Cell. Mol. Life Sci. (2010) 67:255–276
DOI 10.1007/s00018-009-0164-6

Cellular and Molecular Life Sciences

REVIEW

Reconciling theories of chaperonin accelerated folding with experimental evidence

Andrew I. Jewett · Joan-Emma Shea

Received: 20 July 2009 / Revised: 14 September 2009 / Accepted: 25 September 2009 / Published online: 23 October 2009
© Birkhäuser Verlag, Basel/Switzerland 2009

Abstract For the last 20 years, a large volume of experimental and theoretical work has been undertaken to understand how chaperones like GroEL can assist protein folding in the cell. The most accepted explanation appears to be the simplest: GroEL, like most other chaperones, helps proteins fold by preventing aggregation. However, evidence suggests that, under some conditions, GroEL can play a more active role by accelerating protein folding. A large number of models have been proposed to explain how this could occur. Focused experiments have been designed and carried out using different protein substrates with conclusions that support many different mechanisms. In the current article, we attempt to see the forest through the trees. We review all suggested mechanisms for chaperonin-mediated folding and weigh the plausibility of each in light of what we now know about the most stringent, essential, GroEL-dependent protein substrates.

Keywords GroEL · Passive · Active · Obligate substrates · ACM · Stationary IAM

Introduction

Proteins in vivo fold in a very different environment than they do in in vitro experiments. In vivo proteins are manufactured by ribosomes in a highly crowded environment that does not resemble buffers used in in vitro experiments [1, 2]. A significant fraction of the protein chains in the cell do not fold spontaneously, but interact with "chaperones". These proteins (or protein complexes) can interact with the non-native nascent chains as they extend from the ribosome, guiding them to organelles, across membranes, or transporting them to other chaperones, or to the proteolytic machinery [2–8]. Chaperones are believed to facilitate or promote the folding of proteins which are unable to fold on their own under cellular conditions. Promiscuous chaperones such as GroEL/ES, DnaK/J, and trigger factor, are involved in the folding of approximately 10–30% of the proteins in Eubacteria, Eukarya and Archaea [3, 4, 7, 8].

Chaperones typically go through multiple large conformational changes (typically driven by ATP binding and hydrolysis) which alternately increase and decrease their affinity for protein substrates [3–6]. The forces exerted by a chaperone are strong enough to cause traumatic conformational changes in the proteins bound to them [9–12].

The most exhaustively studied chaperone is GroEL. It has been estimated that up to 5% of the proteins in *Escherichia coli* depend on GroEL to fold [13]. GroEL is a member of the HSP60 family of promiscuous type-I chaperonins found in prokaryotes and in eukaryotic mitochondria. *Chaperonins* are a class of hollow cylindrical chaperones. Their singular characteristic is that these hollow cylinders can completely enclose (the majority of) their protein substrates. While held within this container, proteins appear to be able to continue folding [12, 14–18], in some circumstances, at what appears to be an accelerated rate [16, 19–22]. The effect that these cylindrical walls have on the proteins contained inside is the subject of considerable dispute.

GroEL (HSP60) helps a wide variety of proteins fold with the aid of other chaperones. Together with GroES

A. I. Jewett · J.-E. Shea (✉)
Department of Chemistry and Biochemistry,
University of California, Santa Barbara, CA 93106, USA
e-mail: shea@chem.ucsb.edu

A. I. Jewett · J.-E. Shea
Department of Physics, University of California,
Santa Barbara, CA 93106, USA

(HSP10), trigger factor, DnaK (HSP70), DnaJ (HSP40), GrpE (nucleotide exchange factor, NEF), prefoldin (in Eukarya and Archaea), and HSP90 (in Eukarya), GroEL performs regular cell maintenance and is present at high concentrations even in the absence of external stress [2, 3, 5, 7, 12, 23–26]. Under typical conditions, GroEL is backlogged. At least 99% of GroEL chaperones are occupied by substrate proteins [13]. Proteins which interact with GroEL often interact first with other chaperones such as trigger factor, DnaJ, DnaK, and, in the presence of stress, small heat shock proteins. Some of these chaperones, like DnaK/J, are present in much higher concentrations than GroEL. It has been proposed that, in addition to other functions, these chaperones act as a queue, temporarily detaining excess denatured proteins until they either fold, or bind to a GroEL ring once it becomes available [3, 5, 13, 23, 26]. This may help filter out the majority of proteins which do not strictly need GroEL from clogging up the comparatively scarce GroEL machinery [13].

GroEL consists of a pair of open cylinders stacked end to end [27]. Each cylinder, constructed from seven identical 57 kDa peptides [28], is hollow and traditionally is believed to accommodate proteins up to 60 kDa in size [8], although larger potential substrates over 80 kDa have been identified and characterized [13, 29–32]. Each cylinder can also bind to nucleotides ATP/ADP, as well as GroES, a co-chaperone which acts as a lid, closing the container and sealing any substrate protein small enough to fit inside.

GroEL chaperonins, like many other chaperones, feature patches of highly non-polar residues (the “apical domain”) allowing them to selectively recognize proteins with hydrophobic chains which are exposed to the solvent [5, 28, 33–38]. The presence of exposed hydrophobic residues is a characteristic typical of misfolded proteins and that makes them prone to aggregation. The presence of an analogous cluster of hydrophobic residues on the chaperonin enables GroEL to bind to a diverse set of misfolded protein substrates, preventing aggregation and possibly promoting folding.

There are many combinations of ways these molecules could bind to the two rings of GroEL. Nevertheless, GroEL appears to bind to substrate proteins, ATP, and GroES in an ordered, regulated cycle. Many in vitro experiments have been carried out to untangle the order of binding events in a typical GroEL binding and release cycle in vivo, The two rings of GroEL can communicate in an allosteric manner, so that events in one cylinder trigger transitions in the other cylinder [15, 21, 39–53]. Unfortunately, kinetics data from in vitro experiments on GroEL are strongly context-dependent and difficult to interpret. For example, the buffer composition can change protein folding mechanisms and kinetics [18, 54]. In addition, the concentration of denatured substrate proteins [40, 52, 55, 56], the cation (K^+) concentration [56–61], and the gradual accumulation of ADP [55, 62] can also significantly alter the GroEL cycle.

In search of a consensus view for the GroEL binding/release cycle

Although details are still under investigation, the basic cycle of GroEL binding and release is established, and is shown in Fig. 1:

1. Under ordinary conditions, non-native proteins bind rapidly to GroEL, or to other chaperones (typically DnaJ, DnaK, and trigger factor), which deliver them to GroEL [3, 4, 13, 23, 25, 63]. $t_{unbound}$ denotes the total time the protein remains unprotected in the cytosol during each cycle, which depends on the concentration and binding kinetics of these available auxiliary chaperones. (See “The role of HSP70/HSP40 and other ancillary chaperones”.)
2. Proteins experience denaturing stress during, or soon after binding to GroEL [9–11, 64].
3. After binding, there is a lag period [10, 21, 40, 52, 64–68] during which time the protein remains stuck to the opening, and is unable to fold [58, 69–73].
4. The arrival of the co-chaperone, GroES, seals the cylindrical cavity shut, releasing the protein into the center of the cavity (called the “*cis* ring”) during

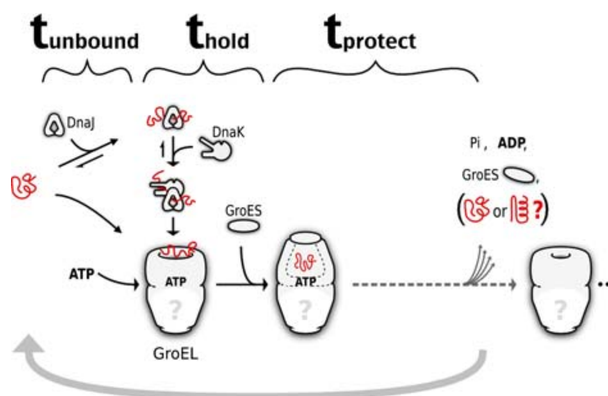


Fig. 1 A drastically simplified view of the GroEL binding and release cycle. Unfolded proteins either bind directly to GroEL, or (more likely) bind to other chaperones (for example, DnaK/J) which transport them to GroEL. Proteins are denatured and immobilized by the chaperonin upon binding to an open ring (the “*trans*-ring”) of GroEL. After a delay, the arrival of a co-chaperone (“GroES”) eventually frees the protein from the GroEL opening, and seals it inside a closed container (the “*cis*-ring”), where it can resume folding. After another delay ($t_{protect}$, during which time ATP hydrolysis occurs), the GroES “lid” disassociates, and the container opens. The protein may or may not have succeeded in folding by this time. If not, it may remain bound to the chaperone (see Fig. 4), or it may be ejected into the cytosol where it will quickly bind to another chaperone and repeat the cycle. The behavior of the opposite ring of GroEL (*bottom*) has been omitted. (See Fig. 2)

which time it is able to fold. After an additional delay, during which ATP undergoes hydrolysis, the GroES "lid" eventually disassociates, and the protein can either escape into the cytosol or remain bound to the chaperone. If the protein has not yet folded, it will bind to a chaperone and the cycle will begin again.

5. ATP binding and hydrolysis provides the energy to repeat this cycle.

A consensus regarding the durations of $t_{unbound}$, t_{hold}, and $t_{protect}$ has not yet been established in vivo. Folding is generally believed to occur at least as rapidly in the presence of GroEL + GroES + ATP as it does in the bulk (in vitro, under permissive conditions) [16] which implies that, in these cases, $t_{protect}$ should be significantly greater than t_{hold} (under those conditions).

As discussed in "Emerging details", the durations of $t_{protect}$, $t_{unbound}$ (and possibly t_{hold}) appear to be concentration dependent. Estimates of $t_{protect}$ range from 3 to 25 s [21, 40, 52, 56, 65, 66, 68], depending on temperature, as well as the concentrations of non-native substrate proteins, ions, and nucleotides [40, 55, 56]. Estimating $t_{unbound}$ is difficult because the probability of binding to a chaperone per unit time is proportional to concentration of available GroEL/ES and DnaK/J chaperones in vivo. The concentration of GroEL, for example, is known to be 2–5 μM in *E. coli* [12, 74] (twice that for GroES [74], 50 μM for DnaK [23, 25], and perhaps an order of magnitude lower for DnaJ [75, 76]). However, 99% or more of GroEL chaperones appear to be fully saturated with substrate protein [13]. This makes it difficult to estimate the concentration of vacant DnaK/J and GroEL rings available for binding.

The entire cycle of binding and release (*for each ring*, as depicted in Fig. 1, in the presence of substrate protein, GroES, and ATP) and requires between 3 s [55, 56, 65], up to 30 s or more [20, 40, 52, 55, 77], depending upon substrate concentration, nucleotide, and (K^+) concentration (see below). The cycle frequency increases from room temperature to body temperature. Hence, under typical conditions in vivo in *E. coli*, the faster estimates listed here are probably more accurate (H. Rye, personal communication). According to the traditional model, the two GroEL rings alternately take turns binding to, and hydrolyzing ATP [17, 39–43, 45–47, 49, 50, 56, 78]. In that case, the period of the entire (two-ring) cycle would be twice as long, requiring from 7 s, to a minute or more. However, some aspects of this model have been recently challenged [62, 79].

Emerging details

Over time, a diverse set of data has accumulated which complicates our understanding of the cycle. GroES, ADP, and substrate protein do *not* necessarily disassociate from GroEL simultaneously (as was depicted in Fig. 1). *Some* substrate proteins remain bound to GroEL after the GroES "lid" unbinds. (See "The stationary iterative annealing model" and "Estimating the fraction of time proteins are exposed to the cytosol".)

ADP can also linger long after GroES has unbound [52, 55, 56] (see Fig. 2a). Additionally, ATP hydrolysis in the cis ring cannot proceed until ADP departs from the opposite (*trans*) ring [45, 50, 52, 55, 80].

The delay in ADP departure appears to be substrate concentration dependent. It is known that with a sufficient excess of denatured substrate proteins (as is likely to exist in vivo [13, 56]; H. Rye, personal communication), GroEL/ES is "substrate driven", cycling at its maximum frequency which is rate-limited by the hydrolysis of ATP [40, 55, 56]. A recent experiment by the Lorimer group [55, 56] shows that the interaction of denatured substrate proteins with the open GroEL ring (the "*trans*" ring) weakens its affinity for ADP and removes this delay [55, 56]. At lower substrate protein concentrations (or excess K^+ ions), the slow release of ADP appears to be the rate-limiting step in the cycle [52]. It was suggested that GroEL's sensitivity to substrate concentration enables it to respond to conditions in the cell, reducing the occupancy time when the demand for chaperones is high, and slowing down ATP consumption when the protein concentration (and risk of aggregation) is low [40, 55].

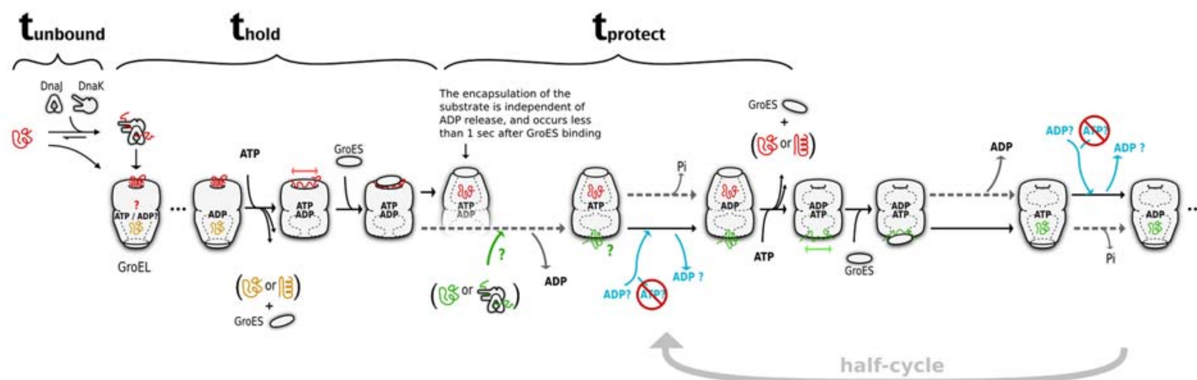
A different model has been recently proposed by the Funatsu and Taguchi groups [62, 79]. At sufficiently low ADP concentrations and high GroES concentrations, both rings of GroEL appear to be simultaneously able to bind to nucleotide and GroES in order to form stable symmetric GroES:GroEL:GroES "footballs" [22, 62]. These footballs coexist with nearly equal quantities of GroEL:GroES "bullets" (where only one side of GroEL binds to GroES) [22, 62]. Such high ATP and relatively low ADP concentrations may be typical in vivo within *E. coli* ([ATP]/[ADP] ≈ 7.90 mM/1.04 mM; see [62, 81].

