div were used for the simulations, corresponding to the miniature NTL9 variant used in experimental folding studies²⁹, while residues 7–53 of 1prb were used, corresponding to those used in the kinetic studies³⁰. Simulation temperatures of 325K (1div), 290K (1prb) and 312K (1pgb) were each chosen slightly above the folding temperature of each protein, to offset the initially stabilizing effect of excluded volume confinement.

Folding simulations were performed with an α -carbon representation of each protein, using a modified version of the GROMACS 4.0.5 program³¹. The system was evolved using Langevin dynamics, with a timestep of 15 fs and a friction coefficient of 0.2 ps⁻¹. Bond lengths were constrained to their initial values using the LINCS algorithm³¹. All simulations were initialised from the folded conformation and run for 15 μ s. Coordinates were saved every 30 ps.

When simulating the proteins in confinement, we add a potential of the form:

$$V(r_i)=4\varepsilon \left[\left(\frac{\sigma}{R-r_i} \right)^{12} - \left(\frac{\sigma}{R-r_i} \right)^6 \right], \quad (1)$$

where r_i is the distance of residue i from the centre of the cavity, ε is the interaction strength, R is the cavity radius and σ is the range of interaction, fixed at 6 Å. For repulsive confinement, we used a value of $\varepsilon = 0.0001$ kJ mol⁻¹. Simulations of the attractive cavity used various interaction strengths and cavity sizes.

Definitions of parameters Q and Θ

In order to calculate the fraction of native contacts Q , we first define the degree of contact formation between each pair of residues that forms a contact in the native state:

$$q_{ij}=(1+\exp[\gamma(r_{ij}-\lambda r_{ij}^{(0)})])^{-1}, \quad (2)$$

where γ defines the steepness of the contact formation curve, set to 2 Å⁻¹, $r_{ij}^{(0)}$ is the distance between the residues in the native structure and r_{ij} their distance in the configuration of interest. The scaling factor $\lambda = 1.2$ is introduced to account for the fact that residues forming a favourable pair interaction may still have $r_{ij} > r_{ij}^{(0)}$. Then the total fraction of native contacts is

$$Q=\frac{1}{N_Q} \sum_{(i,j)} q_{ij}, \quad (3)$$

where the sum runs over native contact pairs (i, j) and N_Q is the total number of native contacts in the polypeptide.

For the analysis of the two-domain titin simulations, the fraction of native contacts for an individual residue i , is given by

$$Q_i^{\text{native}} = \frac{1}{N_i} \sum_{j \in \text{same domain}} q_{ij}, \quad (4)$$

where the index j runs only over the N_i residues in contact with residue i in the native state. Note that there are no native inter-domain contacts as the native structure is of the monomer only³². Similarly, the fraction of misfolded contacts formed by residue i is given by

$$Q_i^{\text{misfolded}} = \frac{1}{N_i^{\text{misf}}} \sum_{j \in \text{other domain}} q_{ij}, \quad (5)$$

where the index j runs only over the N_i^{misf} native-like contacts which can be formed by residue i with residues in the incorrect domain.

The fraction of residues in contact with the cavity, $\Theta(t)$, was defined as $N_c(t)/N_r$ where $N_c(t)$ is the number of peptide residues within 8 Å of the cavity wall at time t and N_r is the total number of residues.

The location of the folded and unfolded minima along Q were determined by computing the potential of mean force $F(Q)$ using the weighted histogram analysis method (WHAM)³³.

Calculating the folding kinetics

The folding rate k_f and unfolding rate k_u were calculated from the native contact trajectories $Q(t)$ using the mean residence times $\langle \tau_f \rangle$ in the unfolded state before folding events, and $\langle \tau_u \rangle$ in the folded state before unfolding events:

$$\langle \tau_{f,u} \rangle = \frac{1}{n_{f,u}} \sum_{i=1}^{n_{f,u}} \tau_{f,u}^i, \quad (6)$$

where $n_{f,u}$ is the number of folding/unfolding transitions and $\tau_{f,u}^i$ is the time spent in the folded/unfolded state before each folding/unfolding transition. The error in $\langle \tau_{f,u} \rangle$ is calculated as $\langle \tau_{f,u} \rangle / \sqrt{n_{f,u}}$. The rate is then

$$k_{f,u} = \langle \tau_{f,u} \rangle^{-1} \pm [\langle \tau_{f,u} \rangle \sqrt{n_{f,u}}]^{-1}. \quad (7)$$

The stability of the folded state, $G^{(\varepsilon)}$, is calculated as

$$\beta \Delta G^{(\varepsilon)} = -\log \frac{k_f^{(\varepsilon)}}{k_u^{(\varepsilon)}}, \quad (8)$$

where $\beta = 1/k_B T$. For the attractive cavity simulations, the stability is calculated relative to the repulsively confined protein: $G^{(\varepsilon)} = G^{(\varepsilon)} - G^{(0)}$.