One of several possible explanations for the coexistence of bullets and footballs is the potentially slow rate of ADP release (see Fig. 2b). This could be tested by varying the concentration of free substrate protein (up to the saturating concentration, approximately 5 μM [56]).

Substrate characterization

In order to properly understand the role of GroEL in protein folding, it is critical to characterize the nature of the proteins that depend on chaperonins to fold. The central question is: what is the obstacle that prevents these proteins from folding?

(A) The traditional GroEL/ES binding/release cycle



(B) A hypothetical cycle that populates both "footballs" and "bullets"

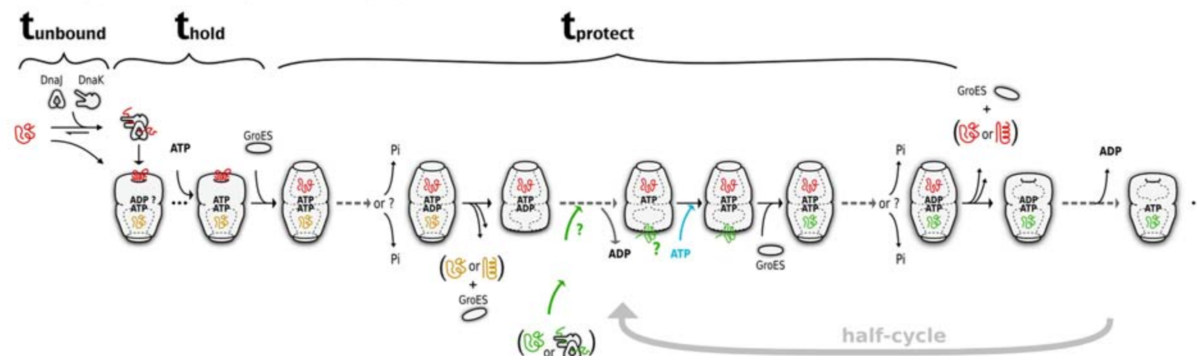


Fig. 2 A composite view of the **a** "bullet", and **b** "football" variants of the GroEL binding and release cycle. Here, we track the journey of one protein (*red*) through the chaperone machinery. Slow, or potentially slow, steps in the cycle are shown using *arrows* with *gray dashed lines*. All other steps involving GroEL are comparatively rapid in vivo (requiring approximately 1 s or less). These figures are a hypothetical synthesis which attempt to smooth out the differences between the conclusions of four recent experiments [52, 55, 62, 79], and do not represent the views of these authors. Uncertainties and conjectures were depicted in the figure with *question marks*. Note: portions of Fig. 2b were sent to us in a personal communication from H. Taguchi. **a** depicts the more traditional picture of the GroEL cycle: each ring of GroEL takes turns alternately binding to GroES and substrate protein. After protein binding, the subsequent binding of ATP [52, 82] can further unfold the substrate protein [10, 11, 44, 52, 64]. ATP binding also provides enough energy (roughly 46 kcal/mol per ring [83, 84]) to allosterically free GroES [85], and possibly the substrate protein (*yellow*) from the opposite ring, less than 1 s later [52]. The subsequent binding of GroES to the top ring eventually frees the protein from the opening of GroEL and into the newly formed "*cis*" chamber, where it can continue folding [10, 11, 21, 40, 52, 64–68, 86]. Meanwhile, ADP disassociates from the other (*bottom*, *trans*) ring. ADP competes with both ATP [62] (*light blue arrow*), and (in effect) substrate protein [55, 56] (*green arrow*) for the open (*trans*) ring. **b** depicts the GroEL cycle at low ADP concentrations. At low [ADP], ATP and GroES are able to bind to both rings of GroEL simultaneously [62, 79]; see also [22, 80, 87]. As an example, in this figure, we depict hydrolysis occurring in the opposite ring following ATP binding, in an alternating fashion. However, it is not clear which ring of GroEL will hydrolyze ATP first.

The majority of GroEL substrates need GroEL to avoid aggregation

A recent exhaustive identification of GroEL substrates identified a set of 85 proteins which are unable to fold in the absence of GroEL/ES [13]. For the majority of these proteins, aggregation emerges as the main obstacle to folding [88]. This supports earlier evidence that GroEL substrates tend to be aggregate prone [2, 4, 89–93]. Aggregation is a pervasive problem that plagues more than just the relatively small percentage of proteins which depend on GroEL to fold. Nearly all proteins are capable of aggregating [90, 94, 95]. When manufactured (at approximately 30 μg/mL concentrations) in a minimal expression system lacking chaperones (the "PURE" system [88, 96, 97]), the majority of cytoplasmic proteins in *Escherichia coli* are more likely to aggregate than fold. This is reflected by the fact that, for the majority of these cytoplasmic proteins, less than 50% were recovered after synthesis and centrifugation [88]. This study focused on measuring protein solubility during translation. It is also possible that some proteins which escape aggregation may fail to reach the native state. For cytoplasmic proteins, the distribution of solubility was bimodal, consisting of two distinct groups, which were designated "Agg" and "Sol", corresponding to highly aggregate-prone and soluble proteins, respectively. The great majority of obligate ("class

III") GroEL substrates appear to belong to the highly aggregate prone "Agg" group [88].

Some GroEL substrates may fail to fold for other reasons

Surprisingly, a few of the proteins identified among the most stringent ("class III") substrates of GroEL [13] belong to the "Sol" group. These consist of *E. coli* proteins which remain mostly soluble when expressed in the absence of chaperones. These proteins are listed in Table 1. The identity of these proteins was provided by H. Taguchi (personal communication). All these proteins were more than 80% recoverable under PURE conditions (see Fig. 4c of [88]). Additionally, two other stringent GroEL substrates have been identified (gatY and dapA) which in vivo neither fold nor precipitate the absence of GroEL [13]. It was suggested that these two proteins simply fold too slowly to escape proteolytic degradation. The fact that these proteins are not highly aggregation prone suggests that GroEL may play an *active* role in folding them. We are not aware of any studies seeking to establish whether or not the soluble proteins listed in Table 1 fold faster in the presence of GroEL. We would like to suggest that such studies may be extremely useful in reconciling the role of GroEL as an aggregation preventer and/or a folding promoter.

In summary, while the nature of the majority of stringent GroEL proteins suggests that the main role of GroEL would be to prevent aggregation, it does not rule out that GroEL may play a role in actively promoting the folding of these molecules.

Overview

The aim of the present paper is to review the proposed mechanisms for GroEL-mediated folding and assess their plausibility. Is GroEL simply an aggregation–prevention device, or can it (occasionally) perform a more active role in folding proteins? A number of theoretical models have been proposed that argue in either direction: Passive models (aggregation prevention) and active models (folding promotion). We outline these models below and assess them in light of what is known experimentally.

Passive models

Passive cage/ACM According to the Anfinsen cage model (ACM), GroEL does not alter protein folding kinetics or the folding pathways. The container formed by GroEL + GroES + ATP simply provides a safe environment in which proteins can fold without associating with other proteins [12, 14, 15, 89, 99–101].

Active models

IAM via ATP The traditional iterative annealing model (IAM) says that the ATP-driven cycles of binding and unbinding to GroEL accelerate folding by periodically disrupting or destabilizing off-pathway misfolded states [9, 10, 22, 72, 73, 77, 80, 102–114]. We suggest that GroEL may be able to accomplish this *without* releasing substrates into the cytosol (the *stationary* IAM).

Table 1 GroEL substrates which are resistant to aggregation while folding

Name	Locustag	*n*	SCOP	Ess	Description
ybaK	ECK0475	159	d.116	No	Protein binding, transcription regulator, DNA-dependent
pflA	ECK0893	246	?	No	Oxidoreductase, glucose metabolism
yfiF	ECK2579	345	c.116	No	Hypothetical methyl transferase, trmH family
suhB	ECK2530	267	e.7	Yes	Hydrolase, inositol-1(or 4)-monophosphatase activity, magnesium binding
smpB	ECK2616	160	b.111	No	Translation, RNA binding, Binds specifically to the ssrA RNA (tmRNA)
tldD	ECK3233	481	?	No	Suppresses the inhibitory activity of the carbon storage regulator (csrA)
rsd	ECK3987	158	?	Yes	Transcription regulator, DNA-dependent
gatY	ECK2089	286	c.1	No	Tagatose-1,6-bisphosphate aldolase gatY (TBPA)
dapA	ECK2474	292	c.1	Yes	Dihydrodipicolinate synthase (DHDPS)

The first seven rows describe the most highly soluble stringent ("class III") GroEL substrates classified by [13, 88]. (Data provided by H. Taguchi, personal communication, and also [13].) For each one of these proteins, fewer than 20% of them precipitated when expressed in a minimal PURE system [88, 96, 97]. The first and second columns contain the K-12 gene names and K-12 locustags, respectively. *n* denotes the number of amino acids in a monomer. The fourth column indicates the likely SCOP structural classification [98]. The fifth column indicates whether the protein is "essential". (Columns 4 and 5 were taken from [13], supplemental table S3. SuhB is monomeric, and dapA is a homotetramer.) The last two rows describe two additional stringent GroEL substrates examined by the Hartl group [13]. Neither of these proteins aggregated in vivo, even in the absence of GroEL [13]. (gatY displayed low solubility when expressed in the PURE system [88]. PURE data for dapA was not obtained.) Not all GroEL substrates were analyzed this way in vivo. These are two examples. We expect these lists of proteins will be subject to future revisions. *After submission, we were informed of new experimental evidence that a number of these proteins may not be obligate GroEL substrates in GroEL/GroES-depleted cells*. (See forthcoming publication by Fujiwara and Taguchi; H. Taguchi, personal communication.)

Active cage Chaperonins may accelerate folding by modifying the static environment in which proteins fold by, for example, smoothing the energy landscape [16, 115, 116], reducing the entropy of the unfolded state, [19, 20, 103, 116–124], providing new pathways for folding [11, 19, 20, 103, 110, 124, 125], or modulating the solvent behavior [122, 126–131].

Disaggregator/depolymerase Along with several other chaperones [132–137], GroEL/ES has been observed to reverse the early stages of aggregation [91]. Other chaperones, such as DnaK/J, are better suited for this task [136] so we omit this topic from this review.

The Anfinsen cage model (passive cage)

On behalf of the set of *aggregate-doomed* GroEL substrates, we ask: what is the most effective method for preventing aggregation? Since aggregation and folding are competing processes [90, 94, 135, 138–140], chaperones like GroEL could thwart aggregation in two ways: (1) by preventing access to other proteins (passive), or (2) by accelerating folding (active). For GroEL, the majority of evidence points towards the passive strategy. A large body of evidence suggests that stringent GroEL substrates fold almost exclusively during the time they spend sequestered inside the cage formed by GroEL + GroES + ATP [12, 14–21, 40, 44, 52]. This also appears to apply to type II chaperonins found in Eukarya [141, 142]. In some cases, chaperonins like GroEL appear to have a minimal effect on the behavior of their protein substrates, altering neither the protein's folding pathway [143–147], nor accelerating the protein's folding kinetics [16–18]. These observations suggest that chaperonins' role is simply to encapsulate their substrates, and prevent association with other proteins (and simulate an "infinitely dilute" environment). Chaperonins reduce the time that proteins spend unprotected in the cytosol ("bulk") before folding by a factor which can be crudely estimated from the rates at which denatured proteins bind to and are released from the various chaperones with which they interact [21, 148]. (See "Estimating the fraction of time proteins are exposed to the cytosol".)

The ATP-driven iterative annealing model

The iterative annealing mechanism (IAM) is probably the earliest and most popular explanation for how chaperones may be able to *actively* accelerate folding [9, 10, 22, 72, 73, 77, 80, 102–111, 113, 114]. Chaperones like GroEL have been observed to distort their substrate's conformation during binding [9]. Evidence that a moderate degree of denaturation occurs upon binding is extensive [9–12, 38, 64, 149–152] (although some evidence suggests otherwise [143, 145–147]). It has been unclear whether the ability of GroEL to denature its substrates upon binding is a critical aspect of its function, or merely a side-effect of the way GroEL binds non-specifically to the exposed hydrophobic portions of non-native proteins. It may be relevant to mention that, after binding to substrate protein, the arrival of ATP cause GroEL to undergo an additional conformational shift [9, 27, 44, 85] that can further denature the proteins that are bound to it [10, 11, 38, 64, 107] (see Figs. 2a and 4, sideways arrows).

ATP binding and consumption provides the energy for GroEL to repeatedly bind to proteins, denature them, and (sometimes) release them [40, 44, 82–85]. Stringent GroEL-dependent proteins typically endure many such ATPase cycles before folding [13] (see also [7, 16, 80, 91, 102]). Obligate ("class III") GroEL substrates fold in 30–60 s *on average* [13], fast enough to encompass at most 15 ATPase cycles, assuming the chaperone operates at near the maximum speed of approximately 4 s per cycle; see "Emerging details". However, some slow-folding obligate substrates like rhodanese and RuBisCo are believed to require as long as 7 min to fold on average, and may endure 100 cycles before folding [7, 16, 55, 56].

Abundant evidence has accumulated that many moderate and large proteins have kinetic intermediates. Given that all known small two-state folders require between 10^{-6} and 1 s to fold, the fact that typical GroEL substrates require a minute to fold, or longer, suggests that the folding of these proteins is rate-limited by kinetic intermediates [93, 153–158]. Numerous compact partially folded conformational states have been discovered at various pH and ion concentrations. These states are believed to correspond to kinetic intermediates typically encountered during folding.