Calculating the diffusion coefficient

The position-dependent diffusion coefficient $D(Q)$ was calculated using the method described by Hummer³⁴. This method has been reviewed recently³⁵, but we summarise it briefly here. The coordinate Q is discretised into 30 – 40 bins, and, for each pair of bins (i, j) , we count the number N_{ji} of transitions from bin i to bin j , after a lag time τ , chosen such that the dynamics appear Markovian after this lag. We found that a lag time of 1.5 ns was sufficient to suppress non-Markovian effects.

A Bayesian method is used to determine $D(Q)$ and $F(Q)$ which are most consistent with the observed transitions N_{ji} in the trajectory. A smoothing prior with a stiffness of 10^{-5} s^{-1} was used, as discussed in ref³⁵.

3 Results and Discussion

Effect of attractive confinement on the folding energy landscape

In order to separate the effects of excluded volume confinement from those due to the attractive interactions, we performed simulations of ntl9, prb and protein G with no confinement, confined in a spherical cavity with repulsive walls and confined in a spherical cavity with attractive interactions between the residues and wall. By varying the interaction strength ϵ and the cavity size R , we explored the parameter space for protein folding inside attractive cavities. The cavity radii were chosen according to the size of the folded protein: we used 24, 28, 32 Å for ntl9; 28, 30, 32 Å for prb; and 26, 28, 30 Å for protein G. These were chosen such that the smallest cavity radius was roughly $3R_g$, except in the case of protein G where this resulted in only a modest stabilization by the repulsive cavity, and a slightly smaller cavity was used.

In our spherical cavities, a reduction in the conformational space available to the unfolded state should result in a destabilisation of the unfolded relative to the folded state; the effect on the folded state is expected to be small by comparison. This leads to the well-known increase in protein stability due to excluded volume confinement^{9,11,12,36–38}, e.g. for an ideal chain, the stability will vary approximately as $\sim R^{-2}$. Our aim is to explore the effect of additional attractive interactions with the cavity wall. It is clear that in the limit of strong interactions, the unfolded state will be stabilised due to its ability to form more contacts with the surface (Fig. 1 C). However, the behaviour for intermediate values of interaction is less clear.

We initially characterise the effect of confinement on the folding energy landscape by calculating free energy surfaces F for projection onto the fraction of native contacts Q , a natural variable for describing protein folding^{39–43}, particularly for a G-like model. In Fig. 2, we show representative plots of $F(Q)$ for each protein in the unconfined, repulsive and attractive scenarios, for the largest cavity size (in order to emphasise the effects of attractive interactions over volume exclusion). As has been shown previously^{9,13,18,19,44,45}, encapsulation in a repulsive sphere results in a destabilisation of the unfolded state relative to the folded state when compared with the unconfined protein. More unexpectedly, however, for weakly attractive interactions with the cavity, we observe a further stabilisation of the folded state, over that seen in the repulsive cavity¹³. At large ϵ , we obtain the

predicted result of stabilisation of the unfolded state relative to the folded state. Thus, our results indicate that there are some attraction strengths at which the protein is stabilised to a significant degree when compared with the unconfined or repulsively confined scenarios. It must also be noted that the effect of the attractive interactions is highly non-linear in Q , such that in some cases the relative stability of unfolded and folded states may be similar, but the barrier height can be quite different (compare for example the surface for protein G in a repulsive cavity with that for $\varepsilon = 0.15 k_B T$ in Fig. 2); i.e. the change in the potential of mean force $F(Q)$ would not be well described by a linear free energy relation. Although this might be attributed to a change in the quality of Q as a folding reaction coordinate, we have applied a Bayesian criterion^{42,46} to assess Q as a coordinate and find that in most cases it remains a good coordinate, or at least changes in a way opposite from that which could explain the increased barrier. The change must therefore arise from some more subtle effect, to which we return below.

Change in stability and folding rates

Having described the effect of attractive confinement on the equilibrium properties of the proteins, we next determined its influence on the folding and unfolding kinetics of the different proteins in each confinement scenario. Folding rates k_f (and unfolding rates k_u) were determined from long equilibrium folding trajectories, using the mean waiting times before folding (or unfolding) transitions, assessed using the native contact trajectories $Q(t)$. Folding transitions were determined as segments of $Q(t)$ during which Q crosses from Q_u to Q_f without recrossing, and vice versa for unfolding transitions, with Q_u and Q_f being set points in the centre of the unfolded and folded states: values are $Q_f = 0.7$ (ntl9), 0.9 (prb), 0.85 (protein G) and $Q_u = 0.3$ (ntl9), 0.5 (prb), 0.2 (protein G).