As mentioned earlier, some proteins are susceptible to making inappropriate interactions with other proteins, which can interrupt the folding process, and lead to aggregation. However some proteins, especially large, multi-domain proteins, may be susceptible to misfolding due to energetic frustration, which occurs when the chain makes inappropriate interactions with itself [3, 159–165]. For additional theoretical perspectives, see [166–173]. These energetically favorable interactions could trap the protein, and retard the folding process.

Proteins may also be susceptible to topologically frustration, prevented from reaching the native state by steric hindrance, trapped by the formation of (correct) intra-chain contacts made too early [172, 174–178].

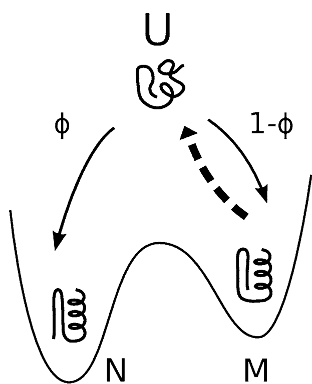


Fig. 3 An illustrative example of a protein that would benefit from iterative denaturation. *N* and *U* refer to the "native" and "unfolded" states, respectively. This protein has a single dominant long-lived off-pathway kinetic trap (denoted *M* for "misfolded"). ϕ denotes the probability that a denatured protein is able to avoid kinetic traps on the way to the folded state. If trapped, the protein must escape before it can fold, resulting in a much slower effective rate of folding. Binding to GroEL destabilizes the protein's conformation, ousting it from all of these long-lived, stable, misfolded conformations, and giving it another chance to fold (*dashed arrow*)

A cartoon depicting a protein which has a single misfolded conformation is shown in Fig. 3 [103, 104, 179].

The IAM stipulates that the repeated denaturation caused by multiple cycles of binding and release may free proteins from these kinetic traps. The more frequently this occurs, the more opportunities a protein has to escape, leading to faster folding, for this class of frustrated, trapped proteins. Accelerated folding due to iterative denaturation has been predicted mathematically [104, 109, 114, 148], and has been observed in minimalist polymer simulations (on a lattice [103, 104, 110] and off-lattice [111, 125]).

A protein's rate of folding can be accelerated by a factor which is *at most* proportional to the cycle frequency, and even then, only if the protein's folding is rate-limited by escape from long-lived off-pathway kinetic traps (longer-lived than the cycle interval). On the other hand, proteins with *on-pathway* intermediates are likely to be *decelerated* by iterative denaturation, which erases whatever progress has been made along the folding pathway. A general formula for a protein's folding time under the influence of cycles of forced unfolding is given in [148] and in "A review of the effects of iterative denaturation".

Unfortunately, there is a paucity of evidence directly demonstrating that forced unfolding by GroEL actually leads to faster folding. Indeed, folding rate accelerations in the presence of GroEL (or GroEL mutants) tend to be modest: less than a factor of 4 for wild-type proteins, and this acceleration may be due to other reasons. There are, however, examples of proteins that do not fold faster in the presence of GroEL [see "The Anfinsen cage model (passive cage)"].

Traditional IAM does not explain aggregation reduction

However, there is another serious problem with the IAM theory. Because the rate at which proteins are bound (denatured) and released is almost entirely determined by the rate of ATP hydrolysis which drives the cycle forward, one of the main predictions of the IAM is that increasing this rate of hydrolysis will accelerate folding, and consequently, reduce aggregation and increase yield. We have examined this argument and concluded that this mechanism could never increase the *yield* of *aggregate-doomed* proteins [148]. Although, rapid ATP hydrolysis could accelerate folding (see "A review of the effects of iterative denaturation"), it would also accelerate the rate at which proteins are ejected into the cytosol where they are at risk of aggregating. (Rapid ATP hydrolysis reduces $t_{protect}$.) This later effect always dominates, more than canceling the benefits of iterative denaturation, increasing aggregation and decreasing yield. This remains generally true even in the presence of other chaperones (see "Estimating the fraction of time proteins are exposed to the cytosol").

Such a mechanism *could* explain how GroEL enables the folding of proteins which are *not* susceptible to aggregation (for example, see Table 1). But, for the set of aggregate-doomed proteins, the optimal way to improve protein yield would be to bind to the protein once, and to keep it encapsulated until folding [148].

The stationary iterative annealing model

This leads us to suggest another possible model for GroEL-mediated folding (the stationary iterative annealing model). In this model, substrates annealed by cycles of ATP hydrolysis while remaining bound to GroEL until folding (Fig. 4). This scenario (Fig. 4) is qualitatively different from the mechanism we discussed above, and immune to some of our criticisms.

Early observations of GroEL-mediated folding of the RuBisCo and MDH proteins lead scientists to believe that GroEL ejects its substrates into the cytosol every cycle [21, 80, 102]. Later observations revealed that rhodanese can remain bound to the ring even after it opens (following ATP hydrolysis) [12, 180, 181]. This issue is complicated the fact that, in vitro, a small, but significant fraction (25%) of non-native rhodanese proteins do "leak" out during cycle [180, 181] (in the absence of crowding agents). Because rhodanese is a slow folder, requiring 7 min (between 7 and 60 full two-ring ATPase cycles [7, 16, 55, 56, 149]), this would mean that rhodanese is typically released from GroEL in a non-native state multiple times before folding. In other words, rhodanese appeared to behave like other substrates that interact with GroEL, albeit

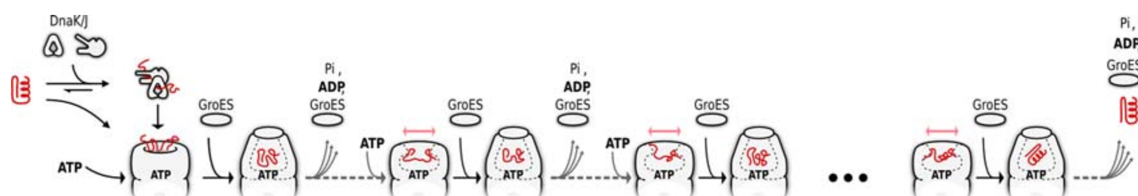


Fig. 4 The stationary iterative annealing model: the traditional iterative annealing model (IAM) assumes that proteins are released into the bulk during each cycle of ATP hydrolysis (as depicted in Fig. 1). This might accelerate folding, but in the end would be counter-productive for aggregate-prone proteins, ejecting them into the bulk where they may aggregate. As an alternative, we point out that some substrates remain bound to GroEL during multiple rounds of of ATP hydrolysis [7, 12, 180]. Here, we show a hypothetical protein which remains bound to GroEL until folding, experiencing multiple ATPase cycles without risking aggregation. The *red sideways arrows* depict the denaturation of the protein upon ATP binding. (See Fig. 2)

with (effectively) a slower unbinding rate. (See Eqs. 1 and 3 from "Substrates do not always unbind from GroEL".)

However, followup experiments revealed that, in vivo, leakage is reduced to the point where rhodanese essentially remains bound to GroEL until folding [7]. This is likely due to macromolecular crowding in the cytosol [24, 180, 182].

The distinction matters. Proteins which do not unbind from GroEL are still subjected to cycles of forced unfolding (see Fig. 4). Upon successfully folding, proteins conceal their hydrophobic residues which could help them permanently escape GroEL [12], i.e., by passing through the opening of the GroEL *trans*-ring without sticking to it. In this way, GroEL might be able accelerate the folding of certain proteins without increasing their risk of aggregation.

Hence, in principle, iterative denaturation may explain the GroEL-mediated folding, even for some aggregate-doomed substrates. While rhodanese might not be the best example (see "Experimental studies of confinement" and "Conclusion"), there are many uncharacterized obligate GroE substrates that might benefit from stationary iterative annealing.

Active cage theories

A number of experiments have demonstrated that the GroEL/ES complex can assist protein folding in the absence of cycling (multiple rounds of binding and release of GroES/ATP) [14–20, 40, 44]. The folding of some proteins can also be enabled by GroEL alone, in the absence of GroES and ATP, for example, hen lysozyme [144], barnase [183, 184], and rhodanese [185]. Unless otherwise specified, all other experiments were carried out in the presence of GroES and nucleotide.

Some proteins appear to fold more rapidly in the presence of chaperonins. MDH folds 3.7× faster in the presence of GroEL than it does unassisted in the bulk [21]. Larger rate accelerations have been observed for mutants of maltose binding protein (MBP) in the presence of GroEL [19, 22] and SR-EL [19] (see below).

Perhaps the most compelling experiment to demonstrate accelerated folding in the *absence of cycling* was performed by Hartl and co-workers [16] using a single-ring mutant of GroEL (SR-EL [17]) which is unable to release GroES (or the enclosed substrate protein) after binding. They showed that the obligate GroEL substrate (RuBisCo) appears to fold 4× faster [16] when sealed inside the SR-EL + GroES cage than it does in the bulk under folding permissive conditions.

Altered folding rates in the cage can be due to a number of factors, including (1) steric confinement effects, (2) interaction of the protein with the mildly hydrophobic walls of the chaperonin, and (3) modulation of the hydrophobic effect due to confined water. The GroEL chaperone is complicated, and separating the effects of pure confinement on folding from the other effects at play is a nearly impossible task from an experimental perspective. Experiments of folding within the chaperonin cage are difficult to decipher, and, as we will show below, the same experiment is subject to different interpretations by different research groups. We begin by reviewing the theoretical studies of protein folding in a cage. While these studies omit many of the details of realistic protein and chaperonin systems, they have the great advantage of being able to deconvolute confinement from other cage effects.

Confinement

Effects of confinement on folding

Thermodynamic reasoning would allow us to argue the following: One of the primary roles of confinement is to eliminate extended conformations, thus reducing the conformational entropy of the unfolded state. A direct result of a decrease in the entropy gap between the collapsed and coiled states is an increase in the melting temperature of the protein (considering that the energy gap is unaffected by confinement). Hence, confinement should increase the

stability of the folded state. The role of confinement on folding rates is more subtle, but one can argue that eliminating conformations should increase folding rates (if the transition state is unaffected). These effects would only hold true for cage sizes slightly larger than the protein itself (i.e., in between the radius of gyration for the folded and unfolded states). For cages that are large compared to the size of the unfolded protein, one would expect no effect on the protein stability. On the other hand, for cages similar in size to the protein, one would expect that steric effects would lead to destabilization of the protein and decreased folding rates.

Theoretical studies of confinement

A number of theoretical models have probed the effect of confinement on folding. Zhou and Dill [186] used an analytical theory in which the unfolded protein was described by a Gaussian chain and the folded protein by a sphere. By solving the diffusion equation with different boundary conditions corresponding to different confining cages, the authors calculated the effect of confinement on the folding free energy and showed that confinement would lead to a gain in stability of at least 9 kcal/mol. This stabilization is seen only for cavity sizes slightly larger than the protein. Experimental confirmation of increased stability in cages can be found in the work of Eggers and Valentine [187] in which the melting temperature of alpha-lactalbumin was seen to increase upon encapsulation in silica pores. Using the same Gaussian chain model for the protein [188], Zhou showed that confinement leads to an overall increase in rate of contact formation between residues of the protein. Reduction in the diffusion of the chain was found to be small compared to the increase in rate of contact formation, resulting in an overall increase in folding rates upon confinement (for cage sizes slightly larger than the protein).

Following suit, a number of simulations have probed the effects of confinement on protein stability and folding rates. Simulations ranged from Monte Carlo simulations using lattice models [103], to more sophisticated C_α based off-lattice Go-models [118–123, 189, 190]. The Go-model is a simple coarse-grained view of proteins in which attractive interactions are present only between those residues that would form native contacts in the folded state. The energetic frustration of such proteins is minimal, and these proteins tend not to be kinetically trapped. In agreement with the analytical results of Zhou, these simulations show that encapsulation in a purely repulsive container (sphere or cylinder) would lead to increased folding rates and stabilities for containers slightly larger than the protein itself. This increase in folding rate is dependent on the topology of the protein [189] and more pronounced for proteins with larger numbers of long-ranged tertiary contacts [118]. Additional simulations on confined (unfrustrated) Go-models [120, 191] found that the transition state was slightly destabilized by confinement, although it remained very similar in terms of the number of native contacts to the bulk case. The above studies all indicate that such unfrustrated folders are stabilized by confinement and do not exhibit a dramatic change in the folding mechanism (other than a compaction of the unfolded state and a very minor compaction of the transition state). Simulations on more realistic non-Go models [116] also showed that proteins that are not significantly kinetically trapped will experience both an increase in stability and folding rates upon confinement in an inert pore slightly larger than the protein. However, proteins whose folding is severely hindered by deep, misfolded kinetic traps experienced a decrease in folding rates upon confinement at temperatures at which misfolded structures were significantly populated. These proteins require large conformational rearrangements to escape these traps. Confinement slows folding by preventing proteins from adopting the kind of extended conformations required to bring them out of their trapped state. These simulations indicate that pure confinement of a highly frustrated protein can significantly inhibit folding and does not lead to folding acceleration. However, it could still serve the role of reducing the concentration of aggregate-prone proteins in the crowded cellular milieu. It is critical to establish whether GroEL substrates have stable kinetic traps.

Experimental studies of confinement

A number of experiments, driven by work in the Hartl and Horwich groups, have focused on elucidating the role of the GroEL cavity in folding. Experiments using a non-cycling single-ring GroEL mutant (SR-EL) were particularly useful in this respect. Some of these experiments are discussed below.

Hartl and co-workers performed experiments on the wild-type GroEL and on the single-ring mutant SR-EL under permissive conditions so that folding rates in and out of the cage could be compared without needing to consider the effect of aggregation [16]. The results of these experiments appear to be in good agreement with theoretical predictions of confinement effects. For the proteins studied, those significantly smaller than the 60 kDa GroEL (such as the 33 kDa rhodanese) did not show any change in folding rate in the presence of GroEL or SR-EL. Proteins slightly smaller than GroEL experienced an increase in folding rates. In the case of the 50 kDa RuBisCo, a rate increase of a factor of four was seen both with the regular GroEL and the single-ring mutant. (Additional experiments by Hartl ruled out the possibility that a single unfolding event by GroEL would allow RuBisCo to fold efficiently in the bulk). The observed rate increase in the chaperonin cavity

was attributed to confinement effects leading to a "smoothing of the energy landscape" [16]. This could be interpreted to mean the loss of stable extended misfolded conformations [116].

Further experiments by Hartl and co-workers [19] modulated the size of the SR-EL cavity by altering the C-terminal GGM repeats to further probe the role of confinement on folding rates. The size of the cavity was varied by either deleting or adding to the $[GGM]_4M$ C-terminal segments of each GroEL subunit. Although not seen in the crystal structure, these segments are believed to protrude into the cavity. In some constructs, G and M were mutated to A, to rule out the role of the specific sequence or hydrophobicity in promoting or preventing folding. Deletion of a single segment would increase the volume by 4.4%, while adding a segment would decrease the volume by 4.4%.

Experiments were performed under both permissive conditions (on a slow folding mutant of the MBP, mw 41 kDa) and under non-permissive conditions for rhodanese (33 kDa), and RuBisCo (50 kDa). In the case of the smallest protein, rhodanese, reduction of the cage size up to 4.4% was seen to increase folding rates, again in agreement with confinement theory. The same effect was observed when the M residues in the repeat sequences were mutated to A, indicating that the rate increases could be attributed solely to cage size effects and not to interactions of the protein with the wall of the cavity. Even further reduction in cage size led to a decrease in folding rate with no loss of yield. Eventually, further size reduction led to a greater decrease in folding rate, accompanied by a 40–70% loss of yield. Importantly, the ability of the chaperonin to encapsulate the protein was unaffected, intimating that the reduction in folding rates was due to steric hindrance effects (for example, the inability of trapped conformations to rearrange their structure in the tight space of the confined chaperonin). Later work probed this notion of steric restriction using steady-state fluorescence anisotropy measurements. These studies showed restricted mobility in the cage when inserts were present [for GFP (27 kDa) and DHFR-GFP]. Such restriction of mobility in small cages has been observed in confinement simulations [116, 121].

In the case of DM-MBP (a double-mutant of maltose-binding protein [192, 193]), similar rate enhancements (13-fold) were found in the presence of the wild-type cycling GroEL and in the presence of the single-ring mutant SR-EL, intimating that the cage itself (and not cycling) is responsible for decreases in folding times [19]. Reducing the size of the GroEL cavity (wild-type and SR-EL) was seen to slow down the folding of DM-MBP by 40%. An even more dramatic reduction in folding rates was seen for the larger RuBisCo protein. While a small reduction in cage size slowed down folding without altering the yield, more dramatic cage restrictions affected both rates and yields and were associated with a dramatic drop in the ability of smaller chaperonins to encapsulate the protein.

The overarching conclusion reached by the Hartl group based on the above experiments is that confinement effects can explain the observed rate changes of folding in the presence of chaperonins.

These conclusions were challenged by Horwich and co-workers [18, 99] who argued that the rate accelerations in SR-EL seen in the work of Hartl and co-workers [16] under permissive conditions were not due to confinement effects. Rather, the chaperonin would increase folding rates by preventing multimeric association that would otherwise occur in the bulk. Both RuBisCo and DM-MBP were seen to associate in solution in the experiments of Horwich (based on gel filtration and light scattering studies). According to this study, the rate of DM-MBP folding slows down at higher DM-MBP concentrations. However, under chlorine-free conditions where DM-MBP could no longer aggregate (as evidenced by lack of light scattering), the protein showed the same folding rate in the bulk as in the chaperonin cavity.

However, these results are disputed. The DM-MBP folding kinetics data published by the Hartl group (figure S1 from the supplemental section of [19]) does not show concentration dependent folding rates for DM-MBP. It has also been brought to our attention by an anonymous reviewer that the light scattering signal in the Horwich paper [18] does not decay over the hour-long observation period, during which time the majority of RuBisCo and DM-MBP should have refolded. The persistence of light scattering may, in part, be due to the fact that the aggregation of RuBisCo and DM-MBP was not fully reversible [18]. However, the full reason is not yet known (A. Horwich, personal communication). We hope for future clarification from both groups.

The controversy surrounding these experiments points to the difficulty of monitoring the folding of GroEL substrates, which are often multimeric and typically highly aggregate prone. For example, the critical aggregation concentration for RuBisCo is less than 10 nM at 25°C [16, 92, 93]. Even when folding is possible, transient aggregates can be mistaken for intermediates [194, 195]. More exotic artifacts are possible. For example, for reasons not fully understood, rhodanese may fold more rapidly at higher concentrations [156]. Standard efforts to reduce aggregation routinely require use of buffers containing molecules like bovine serum albumin (BSA), an artificial chaperone used in all the RuBisCo and DM-MBP experiments discussed here [16, 18, 19]. BSA reduces aggregation by binding non-specifically to exposed hydrophobic patches on the surface of partially folded proteins. In principle, these heterologous associations could alter protein refolding kinetics. Ideally, folding kinetics of

such proteins should be studied under dilute conditions, and in the absence molecules like detergents or BSA which can retard or otherwise change its folding kinetics.

Horwich and co-workers [196] also re-examined the work of Hartl and co-workers [19], probing the effect of altered GroEL C-terminal repeats on protein folding rates. They found that the effect of altering the cage size was much more dramatic for the double-ring mutant than for the single-ring mutant. In most instances, little change in folding rate was observed for the single-ring mutant. In the case of wild-type MDH, there was no change in rate in either the tail-multiplied double-ring or single-ring GroEL variants. This is in contrast to the rate change for DM-MBP reported by Hartl [19]. The experiments indicate that altering the cage size may not have an effect on all substrates of a given size. We note that confinement may still explain the increased folding rates for DM-MBP if the mutations lead to more extended conformations in the ensemble of unfolded states. The experiments presented in reference [196] lead Horwich and co-workers to conclude that confinement may not explain the folding rate changes in the C-terminal repeat GroE-mutants. They propose instead that observed rate changes are due to altered ATPase activity upon adding C-terminal inserts. Their experiments showed that ATPase activity increases linearly with the C-terminal insert lengths. The result of increased rates of ATP turnover in tail-multiplied double-ring GroEL mutants is a faster rate of cycling. This would not necessarily translate into faster folding. As mentioned in "The ATP driven iterative annealing model", the folding of proteins can be accelerated (or decelerated) by more frequent cycles of denaturation depending upon whether they populate long-lived off-pathway (or on-pathway) intermediate states while folding. We refer the reader to "The ATP driven iterative annealing model" and "A review of the effects of iterative denaturation" for a discussion of the effects of increasing cycling rates on folding.

Hartl and co-workers [20] responded with a new set of experiments in which they did not find a linear relationship between ATPase activity and C-terminal insert lengths; rather, they found that the ATPase activity reached a plateau at two $[GGM]_4$ inserts. Further experiments to probe whether the rate of the ATPase reaction affected folding rates were performed using the D398A mutant of GroEL that binds ATP and GroES, but hydrolyses ATP very slowly. Their experiments on GFP and rhodanese showed the same folding acceleration for the C-terminal insert D398A mutant as for the wild-type GroEL with C-terminal inserts and the SR-EL with the same tail extensions. Hartl argued that these experiments indicate that confinement and not altered ATPase activity is the reason for the observed changes in folding rate. We note that although rhodanese is not a two-state folder [156, 157], its folding kinetics appear to be crudely single-exponential and [7, 16, 20], and it is not likely to be effected by a change in the ATPase cycle rate. Again, rhodanese folds just as rapidly in non-cycling SR-EL as it does in GroEL [14, 16]. In these examples, the ATPase cycle does not appear to serve any other purpose than to cause the configuration changes that enable the chaperonin to reversibly capture and release substrate proteins.

Interaction with walls

Confinement is just one aspect of caging. Confinement theory ignores the actual nature of the interior of the chaperonin. In theoretical studies/simulations of confinement, the surface of the chaperone is modeled by a purely sterically repulsive (non-interacting hydrophilic) surface. The interior of the chaperonin is, of course, more complex, consisting of a variety of hydrophilic, hydrophobic, and charged residues that can interact with the protein and solvent, altering the folding environment and the folding pathway.

Stochastic cycling–cage-mediated annealing

It has been estimated that between 20 and 40% of the residues which in the interior of the GroEL *cis*-cavity are hydrophobic [27, 103]. The role of these interactions was investigated theoretical using lattice [103] and off-lattice [124, 125] models. These studies showed that including moderate attractive interactions with the cage can increase (or decrease) folding rates.

A novel mechanism of "stochastic cycling" (also known as "passive destabilization") has been introduced to explain how chaperonins could accelerate the folding of frustrated proteins inside the chaperonin cage [103, 125]. This is closely related to the "transient binding and release" (TBR) model proposed for minichaperones [179, 197]. In principle, a moderately attractive surface can destabilize a protein, reducing the lifetime of otherwise stable partially folded conformations [103, 109, 110, 125]. As visualized in coarse-grained computer simulations, attractive interactions with the wall can rapidly periodically unfold the proteins to which it binds [125]. The mechanism resembles in some respects the standard "iterative annealing mechanism"; however, folding occurs inside the cavity and involves thermally driven stochastic (rather than ATP-driven) cycles of binding, and unbinding from the chaperonin [103, 125]. As with the IAM, to accelerate folding, these cycles must occur more frequently than the lifetime of the protein's kinetic traps. Viewed another way, these collapsed, denatured, wall-bound states can provide an alternate route to the native state. Folding proceeds via a new pathway (through a bound intermediate state) of lower

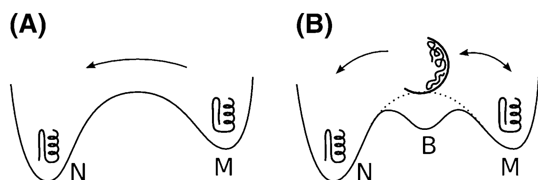


Fig. 5 One-dimensional schematic of a frustrated protein in **a** the absence and **b** the presence of a chaperone. **a** Once the protein is in the incorrect local energy minima (*M*), the protein may never be able to reach the native state (*N*) on biologically relevant time scales. **b** This simple one-dimensional cartoon is meant to convey that a sufficiently weak hydrophobic surface effectively lowers the barrier to folding by providing an alternative route to the native state. If the surface is not too hydrophobic, the energy well at the chaperonin-bound state (*B*) can be sufficiently shallow to enable spontaneous unbinding and, hence, refolding

energy than would be the case for folding in the absence of a chaperonin (see Fig. 5).

Experimental confirmation of the importance of hydrophobic residues lining the interior of GroEL can be found in the work of Hartl and co-workers [19] in which removal or mutation of the mildly hydrophobic GGM repeats in SR-EL was seen to affect folding rates. The folding of DM-MBP (a particularly energetically frustrated protein according to Hartl) was decelerated upon mutation of $[GGM]_4M$ to $[AAA]_4A$ or $[GGA]_4A$ more so than when this segment was deleted. Tail insert GroEL mutants with the [GGA] sequence did not promote folding of DM-MBP as successfully as the [GGM] mutants. Furthermore, substitution of [GGM] by [AAA] lead to restricted mobility of DM-MBP.

Stochastic cycling was originally proposed to occur within the interior of the closed GroEL + GroES + ATP cavity [103, 125]. Generally, most evidence indicates that the majority of GroEL substrates are immobilized while bound to the open GroEL *trans*-ring [58, 70–73]. However, there are experimental examples [11, 198] of denatured proteins which are not completely immobilized while bound to the open GroEL *trans*-ring. Observed motion in proteins bound to GroEL may be due to local rather than long-range fluctuations in the molecule [69]. Nevertheless, hydrophobic segments of proteins tend to be more mobile and are released from the apical domain earlier (possibly altering the order in which the protein sections are constructed in the cavity) [11]. Given this evidence of mobility, it is tempting to speculate that *some* proteins (or portions of these proteins) might be able to repeatedly bind to, and (partially) free themselves from, the open GroEL *trans*-ring during this time.

We point out that all proteins spend a significant fraction of their time (approximately 1 s per cycle) bound to the GroEL opening before GroES displaces them into the cavity [40, 52, 82]. Furthermore, some *large* GroEL substrates spend nearly all of their time bound to the open GroEL *trans*-ring before folding, because they are too big to fit inside the closed GroEL + GroES + ATP *cis*-chamber [29, 31]. However, so far, the evidence collected for these substrates suggests that GroEL passively prevents aggregation and does not accelerate their folding [31]; GroEL may even prevent folding [29]. The issue is further complicated by new evidence which suggests that the closed *cis*-chamber can expand to encapsulate some oligomeric proteins as large as 86 kDa in size [30]. The manner in which GroEL assists the folding of these large proteins is a compelling topic deserving future investigation.

Cavity may alter solvent behavior

The lining of the GroEL cavity has an overall negative charge of −42, with a number of negatively charged amino acids highly conserved (and several negatively charged clusters located near the apical domain). Mutations that altered the net charge of the cavity affected folding rates in a substrate specific manner. For instance, a D359K mutation in the context of the single-ring mutant led to rate decreases in the case of DM-MBP, a mild rate increase for rhodanese, and no effect on RuBisCo [19]. Certain mutations that reduced the negative charge also affected the mobility of DM-MBP. Neutralizing the cavity (SR-KKK(2) mutation) led to an increase in folding rates for rhodanese and a decrease in folding rates for DM-MBP. These experiments suggest that having an overall negative charge inside the chaperonin is productive for folding. From a theoretical perspective, Pande and co-workers [127] developed a novel phenomenological model for studying a confined protein that incorporates the effect of solvent through a field theory formulation of the free energy functional for water. Their studies suggest that the hydrophilic (charged) interior of the GroEL cavity upon complexation with ATP and GroES increases the density of water in the cavity (as compared to the more hydrophobic lining of the unbound GroEL). As a result, the hydrophobic effect is enhanced with respect to the bulk [129, 187], facilitating (accelerating) folding. Fully atomic simulations on charged mutants of GroEL [128] in which the number of water molecules within 1 nm of the surface was monitored further support a picture in which the charged cavity walls can alter the solvent environment and modulate the hydrophobic effect for folding.