Figure 3 presents the dependence of $\log k_f$, $\log k_u$ and G on the strength ε of the attraction, all values being relative to those for repulsive confinement. The change in stability G shows a turnover as a function of ε at low attraction strengths (fig. 3), a result which has been previously noted using a simplified lattice model¹³ of protein confinement, as well as models for protein-protein interactions in the presence of attractive crowders⁴⁷. The value of ε at which the difference between the folding and unfolding rates relative to the repulsively confined scenario is maximised is also dependent on the cavity size, with larger cavity sizes reducing the attraction strength at which the folded state is most stable. At larger cavity sizes, more extended unfolded conformations become accessible and are stabilised by interactions with the attractive walls.

Because the fraction of native contacts Q is generally a good coordinate for describing protein folding dynamics⁴⁸, folding kinetics is often modelled as one-dimensional diffusion along Q ^{35,41,48,49}. In this picture, the diffusion coefficient D along Q (possibly a function of Q) and free energy $F(Q)$ completely determine the rates. We first examine the variation in diffusion coefficient with changing attraction strength, by calculating position-dependent diffusion coefficients $D(Q)$, shown in Figure 4. The change in diffusion coefficient in all cases is by a factor of less than ~ 2 , most importantly in the region near the folding transition state, which controls the rate.

The changes in $D(Q)$ experienced by the substrate proteins as ϵ is increased are sufficiently small that we will assume in the discussion below that the change in folding and unfolding rates is primarily due to the energetic and entropic effects of interacting with the cavity walls. While this may be true here, our potential is relatively smooth, and so we cannot rule out the possibility that effects on the diffusion coefficient may be relevant for a more rugged energy surface.

Rationalising “turnover” in stability and folding rate

We attempt to rationalise the observed trends in stability and folding and unfolding rates from the perspective of a one-dimensional projection of the dynamics onto a good reaction coordinate. In this way, by analysing separately the changes in the folded, unfolded and transition states in this projection, we should be able to explain any changes in stability and barrier heights—and hence the folding rate.

Some qualitative insights into the changes in properties of the energy landscape can be obtained by constructing two-dimensional potentials of mean force $F(Q, \Theta)$, where Θ is the fraction of residues adsorbed to the surface (defined as those within $8 \text{ \AA}$ of R), shown in Figure 5. On these 2D free energy surfaces, the folding coordinate Q resolves the folded and unfolded states, whilst Θ reveals how each state responds to changes in attraction ϵ . For ntl9, the folded and unfolded states are defined with minima at $Q = 0.7$ and $Q = 0.3$, respectively. In the repulsive cavity, the bulk of folded and unfolded state conformations form a minimal number of contacts with the cavity ($\Theta < 0.2$). At $\epsilon = 0.09k_B T$, the landscape shifts uniformly towards larger Θ as both folded and unfolded conformations form contacts with the cavity, with the folded state becoming more stable. As attraction strengths are increased further, the unfolded state starts to become stabilised and the minimum shifts towards $\Theta = 0.8$. The folded state is destabilised at $\epsilon = 0.31k_B T$ and the height of the folding barrier is increased significantly. While a generally similar picture is seen for the unfolded states of prb and protein G, there are some differences in the properties of the folded states: instead of continuously shifting toward larger Θ with increasing ϵ , the folded states of prb and protein G can be resolved into two distinct populations. One population is not bound to the wall, while the other forms the maximum number of contacts. The effect of changing ϵ is less to move the centres of these two populations toward higher Θ than to shift the system from one population to another. Thus, while the overall trends in stability and folding rate may be similar from protein to protein, there are subtleties in the way in which the cavity modulates the folding energy landscape.

The behaviour of the folded state for prb and protein G is approximately “two-state”, with adsorbed and non-adsorbed populations. This might be expected if the folded protein were very rigid: the effect of changing ϵ would be a simple energetic stabilisation of the adsorbed population, leading to a shift between populations, as seen particularly clearly for prb.

Simple model for attractive confinement

Using the insight into the phenomenology of attractive confinement from the 2D projections, we have devised a simple analytical model which captures the essential features. We consider only the change relative to excluded volume confinement, as that case has already

been extensively considered in the literature^{9,11,12,36–38,50}. For the folded state, we adopt a “two-state” model, with the protein either associated with the wall, or dissociated. We denote the volumes of configuration space associated with the wall-associated and dissociated states as V_{ass} and V_{diss} respectively. If n contacts are formed with the cavity wall, each with energy $-\varepsilon$, the adsorption energy is $U_{\text{cont}} = -n\varepsilon$. Although such a model, in which the protein is approximately rigid (so the number of contacts with the wall changes little with ε in the wall-associated state), might describe the data for prb and protein G, for NTL9 the number of contacts formed in the associated state clearly changes with ε . To account for this, we include an elastic (free) energy associated with the internal protein degrees of freedom. This we approximate as being harmonic in the number of contacts with the wall, n . However, a certain number of contacts, $n^0 = N_{\text{res}}\Theta^0$, (where N_{res} is the number of residues and Θ^0 the fraction of bound residues corresponding to n^0) can be formed at no energetic cost, so that the elastic energy of the folded state is:

$$\beta U_{\text{el}}(n) = \frac{1}{2}\kappa(n - n^0)^2. \quad (9)$$

Minimising the energy of the adsorbed species, for a given value of ε , yields for the total change of energy on adsorption:

$$\beta U_{\text{cont}}(\varepsilon) + \beta U_{\text{el}}(\varepsilon) = -n^0\varepsilon - \frac{\varepsilon^2}{2\kappa}. \quad (10)$$

Combining this with the assumption of a two-state model, we obtain for the overall change in free energy (including both adsorbed and non-adsorbed states) at a given ε :

$$\beta \Delta G = -\ln \left[\frac{1 + A \exp(n^0\varepsilon + \varepsilon^2/2\kappa)}{1 + A} \right], \quad (11)$$

where $A = V_{\text{ass}}/V_{\text{diss}}$. We are assuming here that A depends negligibly on the strength of attraction ε . Similarly, the average fraction of contacts at any given ε is:

$$N_{\text{res}}\Theta = \frac{A(n^0 + \varepsilon/\kappa) \exp(n^0\varepsilon + \varepsilon^2/2\kappa)}{1 + A \exp(n^0\varepsilon + \varepsilon^2/2\kappa)}, \quad (12)$$

where N_{res} is the total number of residues in the protein. Thus a more “rigid” structure with large κ would behave like an ideal two-state model and have a sigmoidal binding curve (i.e. Θ vs ε), while a more “flexible” structure would have a flatter, less cooperative curve.

The same model can reasonably be applied to the transition state and to the folded state. It is less clear that such a model would be applicable to the unfolded state. However, we note that a simple calculation by de Gennes⁵¹ for the change of free energy of a chain upon weak adsorption to a surface shows that it scales approximately as $\sim \varepsilon^2$, while the number of contacts with the wall scales as $\sim \varepsilon$. We therefore describe the change in free energy of the unfolded state simply as $\beta G_{\text{u}} = -B_{\text{U}}\varepsilon^2$.

We have used the simple model to fit the data in Fig. 3, where the best fit curves have also been plotted. The folding rate curves are fitted to the equations

$$-\log[k_f/k_f^0] = \beta \Delta G_{\ddagger} - \beta \Delta G_u \quad (13)$$

for the folding rates k_f , with the change of transition state free energy $\beta \Delta G_{\ddagger}$ being given by Eq. 11, and change of unfolded state free energy $\beta \Delta G_u$ by the de Gennes model described above. An analogous expression is used for unfolding rates. For the stabilities, we use the expression

$$\beta \Delta \Delta G = \beta \Delta G_f - \beta \Delta G_u \quad (14)$$

where the quantities are as defined above. The curve fits are generally good, considering the simplicity of the model; the best fit parameters are given in Table 1 for the smallest cavity for each protein and in Table S1 for all cavity sizes. The trends in the parameters are physically reasonable: $A_{\ddagger} > A_F$, $\kappa_{\ddagger} > \kappa_F$ as expected for a transition state which is more expanded and deformable than the folded state, and $n_{\ddagger}^0 < n_F^0$. This last trend can be understood as a consequence of the lower density of residues on the surface of a more expanded transition state. Similarly, the number of contacts formed by the folded state (in the limit $\varepsilon \rightarrow 0$), n_F^0 and the prefactor for the unfolded state adsorption free energy, B_U , increase with the size of the protein.

The model can explain a number of the unexpected features of the data. For example, the “turnover” in stability is a consequence of the different response of the folded and unfolded proteins to the attractive interactions. For a rigid folded state where κ_F is very big, the energy of the adsorbed folded state changes as $n_F^0 \varepsilon$, while the free energy of adsorption of the unfolded state is $B_U \varepsilon^2$. Thus, at low ε , the folded state will be stabilised more, but at high ε , the unfolded state stabilisation will dominate. The physical reason for this difference is that at low ε , the unfolded state must pay a large entropic penalty for residues bound to the surface, yet gains little favourable interaction energy as a result. On the other hand, the folded state, having many residues in close proximity, can readily make a large number of favourable contacts without any entropic cost to the bound state, and so is stabilised at low ε . However, at high ε , the unfolded state can always be stabilised more by adsorbing more residues to the surface (until they are all adsorbed), while the number of residues in the folded state which can contact the surface is more limited. The model can also explain why the barrier height may be increased as the result of attractive interactions, even when (as described above) the change in stability is relatively small. In our model, the reason is that the transition state does not benefit fully from either stabilisation mechanism: it cannot form as many contacts at low ε as the folded state, being more disordered. However, although it is more easily deformed than the folded state, as evidenced by $\kappa_{\ddagger}$, it is less easily deformed than the unfolded state, presumably because a certain degree of native structure formation is essential.