Conclusion

The process of reviewing some of the recent available experimental data raised several questions of our own. These are the questions which would settle (in our minds, at least) whether or not GroEL can function as an active folding promoter, or if it only passively prevents aggregation:

1. *Do obligate substrates fold faster with GroEL/ES?* GroEL/ES spends 80% of the time bound to obligate ("class III") substrates [13]. However, the majority of experiments have been carried out on non-obligate ("class I" or "class II") substrates (proteins which interact with GroEL, but can fold without it in vivo), or on substrate mutants, or on model protein systems which are not natural chaperonin substrates. Thus, it is not clear whether any obligate GroEL substrate proteins actually fold faster in the presence of wild-type GroEL + ES + ATP. For *E. coli*, the proteins in Table 1 might be a good place to look. It has already been reported that one of these, dapA/dihydrodipicolinate synthase, may be accelerated by a factor of 10× in the presence of GroEL + ES + ATP [13].
2. *Is folding acceleration significant?* Accelerations reported so far for obligate substrates (other than dapA) are relatively modest: less than 1.5× for rhodanese (and again, only using a GroEL mutant). RuBisCo *may* fold 4× faster, although disagreement remains over whether GroEL/ES accelerates RuBisCo and DM-MBP folding [16, 18, 19]; neither rhodanese or RuBisCo are present in *E. coli*. Folding is accelerated by a factor of 4× for non-obligate substrates [19, 20, 44, 143], such as MDH and DHFR, which can be renatured without chaperonins [143, 199]. Folding can be accelerated by up to 50× for mutants of non-GroEL substrates [19, 22]. The most dramatic acceleration, 50×, was observed for mutant "MBPY283D" at elevated temperatures near 40°C [22].

For comparison, early estimates have suggested that GroEL may reduce the time proteins would spend in the bulk before folding by as much as a factor 50–100× [21, 148], simply by providing proteins with a safe place to fold. This factor depends on $\langle t_{\mathrm{unbound}} \rangle$, as explained below. (See also "Estimating the fraction of time proteins are exposed to the cytosol".)

Whether a modest rate acceleration (for example, a factor of 2× to 4×) is significant also depends on a related question: Approximately what fraction of GroEL substrates are accelerated? If a substantial fraction of GroEL substrates fold faster in the presence of GroEL/ES, this reduction in occupancy could free up a substantial number of GroEL chaperones for use (and their supporting machinery), even if the acceleration is only 2× or less. However, if GroEL accelerates the folding of only 1% of its protein substrates (for example), this would have a minimal effect on the on chaperone availability.

3. *Are GroEL obligate substrates sufficiently kinetically trapped to benefit from iterative denaturation?* Assuming an obligate GroEL substrate is found which folds significantly faster in the presence of GroEL, there are still many possible mechanisms to explain how this occurs. One way to distinguish between them is to search for *long-lived, monomeric* kinetic traps. Iterative denaturation by GroEL (caused by cycles of ATP hydrolysis or stochastic cycling inside the cavity) can only accelerate the folding of proteins trapped in long-lived misfolded conformations (see "A review of the effects of iterative denaturation"). Preferably, these should be monomeric kinetic traps, since the cell is more likely to use other chaperones like DnaK/J to break apart any small aggregates which can delay folding [136]. In order to accelerate folding, the lifetime of these traps must exceed the period between cycles of denaturation. While it has been established that large aggregate-prone proteins (including rhodanese and RuBisCo) have (and fold slowly because of) kinetic intermediates [93, 153–158], it is not clear that these intermediates are off-pathway intermediates, and that they are *long-lived* compared to the overall timescale required for folding. It is difficult to deduce the lifetime of these intermediate states from experiments carried out using denaturants under equilibrium conditions. These traps should be evident from bulk experiments that monitor the population of folded protein as a function of time. Let "$P(t)$" denote the fraction of proteins which have not yet committed to folding as a function of time, t. This is a decreasing function which begins at 1 and decays to 0. For most proteins, especially small single-domain, two-state folders, $P(t)$ resembles a decaying exponential: $P(t) = e^{-k_{\mathrm{F}} t}$, where k_{F} is the folding rate. Supposing that GroEL denatures proteins every 3 s (i.e., the "substrate driven" limit [40, 55, 56]), in order for this to accelerate folding, these traps must be plainly visible in the graph of $P(t)$ (with lifetimes in excess of 3 s, in this example). $P(t)$ must fail to fit to a single exponential, requiring instead a stretched or a weighted sum of decaying exponentials, using only positive weights, and at least one rate constant slower than (1/3) s^{-1}. To detect the range of fast and slow processes, ideally multiple measurements of $P(t)$ could be be taken during the first cycle period (3 s) of folding. With some exceptions [144], the majority of protein folding kinetics data we know of do not show obvious evidence of such long-lived traps. While some of the GroEL substrates studied so far, like RuBisCo and rhodanese, are kinetically trapped [156, 157], and fold very slowly, they also *appear* to have ordinary decaying exponential folding kinetics (at least under conditions where such measurements are possible) [7, 16, 143]. It is not surprising that these proteins do not seem to benefit from cycles of iterative denaturation (under these same conditions, see below).

The folding GroEL substrates is especially challenging to monitor, because these are precisely the proteins which have the most difficulty folding unassisted under relevant physiological conditions. The population of folded protein $[1-P(t)]$, is usually difficult to measure directly under dilute enough conditions to avoid self-association. Artificial buffers which are used to get around these problems could (in principle at least) introduce serious artifacts into the protein's folding kinetics, for example by altering the population of various intermediate states in the folding process [18, 156]. In order to test the IAM (or stationary IAM) theories, it is probably sufficient to measure the folding kinetics ($P(t)$) in the presence of non-cycling chaperonin mutants (like SR-EL [16, 17]), since any resulting confinement effects are unlikely to completely mask the symptoms of kinetic frustration [116, 125]. More generally, single molecule experiments may be necessary to investigate the folding of highly aggregate prone proteins.

We note that difficulty in calibrating the "base lines", i.e., the signals corresponding to $P(t) = 0$ and $P(t) = 1$, could falsely make a kinetically trapped protein appear like a two-state folder with single-exponential folding kinetics.

4. *How many substrates remain bound to GroEL until folding?* ATP-driven denaturation (the traditional IAM) cannot increase the yield of kinetically trapped aggregate-doomed proteins [148] unless proteins remain bound to GroEL until folding (the stationary IAM). In vivo (or under conditions of extreme macromolecular crowding [180]), at least one protein (rhodanese) appears to remain bound to a single GroEL chaperonin ring until folding [7]. How widespread is this behavior among other stringent GroEL substrates?
5. *What is $\langle t_{\mathrm{unbound}} \rangle$ in vivo?* The benefits of the protective cage cannot be rigorously assessed until we can estimate the total amount of time obligate GroEL substrates spend unbound in the bulk during each cycle before binding to GroEL, $\langle t_{\mathrm{unbound}} \rangle$. (See "Estimating the fraction of time proteins are exposed to the cytosol".) Unfortunately, this is difficult to predict mostly due to uncertainty in the concentrations of unbound, *available* DnaK/J, GroEL, and other chaperones in the cell, in addition to uncertainty in the way they interact. (See "Estimating the fraction of time proteins are exposed to the cytosol" and [12, 13, 74–76].)

Summary

For the few stringent in vivo GroEL substrates investigated so far, GroEL appears to behave primarily in a passive manner, simply encapsulating misfolded proteins and protecting them from aggregation as they fold. However, in theory, GroEL/ES should be able to accelerate (and also retard) protein folding.

Again, GroEL/ES is a promiscuous chaperone which interacts with a diverse group of 250 protein substrates, 85 of which are strictly unable to fold in cells with GroEL knockouts [13]. We cannot rule out the possibility that GroEL/ES interacts with at least *some* of these proteins in a more interesting way. GroEL/ES must satisfy the competing demands of a large variety of (*obligate*) protein substrates. Introducing mutations that optimize the folding of one substrate are known to harm GroEL's ability to promote the folding of others [200, 201]. There is no need to impose a universal mechanism to explain how GroEL/ES works. While it may not be plausible to exhaustively investigate the folding kinetics of all 85 or so obligate GroEL substrates in the presence of GroEL/ES, we could start with the most soluble substrates [13, 88] (some of which are listed in Table 1).

We initially set out to present a theorist's view of chaperonin-mediated protein folding. As we look back at this body of work, we have come to realize that we are in critical need of experimental folding kinetics data for *obligate* GroEL substrates. These data hold the potential of resolving whether GroEL is ever capable of acting as a folding catalyst in vivo.

Acknowledgments We have greatly benefited from the input of others. I would like to thank Hideki Taguchi, Hays Rye, Ulrich Hartl, Arthur Horwich, and Adrian Apetri for providing thoughtful answers to numerous detailed questions. Hideki Taguchi and Hays Rye made corrections and Hideki Taguchi generously provided figure excerpts and access to recent unpublished solubility data from his laboratory [88]. We would also like to thank Ben Schuler and Ulrich Hartl for helpful advice and discussions. This work was supported by the NSF grant #0642086 and a grant by the David and Lucile Packard Foundation.

Appendix A: Estimating the fraction of time proteins are exposed to the cytosol

In an earlier work, we developed a formula to describe the folding of any protein in the presence of chaperones which cyclically unfold their substrates [148]. We used this to prove that the iterative annealing model (IAM) is not optimal for, and does not explain, the chaperonin-mediated folding of aggregate-prone substrates. Instead, an optimal chaperonin would bind to its substrates only once, releasing it only upon folding, lending support to the Anfinsen cage model. Unfortunately, many simplifying assumptions were made; for example, we ignored the fact that proteins interact with a variety of chaperones other than GroEL, and we ignored the fact that some proteins remain bound to the chaperone over multiple ATPase cycles. We also implicitly assumed that sufficient GroEL chaperonins are present to handle demand, and we ignored transient stress. However,

the conclusions of that study remain valid when these assumptions are relaxed. In the sections that follow we briefly review and generalize these kinetic arguments.

The many-cycle assumption

As explained earlier [21, 148], chaperonins like GroEL can reduce the time that proteins spend unprotected in the cytosol before folding by a fraction (denoted f_{bulk}) which can be estimated from the ratios of the average time spent bound and unbound from GroEL during each cycle:

$$f_{bulk} = \frac{\langle t_{unbound} \rangle}{\langle t_{unbound} \rangle + \langle t_{bound} \rangle} \quad (1)$$

To justify this, we must assume that proteins undergo many cycles of binding to GroEL and release into the cytosol before folding, as suggested by [16, 80, 91, 102]. Here $\langle\rangle$ denotes the average, and t_{bound} is the total time spent bound to the chaperone (while either immobilized or protected).

$$\langle t_{bound} \rangle = \langle t_{hold} \rangle + \langle t_{protect} \rangle \quad (2)$$

(See Figs. 1 and 2, for definitions of $t_{unbound}$, t_{hold}, and $t_{protect}$.) We considered what happens if you abandon this assumption in "The stationary iterative annealing model".

Substrates do not always unbind from GroEL

A complication arises due the fact that some GroEL substrates (rhodanese) do not unbind during every ATPase cycle [7, 12, 180]. Suppose that f_{ub} indicates the probability that the protein substrate can successfully unbind itself from GroEL once the GroES lid departs. (This occurs following ATP hydrolysis; see Figs. 1, caption and 2). If so, then in that case it will require $1/f_{ub}$ cycles for the protein to successfully free itself from GroEL, on average, (note: $1/f_{ub} \geq 1$). This means it would remain bound to GroEL for a duration of approximately $(\langle t_{bound} \rangle + \langle t_{unbound} \rangle)/f_{ub}$ seconds, instead of $\langle t_{bound} \rangle$ seconds. (Minor correction: We note that during the first of these $1/f_{ub}$ cycles, the protein is initially unbound, so to be precise, we should not have included one of these "unbound" time intervals.)

Once finally released, if the protein has not yet folded, it will have to rebind to GroEL requiring $t_{unbound}$ seconds. As long as this process occurs multiple times before folding, the arguments we have made so far continue to apply, and we can replace $\langle t_{bound} \rangle$ in Eq. 1, with $\langle t_{unbound} \rangle\ (1/f_{ub} - 1) + \langle t_{bound} \rangle / f_{ub}$. (The "$-1$" comes from the correction discussed above.) This yields:

$$f_{bulk} = f_{ub} \times \frac{\langle t_{unbound} \rangle}{\langle t_{unbound} \rangle + \langle t_{bound} \rangle} \quad (3)$$

For example, rhodanese in vitro escapes GroEL every four cycles on average ($f_{ub} \approx 0.25$), frequently enough so that it probably escapes GroEL a couple times before folding [180]. Rhodanese is an extraordinarily slow folder, requiring 7 min to fold on average [20], corresponding to 7–60 full, two-ring ATPase cycles and consuming at least 130 molecules of ATP [149].

We note that sometimes unbinding does *not* occur multiple times before folding. For example, in vivo (or in the presence of a crowding agent), rhodanese typically remains bound to GroEL until folding [7]. In that case, the situation is fundamentally different, and we have to consider the issues raised in "The stationary iterative annealing model". Equation 3 does not apply to rhodanese in vivo. We note that this is not an issue for many stringent GroEL substrates. Others (like RuBisCo) unbind from GroEL after every ATPase cycle ($f_{ub} \approx 1$) [16].

Under steady-state conditions

In the absence of stress ("steady-state") conditions (see "The steady-state assumption"), it is more relevant to consider:

$$f^{ss}_{bulk} = f_{ub} \times \frac{\langle t_{unbound} \rangle}{\langle t_{unbound} \rangle + \langle t_{protect} \rangle} \quad (4)$$

During the time interval ($\langle t_{hold} \rangle$) that proteins are either bound to the open GroEL *trans* ring, or bound to auxiliary chaperones like DnaK/J, they are unable to fold or aggregate and for all practical purposes, they are immobilized (although they may in fact be able to move). Under steady-state conditions we neglect to consider any time spent by the protein in these "immobilized" states; in other words, we have substituted $\langle t_{hold} \rangle = 0$ into Eqs. 2, 3. Temporary delays (no matter how long) which have no other effect than to immobilize the protein, should have no effect on a protein's likelihood of eventually folding or aggregating; that is, assuming the risk of aggregation in the bulk remains constant over time. We note that under steady-state conditions, the concentration of denatured proteins, and rate of aggregation, should not fluctuate significantly over time, at least not during the time for most proteins to fold. (See "The steady-state assumption" for details. We elaborate further in "Why should we ignore immobile states?".)

Since $\langle t_{unbound} \rangle$ is typically far more rapid than $\langle t_{protect} \rangle$, this is a considerable reduction ($f^{ss}_{bulk} \ll 1$).

The role of HSP70/HSP40 and other ancillary chaperones

For GroEL substrates, it appears that other chaperones such as trigger factor, HSP70/40 (DnaK/J) and their associated nucleotide exchange factors (GrpE), prevent unfolded

protein chains from aggregating as they wait for GroEL [3, 13, 23, 25]. After being ejected from GroEL, proteins that are still unfolded are likely to bind to chaperones like DnaK/J before rebinding to GroEL. The average of the total time that proteins spend unprotected in the bulk during this time (represented by $\langle t_{\text{unbound}} \rangle$) is a complicated function of the DnaK/J, GrpE, and dimeric trigger factor concentrations, in addition to the nucleotide binding, release, and hydrolysis rates (for example, see [75]), not to mention the length of the substrate protein (as suggested by [134]). However, as far as GroEL substrates are concerned, the only role of these auxiliary chaperones is to reduce $\langle t_{\text{unbound}} \rangle$. Whether they are successful is a separate issue, and it does not effect our conclusion regarding the optimal behavior of GroEL. Equations 1, 3, and 4 still remain remain valid, regardless of the presence of other chaperones.

Appendix B: The steady-state assumption

As mentioned in the "Introduction", GroEL/ES performs maintenance duties in the cell and is always present at high concentrations, even in the absence of external stress [5]. Under non-stress, steady-state conditions, it seems reasonable to assume that concentration of each species of protein remains roughly constant over time; or at least these concentrations do not fluctuate significantly during the course of a single folding event. This is important for understanding the mechanism of GroEL.

Under steady-state conditions, the *only* way to reduce aggregation is to reduce the concentration of non-native proteins in the cell, which can only be done by reducing the average time each protein spends unprotected in the cytosol ("bulk") before folding, $\langle t_{\text{bulk}} \rangle$ [148, 197]. During this time, proteins are susceptible to aggregation. It is convenient to think of this as the product of the average folding time $\langle t_{\text{F}} \rangle$ (under dilute conditions), and the *fraction* of that time which is spent in the bulk f_{bulk}.

$$\langle t_{\text{bulk}} \rangle = \langle t_{\text{F}} \rangle \times f_{\text{bulk}} \tag{5}$$

In this way, we can compare the benefits of folding acceleration (reducing $\langle t_{\text{F}} \rangle$) with the benefits of sequestration/encapsulation (reducing f_{bulk}).

Appendix C: A review of the effects of iterative denaturation

It is useful to ask: under what conditions would iterative denaturation speed up protein folding? Rephrasing earlier arguments [77, 104], let:

t = the time that has elapsed since the protein was first introduced into the cytosol in its unfolded state

$P(t)$ = the probability that the polymer has not yet reached a folding-committed conformation after time t has elapsed under dilute conditions in the absence of chaperones

Assuming that the only effect a chaperone has on the protein is to completely denature it once every τ_{D} s then the probability that the protein has not yet folded after N cycles of binding and release from the chaperone is $[P(\tau_{\text{D}})]^N$. In order for a protein to benefit from chaperone cycling:

$$P(N\tau_{\text{D}}) > [P(\tau_{\text{D}})]^N \tag{6}$$

For any protein which folds with a single well defined folding rate (k_{F}, for example, two-state folders, *or* proteins with only short-lived intermediates) $P(t)$ must resemble a decaying exponential ($P(t) = e^{-k_{\text{F}} t}$). For these proteins, $P(N\tau_{\text{D}}) = [P(\tau_{\text{D}})]^N$. Only proteins for which $P(t)$ decays more slowly at long times (for example, proteins which can fall into kinetic traps) can satisfy this inequality.

It is possible to predict the average folding time, $\langle t_{\text{F}}^{\text{ss}} \rangle$ in the presence of iterative denaturation (at frequency $\lambda_{\text{D}}^{\text{ss}}$) for any protein, assuming the folding kinetics of that protein (under dilute conditions, $P(t)$) are known [148]:

$$\langle t_{\text{F}}^{\text{ss}} \rangle = \frac{1}{\lambda_{\text{D}}^{\text{ss}}} \left[\left(\lambda_{\text{D}}^{\text{ss}} \int_0^{\infty} P(t) e^{-\lambda_{\text{D}}^{\text{ss}} t} \, \mathrm{d}t \right)^{-1} - 1 \right]^{-1} \tag{7}$$

$P(t)$ can be measured directly from bulk experiments, for example using florescence resonance energy transfer spectroscopy, or using enzyme assays applied to aliquots taken at regular intervals. By substituting $P(t) = e^{-k_{\text{F}} t}$, we can see again that proteins with two-state folding kinetics (rate k_{F} s^{-1}) would not benefit from iterative denaturation (in agreement with [104]).

The frequency of denaturation, $\lambda_{\text{D}}^{\text{ss}}$, refers to the frequency at which proteins are denatured as a result of ATP-driven chaperonin binding and release: Specifically:

$$\lambda_{\text{D}}^{\text{ss}} \approx 1/\langle \tau_{\text{D}}^{\text{ss}} \rangle \quad \text{where:} \tag{8}$$

$$\langle \tau_{\text{D}}^{\text{ss}} \rangle = \langle t_{\text{unbound}} \rangle + \langle t_{\text{protect}} \rangle \tag{9}$$

$$= \langle t_{\text{unbound}} \rangle + \langle t_{\text{protect}} \rangle + \langle t_{\text{hold}} \rangle \ldots \text{in the limit that} \langle t_{\text{hold}} \rangle \to 0 \tag{10}$$

The "ss" superscripts are to remind us that under steady-state conditions, we should not consider the time proteins spend while immobilized during each cycle $\langle t_{\text{hold}} \rangle$ (See "Why should we ignore immobile states?". Note that the actual folding time $\langle t_{\text{F}} \rangle$ can be inferred from $\langle t_{\text{F}}^{\text{ss}} \rangle$ by

estimating the fraction of time a protein would have spent immobilized while folding. See Eq. 11 of "Why should we ignore immobile states?")

The chaperone-mediated folding of aggregate prone substrates

Recall that under steady-state conditions, chaperones ability to prevent aggregation is entirely determined by how much chaperones reduce the time proteins spend in the bulk before folding, $\langle t_{\rm bulk}\rangle = \langle t_{\rm F}\rangle \times f_{\rm bulk} = \langle t_{\rm F}^{\rm ss}\rangle \times f_{\rm bulk}^{\rm ss}$ (See "Why should we ignore immobile states?".) Reducing the value of $\langle t_{\rm bulk}\rangle$ reduces aggregation and increases the yield. If we restrict ourselves further to the set of proteins which do not remain bound during *every* ATPase cycle, then we can use Eq. 4. Substituting it, along with Eqs. 7 and 10, results in a formula for $\langle t_{\rm bulk}\rangle$ which *decreases* as the cycle frequency $\lambda_{\rm D}^{\rm ss} \approx 1/(\langle t_{\rm bound}\rangle + \langle t_{\rm protect}\rangle)$ *decreases*; the result is proportional to Eq. 7 from [148]. In other words, for this broad set of proteins, GroEL/ES should cycle *slowly* (maximize $\langle t_{\rm protect}\rangle$). There is no incentive to cycle rapidly, except perhaps to free up chaperones and assist the folding of other proteins. This contradicts the conclusion of the traditional IAM. Again, the cycle for GroEL/ES requires on the order of 10^1 s.

Appendix D: Why should we ignore immobile states?

Simple kinetics models of GroEL/ES behavior assume that the entire time a protein is bound to GroEL it is either able to continue folding [148], or immobilized [113]. In reality, proteins may spend a fraction of their time with GroEL mobile or immobilized. However, under steady-state conditions, these immobile states have no effect. Increasing or decreasing the duration of these frozen states do not tip the balance toward one outcome (folding) or the other (aggregation), at least not under *steady-state* conditions when, presumably, the rate of transition to either of these outcomes is not changing over time.

Motivating example

We have argued Eqs. 12 and 4 without providing an algebraic proof. If it helps the reader, we can motivate Eqs. 12, and 4, by calculating both $\langle t_{\rm F}^{\rm ss}\rangle$ and $f_{\rm bulk}^{\rm ss}$, and show that their product remains equal to $\langle t_{\rm bulk}\rangle$ from Eq. 5:

To motivate this with a concrete example, it is convenient to imagine a hypothetical chaperone system which does not immobilize its substrates, ($\langle t_{\rm bulk}\rangle = 0$), and which otherwise behaves exactly like the GroEL/ES-DnaK/J chaperone system in all other respects, denaturing protein substrates with every ATPase cycle. Of course, the resulting reduction in time spent bound to the chaperone might free up chaperones and increase chaperone availability. However, we ignore this effect here. Here, we imagine a hypothetical chaperone for which $\langle t_{\rm unbound}\rangle$ and $\langle t_{\rm protect}\rangle$ remain unaffected as $\langle t_{\rm hold}\rangle \to 0$.

$\langle t_{\rm F}^{\rm ss}\rangle$ and $f_{\rm bulk}^{\rm ss}$ denote the folding time, and fraction of time spent in the bulk, folding under the influence of this new hypothetical chaperone system (with $\langle t_{\rm hold}\rangle = 0$). The formula for $f_{\rm bulk}^{\rm ss}$ is given in Eq. 4 of "Estimating the fraction of time proteins are exposed to the cytosol". The formula for $\langle t_{\rm F}^{\rm ss}\rangle$ is given in Eq. 7 of "A review of the effects of iterative denaturation". How does this $\langle t_{\rm F}^{\rm ss}\rangle$ compare with the real folding time in vivo, $\langle t_{\rm F}\rangle$?

In the presence of this hypothetical chaperone system, proteins would fold faster because they no longer have to spend a certain fraction of each cycle immobilized and waiting. Assuming many cycles of binding and release, this should reduce the folding time by the fraction of time proteins are *not* immobilized during each cycle (shown in parenthesis in Eq. 11).

$$\langle t_{\rm F}^{\rm ss}\rangle = \langle t_{\rm F}\rangle \times \left(\frac{\langle t_{\rm unbound}\rangle + \langle t_{\rm protect}\rangle}{\langle t_{\rm unbound}\rangle + \langle t_{\rm hold}\rangle + \langle t_{\rm protect}\rangle}\right) \tag{11}$$

Multiplying Eqs. 4 and 11, and substituting Eq. 2, recovers Eqs. 3 and 5:

$$\langle t_{\rm bulk}\rangle = \langle t_{\rm F}^{\rm ss}\rangle \times f_{\rm bulk}^{\rm ss} \tag{12}$$

This shows that ignoring immobilized states (or equivalently, setting $\langle t_{\rm hold}\rangle = 0$) has no effect on $\langle t_{\rm bulk}\rangle$. Under the influence of such a chaperone, proteins would spend *the same amount of time in the bulk* before folding $\langle t_{\rm bulk}\rangle$, and would be no more or less likely to aggregate. Thus, a hypothetical chaperone without immobilized states would prevent just as much aggregation as a real chaperone (under steady-state conditions). Hence, we can justifiably ignore these immobilized states.

References

1. Minton AP (2000) Implications of macromolecular crowding for protein assembly. Curr Opin Struct Biol 10:34–35
2. Frydman J (2001) Folding of newly translated proteins in vivo: the role of molecular chaperones. Annu Rev Biochem 70:603–647
3. Hartl UF, Hayer-Hartl M (2009) Converging concepts of protein folding in vitro and in vivo. Nat Struct Mol Biol 16(6):574–581
4. Hartl UF, Hayer-Hartl M (2002) Molecular chaperones in the cytosol: from nascent chain to folded protein. Science 295:1852–1858
5. Narberhaus F (2002) α-crystallin-type heat shock proteins: socializing minichaperones in the context of a multichaperone network. Microbiol Mol Biol R 66(1):64–93
6. Saibil HR, Horwich AL, Fenton WA (2002) Allostery and protein substrate conformational change during GroEL/GroES-mediated protein folding. Adv Protein Chem 59:45–72