A factor we have not considered to this point, and which is not explicitly included in our model, is the possible influence of the shape of the native protein on the observed confinement effects. This could affect how complementary the protein surface is with that of the idealized cavities considered here. For example, if the folded protein has a surface whose

radius of curvature matches the confining boundary, one would expect that it would be easily stabilized by attractive interactions. A similar argument would apply to the folding transition state, although in that case the shape may differ for different members of the transition-state ensemble (and indeed attractive confinement may then select some transition states over others). In our model, these subtleties are lumped into the fit parameters n_F^0 and n_T^0 , the numbers of contacts which the native and transition states can make without any structural deformation.

Role of attractive confinement in preventing misfolding?

We have shown that, in contrast to confinement in a hard cavity which always stabilises proteins and accelerates folding rates, interactions with the cavity walls in an attractive cavity may either stabilise or destabilise the protein; as noted above, there may even be an increase in the folding barrier when the overall change in stability is small. The recent observation that confinement in GroEL may actually slow the folding of Rhodanese²², indicates that a scenario in which the attractive interactions are strong enough to overcome confinement effects is not merely hypothetical. The fact that folding can be slower in the presence of passive confinement (i.e. without ATP-dependent cycling) may mean that the role of the chaperonin is to prevent aggregation, or that it is essential to consider the ATP-dependent cycle of states. However, it may be that interactions with the cavity in the absence of turnover may help to mediate folding—as has been proposed in both experimental⁵ and coarse-grained simulation¹⁴ studies. We illustrate this for a model system which has recently been shown to be prone to misfolding: a pair of tandem titin I27 domains (Figure 6 (A))⁵². Single molecule fluorescence experiments, together with coarse-grained G β model simulations, have shown that this dimer is able to form domain-swapped misfolded species which are stable for several days (an example is shown in Figure 6 (B)). Since the misfolding observed in the simulations has already been experimentally validated, this seems to be an ideal system for investigating the putative role of interactions with the chaperonin in preventing misfolding. We note here that an amyloid-like misfolding mechanism has recently been proposed as an alternative to the domain-swapped misfolding of titin discussed here⁵³, although that study found that domain-swapping was the predominant misfolding mechanism for other proteins. Clearly, further experiments will be helpful in pinpointing the type of misfolding in titin, but the exact mechanism does not affect the “proof of principle” established in the present work.

The protein model we use is identical to that described in Borgia *et al.*⁵², and the cavity model is the same as described earlier in the present work. We use a cavity radius of 35 Å, in order that the volume be comparable to that of the interior of GroEL. In addition to the unconfined case considered in earlier work, we consider repulsive confinement, and either “weak” ($\epsilon = 0.5k_B T$) or “strong” ($\epsilon = 1.0k_B T$) interactions with the cavity walls. We ran long folding simulations, initialised from a high temperature unfolded state for 5 μ s, with 50 independent runs for each condition studied.

In each case, we classified the final structure, where possible, as either “folded” or “misfolded”; the classification was based on a clustering of the contact maps, as previously described⁵². Each domain-swapped structure may in fact be viewed as two folded titin

domains, where one domain has been inserted into the loop of the other. We thus name the misfolded states according to which loop the imaginary insertion has been performed on. Note that the simulations used are long enough for the protein to misfold, but the protein is never observed to unfold again once it reaches this state (indeed the two time scales are separated by ~ 4 –5 orders of magnitude in experiment).

The results of the folding runs are summarised in Fig. 6 (c)–(f). The first bar in each histogram gives the native population, and the remainder correspond to the most populated misfolded species. Each misfold can be viewed as one folded titin domain inserted into a loop of another folded titin domain, and the misfolds have accordingly been classified as e.g. “C/D” for the case where that loop is the one between the “C” and “D” strands of the native structure⁵⁴. Repulsive confinement, in fact, results in a large increase of misfolding events, presumably because of the increased density of residues in the non-native states. However, what is quite remarkable is the effect of interactions with the cavity wall. Weak interactions already reduce the fraction of misfolding relative to the repulsively confined case, while stronger interactions reduce it still further, to be comparable with the unconfined protein. These results suggest that the attractive interactions, while they may slow folding for simple two-state folders, may accelerate it for proteins which can fall into very stable misfolded traps.