7. Ewalt KL, Hendrick JP, Houry WA, Hartl FU (1997) In vivo observation of polypeptide flux through the bacterial chaperonin system. Cell 90(3):491–500
8. Houry WA, Frishman D, Eckerskorn C, Lottspeich F, Hartl FU (1999) Identification of in vivo substrates of the chaperonin GroEL. Nature 402:147–154
9. Shtilerman M, Lorimer GH, Englander SW (1999) Chaperonin function: folding by forced unfolding. Science 284(5415):822–825
10. Lin Z, Madan D, Rye HS (2008) GroEL stimulates protein folding through forced unfolding. Nat Struct Mol Biol 15(3):303–311
11. Sharma S, Chakraborty K, Müller BK, Astola N, Tang Y, Lamb DC, Hayer-Hartl M, Hartl FU (2008) Monitoring protein conformation along the pathway of chaperonin-assisted folding. Cell 133:142–153
12. Mayhew M, da Silva ACR, Martin J, Erdjument-Bromage H, Tempst P, Hartl FU (1996) Protein folding in the central cavity of the GroEL–GroES chaperonin complex. Nature 379(6564):420–426
13. Kerner MJ, Naylor DJ, Ishihama Y, Maier T, Chang HC, Stines AP, Georgopoulos C, Frishman D, Hayer-Hartl M, Mann M, Hartl FU (2005) Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. Cell 122(2):209–220
14. Weissman JS, Rye HS, Fenton WA, Beechem JM, Horwich AL (1996) Characterization of the active intermediate of a GroEL–GroES-mediated protein-folding reaction. Cell 84(3):481–490
15. Hayer-Hartl MK, Weber F, Hartl FU (1996) Mechanism of chaperonin action: GroES binding and release can drive GroEL-mediated protein folding in the absence of ATP hydrolysis. EMBO J 15(22):6111–6121
16. Brinker A, Pfeifer G, Kerner MJ, Naylor DJ, Hartl FU, Hayer-Hartl M (2001) Dual function of protein confinement in chaperonin-assisted protein folding. Cell 107(2):223–233
17. Weissman JS, Hohl CM, Kovalenko O, Kashi Y, Chen S, Braig K, Saibil HR, Fenton WA, Horwich AL (1995) Mechanism of GroEL action—productive release of polypeptide from a sequestered position under GroES. Cell 83:577–587
18. Apetri AC, Horwich AL (2008) Chaperonin chamber accelerates protein folding through passive action of preventing aggregation. Proc Natl Acad Sci USA 105(45):17351–17355
19. Tang Y-C, Chang H-C, Roeben A, Wischnewski D, Wischnewski N, Kerner MJ, Hartl FU, Hayer-Hartl M (2006) Structural features of the GroEL–GroES nano-cage required for rapid folding of encapsulated protein. Cell 125:903–914
20. Tang Y-C, Chang H-C, Chakraborty K, Hartl FU, Hayer-Hartl M (2008) Essential role of the chaperonin folding compartment in vivo. EMBO J 27:1458–1468
21. Ranson NA, Burston SG, Clarke AR (1997) Binding, enscapsulation and ejection: substrate dynamics during a chaperonin-assisted folding reaction. J Mol Biol 266:656–664
22. Sparrer H, Rutkat K, Buchner J (1997) Catalysis of protein folding by symmetric chaperone complexes. Proc Natl Acad Sci USA 94:1096–1100
23. Teter SA, Houry WA, Ang D, Tradler T, Rockabrand D, Fischer G, Blum P, Georgopoulos C, Hartl FU (1999) Polypeptide flux through bacterial Hsp70:DnaK cooperates with Trigger Factor in chaperoning nascent chains. Cell 97:755–765
24. Ellis RJ, Hartl FU (1996) Protein folding in the cell: competing models of chaperonin function. FASEB 10:20–26
25. Hesterkamp T, Bukau B (1998) Role of the DnaK and HscA homologs of Hsp70 chaperones in protein folding in *E.coli*. EMBO J 17(16):4818–4828
26. Bukau B, Horwich AL (1998) The Hsp70 and Hsp60 review. Cell 92:351–366
27. Xu Z, Horwich AL, Sigler PB (1997) The crystal structure of the asymmetric GroEL–GroES-(ADP)7 chaperonin complex. Nature 388:741–750
28. Braig K, Otwinowski Z, Hegde R, Boisvert DC, Joachimiak A, Horwich AL, Sigler PB (1994) The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. Nature 371:578–586
29. Chaudhuri TK, Farr GW, Fenton WA, Rospert S, Horwich AL (2001) GroEL/GroES-mediated folding of a protein too large to be encapsulated. Cell 107:235–246
30. Chen D-H, Song J, Chuang DT, Chiu W, Ludtke SJ (2006) An expanded conformation of single-ring groel–groes complex encapsulates an 86 kda substrate. Structure 14:1711–1722
31. Ayling A, Baneyx F (1996) Influence of the GroE molecular chaperone machine on the in vitro refolding of *Escherichia coli* β-galactosidase. Prot Sci 5(3):478–487
32. Chuang JL, Wynn RM, Song J, Chuang DT (1999) Groel/groes-dependent reconstitution of $\alpha_2\beta_2$ tetramers of human mitochondrial branched chain α-ketoacid decarboxylase. obligatory interaction of chaperonins with an $\alpha\beta$ dimeric intermediate. J Biol Chem 274(15):478–487
33. Fenton WA, Horwich AL (2003) Chaperonin-mediated protein folding: fate of substrate polypeptide. Q Rev Biophys 36:229–256
34. Saibil HR, Zheng D, Roseman AM, Hunter AS, Watson GMF, Chen S, auf der Mauer A, O'Hara BP, Wood SP, Mann NH, Barnettt LK, Ellis RJ (1993) ATP induces large quaternary rearrangements in a cage-like chaperonin structure. Curr Biol 3(5):265–273
35. Li Y, Gao X, Chen L (2009) GroEL recognizes an amphipathic helix and binds to the hydrophobic side. J Biol Chem 284(7):4324–4331
36. Elad N, Farr GW, Clare DK, Orlova EV, Horwich AL, Saibil HR (2007) Topologies of a substrate protein bound to the chaperonin GroEL. Mol Cell 26(3):415–426
37. Clare DK, Bakkes PJ, van Heerikhuizen H, van der Vies SM, Saibil HR (2009) Chaperonin complex with a newly folded protein encapsulated in the folding chamber. Nature 457:107–111
38. Stan G, Lorimer GH, Thirumalai D, Brooks BR (2007) Coupling between allosteric transitions in GroEL and assisted folding of a substrate protein. Proc Natl Acad Sci USA 104(21):8803–8808
39. Horovitz A, Willison KR (2005) Allosteric regulation of chaperonins. Curr Opin Struct Biol 15(6):646–651
40. Rye HS, Roseman AM, Chen S, Furtak K, Fenton WA, Siabil HR, Horwich AL (1999) GroEL–GroES cycling: ATP and nonnative polypeptide direct alternation of folding-active rings. Cell 97(3):325–338
41. Roseman AM, Chen SX, White H, Braig K, Saibil HR (1996) The chaperonin ATPase cycle: mechanism of allosteric switching and movements of substrate-binding domains in GroEL. Cell 87:241–251
42. Yifrach O, Horovitz A (1994) Two lines of allosteric communication in the oligomeric chaperonin GroEL are revealed by the single mutation Arg196 → Ala. J Mol Biol 243(3):397–401
43. Yifrach O, Horovitz A (1995) Nested cooperativity in the ATPase activity of the oligomeric chaperonin GroEL. Biochemistry 34(16):5303–5308
44. Rye HS, Burston SG, Fenton WA, Beechem JM, Xu Z, Sigler PB, Horwich AL (1997) Distinct actions of *cis* and *trans* ATP within the double ring of the chaperonin GroEL. Nature 388:792–798
45. Kad NM, Ranson NA, Cliff MJ, Clarke AR (1998) Asymmetry, commitment and inhibition in the GroE ATPase cycle impose alternating functions on the two GroEL rings. J Mol Biol 278:267–278

46. Yifrach O, Horovitz A (1996) Allosteric control by ATP of non-folded protein binding to GroEL. J Mol Biol 255(3):356–361
47. Inbar E, Horovitz A (1997) GroES promotes the T to R transition of the GroEL ring distal to GroES in the GroEL–GroES complex. Biochemistry 36(40):12276–12281
48. Horovitz A, Fridmann Y, Kafri G, Yifrach O (2001) Review: allostery in chaperonins. J Struct Biol 135(2):104–114
49. Ma J, Karplus M (1998) The allosteric mechanism of the chaperonin GroEL: a dynamic analysis. Proc Natl Acad Sci USA 95(15):8502–8507
50. Cliff MJ, Kad NM, Hay N, Lund PA, Webb MR, Burston SG, Clarke AR (1999) A kinetic analysis of the nucleotide-induced allosteric transitions of GroEL. J Mol Biol 293:667–684
51. Gresham JS (2004) Allostery and GroEL: exploring the tenets of nested cooperativity. Ph.D. Thesis, University of Maryland, College Park
52. Madan D, Lin Z, Rye HS (2008) Triggering protein folding within the GroEL–GroES complex. J Biol Chem 283(46): 32003–32013
53. Ranson NA, Clare DK, Farr GW, Houldershaw D, Horwich AL, Saibil HR (2006) Allosteric signaling of ATP hydrolysis in GroEL–GroES complexes. Nat Struct Mol Biol 13(2):147–152
54. Jabarak R, Westley J, Dungan JM, Horowitz P (1993) A chaperone-mimetic effect of serum albumin on rhodanese. J Biochem Toxicol 8(1):41–48
55. Grason JP, Gresham JS, Lorimer GH (2008) Setting the chaperonin timer: a two-stroke, two-speed, protein machine. Proc Natl Acad Sci USA 105(45):17339–17344
56. Grason JP, Gresham JS, Widjaja L, Wehri SC, Lorimer GH (2008) Setting the chaperonin timer: the effects of k+ substrate protein on ATP hydrolysis. Proc Natl Acad Sci USA 105(45):17334–17338
57. Melkani GC, Zardeneta G, Mendoza JA (2000) The ATPase activity of GroEL is supported at high temperatures by divalent cations that stabilize its structure. Biometals 16(3):479–484
58. Viitanen PV, Lubben TH, Reed J, Goloubinoff P, O'Keefe DP, Lorimer GH (1990) Chaperonin-facilitated refolding of ribulose bisphosphate carboxylase and ATP hydrolysis by chaperonin 60 (GroEL) are potassium dependent. Biochemistry 29(24): 5665–5671
59. Goloubinoff P, Christeller JT, Gatenby AA, Lorimer H (1989) Reconstitution of active dimeric ribulose bisphosphate carboxylase from an unfolded state. Nature 342(6252):884–889
60. Todd MJ, Viitanen PV, Lorimer GH (1993) Hydrolysis of adenosine 5-triphosphate by *Escherichia coli* GroEL: effects of GroES and potassium ion. Biochemistry 32(33):8560–8567
61. Laminet AA, Ziegelhoffer T, Georgopoulos C, Plückthun A (1990) The *Escherichia coli* heat shock proteins GroEL and GroES modulate the folding of the beta-lactamase precursor. EMBO J 9(7):2315–2319
62. Sameshima T, Ueno T, Iizuka R, Ishii N, Terada N, Okabe K, Funatsu T (2008) Football- and bullet-shaped GroEL–GroES complexes coexist during the reaction cycle. J Biol Chem 283(35):23765–23773
63. Liu C-P, Perrett S, Zhou J-M (2005) Dimeric trigger factor stably binds folding-competent intermediates and cooperates with the DnaK–DnaJ–GrpE chaperone system to allow refolding. J Biol Chem 280(14):13315–13320
64. Lin Z, Rye HS (2004) Expansion and compression of a protein folding intermediate by GroEL. Mol Cell 16:23–34
65. Ueno T, Taguchi H, Tadakuma H, Yoshida M, Funatsu T (2004) GroEL mediates protein folding with a two successive timer mechanism. Mol Cell 14:423–434
66. Taguchi H, Ueno T, Tadakuma H, Yoshida M, Funatsu T (2001) Single-molecule observation of protein–protein interactions in the chaperonin system. Nat Struct Biol 19:861–865
67. Cliff MJ, Limpkin C, Cameron A, Burston SG, Clarke AR (2006) The GroE encapsulation mechanism: elucidation of steps in the capture of a protein substrate. J Biol Chem 281(30): 21266–21275
68. Yokokawa M, Wada C, Ando T, Sakai N, Yagi A, Yoshimura SH, Takeyasu K (2006) Fast-scanning atomic force microscopy reveals the ATP/ADP-dependent conformational changes of GroEL. EMBO J 25(19):4567–4576
69. Hillger F, Hänni D, Nettels D, Geister S, Grandin M, Textor M, Schuler B (2008) Probing protein–chaperone interactions with single-molecule fluorescence spectroscopy. Angew Chem Int Ed 47:6184–6188
70. Badcoe IG, Smith CJ, Wood S, Halsall DJ, Holbrook JJ, Lund P, Clarke AR (1991) Binding of a chaperonin to the folding intermediates of lactate dehydrogenase. Biochemistry 30(38):9195–9200
71. Gray TE, Fersht AR (1993) Refolding of barnase in the presence of GroE. J Mol Biol 234(4):1197–1207
72. Corrales FJ, Fersht AR (1996) Toward a mechanism for GroEL.GroES chaperone activity: an ATPase-gated and -pulsed folding and annealing cage. Proc Natl Acad Sci USA 93(9):4509–4512
73. Jackson GS, Staniforth RA, Halsall DJ, Atkinson T, Holbrook JJ, Clarke AR, Burston SG (1993) Binding and hydrolysis of nucleotides in the chaperonin catalytic cycle: implications for the mechanism of assisted protein folding. Biochemistry 32:2554–2563
74. Lorimer GH (1996) A quantitative assessment of the role of the chaperonin proteins in protein folding in vivo. FASEB 10:5–9
75. Hu B, Mayer MP, Tomita M (2006) Role of the DnaK and HscA homologs of Hsp70 chaperones in protein folding in *E.coli*. Biophys J 91:496–507
76. Tomoyasu T, Ogura T, Tatsuta T, Bukau B (1998) Levels of DnaK and DnaJ provide tight control of heat shock gene expression and protein repair in *Escherichia coli*. Mol Microbiol 30(3):567–581
77. Todd MJ, Lorimer GH, Thirumalai D (1996) Chaperonin-facilitated protein folding: optimization of rate and yield by an iterative annealing mechanism. Proc Natl Acad Sci USA 93(9):4030–4035
78. Horwich AL, Farr GW, Fenton WA (2006) GroEL–GroES-mediated protein folding. Chem Rev 106:1917–1930
79. Koike-Takeshita A, Yoshida M, Taguchi H (2008) Revisiting the GroEL–GroES reaction cycle via the symmetric intermediate implied by novel aspects of the GroEL(D398A) mutant. J Biol Chem 283(35):23774–23781
80. Todd MJ, Viitanen PV, Lorimer GH (1994) Dynamics of the chaperonin ATPase cycle: implications for facilitated protein folding. Science 265(5172):659–666
81. Nelson DL, Cox MM (2000) Lehninger principles of biochemistry, 3rd edn. Worth, New York
82. Motojima F, Chaudhry C, Fenton WA, Farr GW, Horwich AL (2004) Substrate polypeptide presents a load on the apical domains of the chaperonin GroEL. Proc Natl Acad Sci USA 101(42):15005–15012
83. Chaudhry C, Farr GW, Todd MJ, Rye HS, Brunger AT, Adams PD, Horwich AL, Sigler PB (2003) Role of the γ-phosphate of ATP in triggering protein folding by GroEL–GroES: function, structure and energetics. EMBO J 22(19):4877–4887
84. Chapman E, Farr GW, Fenton WA, Johnson SM, Horwich AL (2008) Requirement for binding multiple ATPs to convert a GroEL ring to the folding-active state. Proc Natl Acad Sci USA 105(49):19205–19210
85. Lorimer G (1997) Folding with a two-stroke motor. Nature 388:720–722