What is the origin of the reduced misfolding in the attractive cavities? We have attempted to answer this by looking at how the formation of contacts with the wall affects contact formation within the protein. Intuitively, one would expect that formation of contacts with the walls would favour formation of short-range protein-protein contacts over long-range ones. Since the contacts associated with domain swapping are generally more long-ranged, this would explain the result. Indeed we find that this appears to be the mechanism observed. In Fig. 7, we plot statistics of contact formation in the unfolded state: the average fraction of time which each residue forms contacts with the wall $\langle\Theta_i\rangle$, and the fraction of contacts $\langle Q_i\rangle$ formed by each residue i (i) out of its native contact partners and (ii) out of its misfolded contact partners (Eq. 4–5). It is evident that increasing the magnitude of the interactions with the wall increases formation of protein-wall contacts (as expected), but also progressively disfavours misfolded contacts. The consequent lack of misfolded contact formation in the unfolded state explains the reduced yield of misfolded protein.

4 Conclusions

We have shown that attractive interactions with the surface of a simple cavity modelling the interior of a chaperonin can significantly modulate the folding rate and stability of small proteins: relative to simple excluded volume confinement, weak attractive interactions may lead to a further stabilisation of the protein and reduction in folding rate. However, sufficiently strong interactions will once more destabilise the protein, such that they may even completely offset the excluded volume effect of confinement. Based on the observed trends in contact formation with the cavity, we have developed a simple phenomenological model that explains the observed trends in terms of the different interactions formed by folded and unfolded states with the cavity.

Since the slowdown of folding rates due to attractive interactions means that folding inside a chaperonin may not directly accelerate folding, we ask whether it may confer other benefits. We have tested this using a simulation model of misfolding in multidomain proteins which has been validated against experiment. The model shows that the hypothetical situation of purely repulsive confinement would lead to an unacceptable increase in the population of misfolded states. However, even relatively weak interactions of $\sim 1 k_B T$ can greatly reduce the yield of misfolded protein, which can be explained by the tendency of interactions with the wall to favour short-range over long-range contacts. While this is clearly a simplified model, it suggests a mechanism by which chaperonins (and chaperones more generally⁵⁵) may help to avoid domain-swapped misfolding of newly-synthesised proteins.

One simplification is that we are using only a single average energy scale to characterise interactions with the cavity wall, neglecting sequence specific effects. Inclusion of such effects would obviously be an area for future development.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

AS was supported by a BBSRC doctoral studentship, MK and RBB by BBSRC award BB/H006885/1. RBB was supported by a Royal Society University Research Fellowship while at Cambridge and by the Intramural Research Programme of the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health.

References

1. Thirumalai D, Lorimer GH. *Annu. Rev. Biophys. Biomol. Struct.* 2001; 30:245–269. [PubMed: 11340060]
2. Saibil HR, Ranson NA. *Trends Biochem. Sci.* 2002; 27:627–632. [PubMed: 12468232]
3. Apetri AC, Horwich AL. *Proc. Natl. Acad. Sci. U. S. A.* 2008; 105:17351–17355. [PubMed: 18987317]
4. Todd M, Lorimer G, Thirumalai D. *Proc. Natl. Acad. Sci. U.S.A.* 1996; 93:4030. [PubMed: 8633011]
5. Brinker A, Pfeifer G, Kerner M, Naylor D, Hartl F, Hayer-Hartl M. *Cell.* 2001; 107:223–233. [PubMed: 11672529]
6. Bracher A, Starling-Windhof A, Hartl FU, Hayer-Hartl M. *Nat. Struct. Mol. Biol.* 2011; 18:875–881. [PubMed: 21765418]
7. Saibil HR, Fenton WA, Clare DK, Horwich AL. *J. Mol. Biol.* 2012; 425:1476–1487. [PubMed: 23183375]
8. Gulukota K, Wolynes PG. *Proc. Natl. Acad. Sci. U. S. A.* 1994; 91:9292–9296. [PubMed: 7937758]
9. Takagi F, Koga N, Takada S. *Proc. Natl. Acad. Sci. U.S.A.* 2003; 100:11367–11372. [PubMed: 12947041]
10. Zhou H-X, Rivas G, Minton AP. *Ann. Rev. Biophys.* 2008; 37:375–397. [PubMed: 18573087]
11. Mittal J, Best RB. *Proc. Natl. Acad. Sci. U. S. A.* 2008; 105:20233–20238. [PubMed: 19073911]
12. Hayer-Hartl M, Minton AP. *Biochemistry.* 2006; 45:13356–13360. [PubMed: 17073456]
13. Betancourt MR, Thirumalai D. *J. Mol. Biol.* 1999; 287:627–644. [PubMed: 10092464]
14. Jewett AI, Baumketnet A, Shea JE. *Proc. Natl. Acad. Sci. U. S. A.* 2004; 101:13192–13197. [PubMed: 15331776]

15. Zhuang Z, Jewett AI, Kuttimalai S, Bellesia G, Gnanakaran S, Shea J-E. *J. Phys. Chem. B.* 2011; 100:1306–1315.
16. Sirur A, Best RB. *Biophys. J.* 2013; 104:1098–1106. [PubMed: 23473493]
17. Lucent D, Vishal V, Pande VS. *Proc. Natl. Acad. Sci. U.S.A.* 2007; 104:10430–10434. [PubMed: 17563390]
18. Tian J, Garcia AE. *J. Am. Chem. Soc.* 2011; 133:10141–10149. [PubMed: 21598973]
19. Bhattacharya A, Best RB, Mittal J. *Biophys. J.* 2012; 103:596–600. [PubMed: 22947876]
20. Marino KA, Bolhuis PG. *J. Phys. Chem. B.* 2012; 116:11872–11880. [PubMed: 22954175]
21. Zhou H-X. *J. Chem. Phys.* 2008; 469:245101.
22. Hofmann H, Hillger F, Pfeil SH, Hoffmann A, Streich D, Haenni D, Nettels D, Lipman EA, Schuler B. *Proc. Natl. Acad. Sci. U. S. A.* 2010; 107:11793–11798. [PubMed: 20547872]
23. Wolynes PG. *Q. Rev. Biophys.* 2005; 38:405–410. [PubMed: 16934172]
24. Kumar S, Ma B, Tsai C-J, Sinha N, Nussinov R. *Protein Sci.* 2000; 9:10–19. [PubMed: 10739242]
25. Hoffman DW, Cameron CS, Davies C, White SW, Ramakrishnan V. *J. Mol. Biol.* 1996; 264:1058–1071. [PubMed: 9000630]
26. Johansson MU, de Château M, Wikström M, Forsén S, Drakenberg T, Björck L. *J. Mol. Biol.* 2002; 316:859–865.
27. Gallagher T, Alexander P, Bryan P, Gilliland GL. *Biochemistry.* 1994; 33:4721–4729. [PubMed: 8161530]
28. Kim YC, Hummer G. *J. Mol. Biol.* 2008; 375:1416–1433. [PubMed: 18083189]
29. Horng J-C, Moroz V, Raleigh DP. *J. Mol. Biol.* 2003; 326:1261–1270. [PubMed: 12589767]
30. Wang T, Zhu Y, Gai F. *J. Phys. Chem. B.* 2004; 108:3694–3697.
31. Hess B, Kutzner C, van der Spoel D, Lindahl E. *J. Chem. Theory Comput.* 2008; 4:435–447.
32. Improta S, Politou AS, Pastore A. *Structure.* 1996; 4:323–337. [PubMed: 8805538]
33. Kumar S, Bouzida D, Swendsen RH, Kollman PA, Rosenberg JM. *J. Comp. Chem.* 1992; 13:1011–1021.
34. Hummer G. *New J. Phys.* 2005; 7:34.
35. Best RB, Hummer G. *Phys. Chem. Chem. Phys.* 2011; 13:16902–16911. [PubMed: 21842082]
36. Zhou H-X, Dill KA. *Biochemistry.* 2001; 40:11289–11293. [PubMed: 11560476]
37. Cheung MS, Klimov D, Thirumalai D. *Proc. Natl. Acad. Sci. USA.* 2005; 102:4753–4758. [PubMed: 15781864]
38. Zhou H-X. *Arch. Biochem. Biophys.* 2008; 469:76–82. [PubMed: 17719556]
39. Leopold PE, Montal M, Onuchic JN. *Proc. Natl. Acad. Sci. U. S. A.* 1992; 89:8721–8725. [PubMed: 1528885]
40. Šali A, Shakhovich E, Karplus M. *Nature.* 1994; 369:248–251. [PubMed: 7710478]
41. Succi ND, Onuchic JN, Wolynes PG. *J. Chem. Phys.* 1996; 104:5860–5868.
42. Best RB, Hummer G. *Proc. Natl. Acad. Sci. U. S. A.* 2005; 102:6732–6737. [PubMed: 15814618]
43. Zhang Z, Chan HS. *Proc. Natl. Acad. Sci. U. S. A.* 2012; 109:20919–20914. [PubMed: 23213246]
44. Klimov DK, Newfield D, Thirumalai D. *Proc. Natl. Acad. Sci. U. S. A.* 2002; 99:8019–8024. [PubMed: 12060748]
45. Friedel M, Sheeler D, Shea J. *J. Chem. Phys.* 2003; 118:8106.
46. Hummer G. *J. Chem. Phys.* 2004; 120:516–523. [PubMed: 15267886]
47. Rosen J, Kim YC, Mittal J. *J. Phys. Chem. B.* 2011; 115:2683–2689. [PubMed: 21361356]
48. Best RB, Hummer G. *Proc. Natl. Acad. Sci. U. S. A.* 2010; 107:1088–1093. [PubMed: 20080558]
49. Best RB, Hummer G. *Phys. Rev. Lett.* 2006; 96:228104. [PubMed: 16803349]
50. Casassa DF. *Polym. Lett.* 1967; 5:773–778.
51. de Gennes, P-G. *Scaling Concepts in Polymer Physics.* Cornell University Press; 1978.
52. Borgia MB, Borgia A, Best RB, Steward A, Nettels D, Wunderlich B, Schuler B, Clarke J. *Nature.* 2011; 474:662–665. [PubMed: 21623368]

53. Zheng W, Schafer NP, Wolynes PG. Proc. Natl. Acad. Sci. U. S. A. 2013; 110:1680–1685. [PubMed: 23319605]
54. Fowler SB, Clarke J. Structure. 2001; 9:355–366. [PubMed: 11377196]
55. O'Brien EP, Christodoulou J, Vendruscolo M, Dobson CM. J. Am. Chem. Soc. 2012; 134:10920–10932. [PubMed: 22680285]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 1. Different confinement scenarios for protein folding. (A) unconfined chain, with two-state free energy surface; (B) chain confined in a repulsive cavity; (C) chain confined in a cavity with short-range attractive interactions.

Fig. 2.

Representative folding free energy surfaces for three model proteins, at selected values of the strength of contact interactions with the cavity. The model proteins are NTL9 (upper), prb₇₋₅₃ (centre) and protein G (lower). For clarity, error bars are not shown on the plots, but we report the maximum error on $F(Q)$ out of the unfolded minimum, the folded minimum and the barrier top, over all the scenarios. These maximum errors are $0.02 k_B T$ for NTL9, $0.55 k_B T$ for prb₇₋₅₃ and $0.34 k_B T$ for protein G.

Fig. 3.

Systematic dependence of folding rates and stabilities on the strength of interaction with the cavity for the three model proteins: (a–c) ntl9, (d–f) prb, (g–i) protein G. Blue lines and points are for the smallest cavity size, green the middle cavity size and red the largest cavity size. Points are values computed from simulation and curves correspond to the fit of the phenomenological model, i.e. Eq. 13 for the rates and Eq. 14 for stabilities.

Fig. 4. Diffusion coefficient along the folding coordinate Q . Position dependent diffusion coefficients have been determined as a function of the interaction with the cavity for NTL9 (upper), prb (centre) and protein G (lower).

Fig. 5.

Two-dimensional potentials of mean force as a function of the fraction of residues Θ in contact with the walls and the fraction of native contacts Q for the protein. (a–c) are PMFs in the repulsive cavity, (d–f) are PMFs for “weak” attractions with the wall ($\epsilon = 0.09, 0.08, 0.06$ for NTL9, prb_{7–53} and protein G respectively), (g–i) are for “intermediate” attractions with the wall ($\epsilon = 0.25, 0.17, 0.15$ for NTL9, prb_{7–53} and protein G respectively), and (j–l) are for “strong” attractions to the wall ($\epsilon = 0.31, 0.25, 0.21$ for NTL9, prb_{7–53} and protein G respectively).

Fig. 6. Misfolding of titin. (a) Structure of folded titin dimer; (b) structure of a representative domain-swapped misfolded state. Yields of folded and misfolded structures are shown for (c) unconfined case, (d) repulsive confinement, (e) weak attraction to cavity and (f) strong attraction to cavity.

Fig. 7.

Effect of cavity interactions on formation of misfolded contacts. (a) Unconfined, (b) repulsive confinement, (c) confinement in a weakly interacting cavity and (d) confinement in a strongly interacting cavity. Blue curves give the average fraction of native contacts formed by each residue. Orange curves give the corresponding fraction of possible misfolded domain-swapped contacts. Black curves give the fraction of time for which each

residue is in contact with the cavity wall. All curves are averaged over a window of 5 residues.

Table 1

Parameters for fit of phenomenological model to rates and stabilities. Cavity size is given below each protein name.

Parameter:	NTL9	prb ₇₋₅₃	Protein G
Cavity:	2.4 nm	2.8 nm	2.6 nm
A_F	0.46	0.29	0.15
n_F^0	9.8	26	27
$\kappa_F / k_B T$	0.2	10	30
$A_{\ddagger}$	13	0.38	8.3
$n_{\ddagger}^0$	2.8	12	8.2
$\kappa_{\ddagger}^+ / k_B T$	0.053	0.030	0.020
$B_U / k_B T$	14	34	58