86. Taguchi H (2005) Chaperonin GroEL meets the substrate protein as a “load” of the rings. J Biochem 137:534–539
87. Sparrer H, Buchner J (1997) How GroES regulates binding of nonnative protein to GroEL. J Biol Chem 272(22):14080–14086
88. Niwa T, Ying B, Saito K, Jin W, Takada S, Ueda T, Taguchi H (2009) Bimodal protein solubility distribution revealed by an aggregation analysis of the entire ensemble of *Escherichia coli* proteins. Proc Natl Acad Sci USA 106(11):4201–4206
89. Ellis JR, Hemmingsen SM (1989) Molecular chaperones: proteins essential for the biogenesis of some macromolecular structures. Trends Biochem Sci 14:339–342
90. Thirumalai D, Klimov DK, Dima RI (2003) Emerging ideas on the molecular basis of protein and peptide aggregation. Curr Opin Struct Biol 13:146–159
91. Ranson NA, Dunster NJ, Burston SG, Clarke AR (1995) Chaperonins can catalyze the reversal of early aggregation steps when a protein misfolds. J Mol Biol 250(5):581–586
92. Schmidt M, Buchner J, Todd MJ, Lorimer GH, Viitanen PV (1994) On the role of GroES in the chaperonin-assisted folding reaction—3 case studies. J Biol Chem 269(14):10304–10311
93. van der Vies SM, Viitanen PV, Gatenby AA, Lorimer GH, Jaenicke R (1992) Conformational states of ribulose bisphosphate carboxylase and their interaction with chaperonin 60. Biochemistry 31(14):3635–3644
94. Dobson CM (2003) Protein folding and misfolding. Nature 426:884–890
95. Dobson CM (1999) Protein misfolding, evolution and disease. Trends Biochem Sci 24:329–332
96. Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, Nishikawa K, Ueda T (2001) Cell-free translation reconstituted with purified components. Nat Biotechnol 19:751–755
97. Shimizu Y, Kanamori T, Ueda T (2005) Protein synthesis by pure translation systems. Methods 36(3):299–304
98. Murzin AG, Brenner SE, Hubbard T, Chothia C (1995) Scop: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol 247(4):536–540
99. Horwich AL, Apetri AC, Fenton WA (2009) The GroEL/GroES *cis* cavity as a passive anti-aggregation device. FEBS Lett 583(16):2654–2662
100. Ellis RJ (2003) Protein folding: importance of the Anfinsen cage. Curr Biol 13(22):R881–R883
101. Agard DA (1993) To fold or not to fold. Science 260(5116):1903–1904
102. Weissman JS, Kashi Y, Fenton WA, Horwich AL (1994) GroEL-mediated protein-folding proceeds by multiple rounds of binding and release of nonnative forms. Cell 78:693–702
103. Betancourt MR, Thirumalai D (1999) Exploring the kinetic requirements for enhancement of protein folding rates in the GroEL cavity. J Mol Biol 287(3):627–644
104. Sfatos CD, Gutin AM, Abkevich VI, Shakhnovich EI (1996) Simulations of chaperonin-assisted folding. Biochemistry 35(1):334–339
105. Siabil H (2000) Molecular chaperones: containers and surfaces for folding, stabilising or unfolding proteins. Curr Opin Struct Biol 10(2):251–258
106. Thirumalai D, Lorimer GH (2001) Chaperonin-mediated protein folding. Annu Rev Biophys Biomol Struct 30:249–269
107. Stan G, Thirumalai D, Lorimer GH, Brooks BR (2003) Annealing function of GroEL: structural and bioinformatic analysis. Biophys Chem 100:453–467
108. Wang JD, Weissman JS (1999) Thinking outside the box: new insights into the mechanism of GroEL-mediated protein folding. Nat Struct Biol 6(7):597–600
109. Gulukota K, Wolynes PG (1994) Statistical mechanics of kinetic proofreading in protein folding. Proc Natl Acad Sci USA 91(20):9292–9296
110. Chan HS, Dill KA (1996) A simple model of chaperonin-mediated protein folding. Proteins 24(3):345–351
111. Gorse D (2001) Global minimization of an off-lattice potential energy function using a chaperone-based refolding method. Biopolymers 59:411–426
112. Hubbard TJP, Sander C (1991) The role of heat-shock and chaperone proteins in protein folding. Protein Eng 4(7):711–717
113. Tehver R, Thirumalai D (2008) Kinetic model for the coupling between allosteric transitions in GroEL and substrate protein folding and aggregation. J Mol Biol 377:1279–1295
114. Orland H, Thirumalai D (1997) A kinetic model for chaperonin assisted folding of proteins. J Phys I (France) 7:553–560
115. Altschuler GM, Willison KR (2008) Development of free-energy-based models for chaperonin containing TCP-1 mediated folding of actin. J R Soc Interface 5(29):1391–1408
116. Baumketner A, Jewett A, 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(3):701–713
117. Hayer-Hartl M, Minton AP (2006) A simple semiempirical model for the effect of molecular confinement upon the rate of protein folding. Biochemistry 45:13356–13360
118. Takagi F, Koga N, Takada S (2003) How protein thermodynamics and folding mechanisms are altered by the chaperonin cage: molecular simulations. Proc Natl Acad Sci USA 100(20):11367–11372
119. Klimov DK, Newfield D, Thirumalai D (2002) Simulations of β-hairpin folding confined to spherical pores using distributed computing. Proc Natl Acad Sci USA 99(12):8019–8024
120. Mittala J, Best RB (2008) Thermodynamics and kinetics of protein folding under confinement. Proc Natl Acad Sci USA 105(51):20233–20238
121. Friedel M, Sheeler DJ, Shea J-E (2003) Effects of confinement and crowding on the thermodynamics and kinetics of folding of an off-lattice protein model. J Chem Phys 118:8106–8113
122. Cheung MS, Klimov D, Thirumalai D (2005) Molecular crowding enhances native state stability and refolding rates of globular proteins. Proc Natl Acad Sci USA 102(13):4753–4758
123. Xu W-X, Wang J, Wang W (2005) Folding behavior of chaperonin-mediated substrate protein. Proteins 61:777–794
124. Cheung MS, Thirumalai D (2006) Nanopore–protein interactions dramatically alter stability and yield of the native state in restricted spaces. J Mol Biol 357(2):632–634
125. Jewett AI, Baumketner A, 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(36):13192–13197
126. England J, Lucent D, Pande V (2008) Rattling the cage: computational models of chaperonin-mediated protein folding. Curr Opin Struct Biol 18:163–169
127. England J, Pande V (2008) Potential for modulation of the hydrophobic effect inside chaperonins. Biophys J 95:3391–3399
128. England J, Lucent D, Pande V (2008) A role for confined water in chaperonin function. J Am Chem Assoc 130(36):11838–11839
129. Lucent D, Vishal V, Pande VS (2007) Protein folding under confinement: a role for solvent. Proc Natl Acad Sci USA 104(25):10430–10434
130. Vaitheeswarana S, Thirumalai D (2008) Interactions between amino acid side chains in cylindrical hydrophobic nanopores with applications to peptide stability. Proc Natl Acad Sci USA 105(46):17636–17641
131. Xu W, Mu Y (2008) Polar confinement modulates solvation behavior of methane molecules. J Chem Phys 128:234506
132. Diamant S, Peres BA, Bukau B, Goloubinoff PA (2000) Size-dependent disaggregation of stable protein aggregates by the DnaK chaperone machinery. J Biol Chem 275(28):21107–21113

133. Glover JR, Lindquist S (1998) HSP104, HSP70, and HSP40: a novel chaperone system that rescues previously aggregated proteins. Cell 94(1):73–82
134. Ben-Zvi A, De Los Rios P, Dietler G, Goloubinoff P (2004) Active solubilization and refolding of stable protein aggregates by cooperative unfolding action of individual Hsp70 chaperones. J Biol Chem 279(36):37298–37303
135. Goloubinoff P, Mogk A, Zvi APB, Tomoyasu T, Bukau B (1999) Sequential mechanism of solubilization and refolding of stable protein aggregates by a bichaperone network. Proc Natl Acad Sci USA 96(24):13732–13737
136. Mogk A, Tomoyasu T, Goloubinoff P, Rüdiger S, Röder D, Langen H, Bukau B (1999) Identification of thermolabile *Escherichia coli* proteins: prevention and reversion of aggregation by DnaK and ClpB. EMBO J 18(24):6934–6949
137. Krzewska J, Langer T, Liberek K (2001) Mitochondrial Hsp78, a member of the Clp/Hsp100 family in *Saccharomyces cerevisiae*, cooperates with Hsp70 in protein refolding. FEBS Lett 489(1):92–96
138. Zettlmeissl G, Rudolph R, Jaenicke R (1979) Reconstitution of lactic dehydrogenase. noncovalent aggregation vs. reactivation. 1. Physical properties and kinetics of aggregation. Biochemistry 18(25):5567–5571
139. Chiti F, Taddei N, Baroni F, Capanni C, Stefani M, Ramponi G, Dobson CM (2002) Kinetic partitioning of protein folding and aggregation. Nat Struct Biol 9(2):137–143
140. Finke JM, Roy M, Zimm BH, Jennings PA (2000) Aggregation events occur prior to stable intermediate formation during refolding of interluekin 1β. Biochemistry 39:575–583
141. Farr GW, Scharl EC, Schumacher RJ, Sondek S, Horwich AL (1997) Chaperonin-mediated folding in the eukaryotic cytosol proceeds through rounds of release of native and nonnative forms. Cell 89(6):927–937
142. Thulasiraman V, Yang C-F, Frydman J (1999) In vivo newly translated polypeptides are sequestered in a protected folding environment. EMBO J 18(1):85–95
143. Horst R, Fenton WA, Englander SW, Wührich K, Horwich AL (2007) Folding trajectories of human dihydrofolate reductase inside the GroEL–GroES chaperonin cavity and free in solution. Proc Natl Acad Sci USA 104(52):20788–20792
144. Coyle JE, Texter FL, Ashcroft AE, Masselos D, Robinson CV, Radford SE (1999) GroEL accelerates the refolding of hen lysozyme without changing its folding mechanism. Nat Struct Biol 6(7):683–690
145. Chen J, Walter S, Horwich AL, Smith DL (2001) Folding of malate dehydrogenase inside the GroEL–GroES cavity. Nat Struct Biol 8(8):721–728
146. Goldberg MS, Zhang J, Sondek S, Matthews CR, Fox RO, Horwich AL (1997) Native-like structure of a protein-folding intermediate bound to the chaperonin GroEL. Proc Natl Acad Sci USA 94(4):1080–1085
147. Lilie H, Buchner J (1995) Interaction of GroEL with a highly structured folding intermediate: iterative binding cycles do not involve unfolding. Proc Natl Acad Sci USA 92(18):8100–8104
148. Jewett AI, Shea J-E (2008) Do chaperonins boost protein yields by accelerating folding or preventing aggregation? Biophys J 94(8):2987–2993
149. Martin J, Langer T, Boteva R, Schramel A, Horwich AL, Hartl FU (1991) Chaperonin-mediated protein folding at the surface of GroEL through a molten globule like intermediate. Nature 352:36–42
150. Viitanen PV, Donaldson GK, Lorimer GH, Lubben TH, Gatenby AA (1991) Complex interactions between the chaperonin 60 molecular chaperone and dihydrofolate reductase. Biochemistry 30:9716–9723
151. Zahn R, Perrett S, Fersht AR (1996) Conformational states bound by the molecular chaperones GroEL and SecB: a hidden unfolding (annealing) activity. J Mol Biol 261(1):43–61
152. Zahn R, Perrett S, Stenberg G (1996) Catalysis of amide proton exchange by the molecular chaperones GroEL and SecB. Science 271(5249):642–645
153. Jaenicke R, Seckler R (1997) Protein misassembly in vitro. Adv Protein Chem 50:1–59
154. Wetzel R (1996) For protein misassembly, it's the "I" decade. Cell 86:699–702
155. Fink AL (1995) Compact intermediate states in protein folding. Annu Rev Biophys Biomol Struct 24:495–522
156. Gorovits BM, McGee WA, Horowitz PM (1998) Rhodanese folding is controlled by the partitioning of its folding intermediates. Biochimica et Biophysica Acta 1382(1):120–128
157. Tandon S, Horowitz PM (1989) Reversible folding of rhodanese: presence of intermediate(s) at equilibrium. J Biol Chem 264(17):9859–9866
158. Jennings PA, Wright PE (1993) Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin. Science 262(5135):892–896
159. Jackson SE, Fersht AR (1991) Folding of chymotrypsin inhibitor-2.1. Evidence for a two-state transition. Biochemistry 43:10428–10435
160. Brockwell DJ, Radford SE (2007) Intermediates: ubiquitous species on folding energy landscapes? Curr Opin Struct Biol 17:30–37
161. Jahn TR, Radford SE (1991) The Yin and Yang of protein folding. FEBS J 43:10428–10435
162. Englander SW, Mayne L, Krishna MMG (2007) Protein folding and misfolding: mechanism and principles. Q Rev Biophys 40:1–41
163. Lindberg MO, Oliveberg M (2007) Malleability of protein folding pathways: a simple reason for complex behaviour. Curr Opin Struct Biol 17(1):21–29
164. Netzer WJ, Hartl FU (1997) Recombination of protein domains facilitated by co-translational folding in eukaryotes. Nature 388:343–349
165. Wright CF, Teichmann SA, Clarke J, Dobson CM (2005) The importance of sequence diversity in the aggregation and evolution of proteins. Nature 438:878–881
166. Bryngelson JD, Wolynes PG (1987) Spin glasses and the statistical mechanics of protein folding. Proc Natl Acad Sci USA 84(21):7524–7528
167. Bryngelson JD, Wolynes PG (1989) Intermediates and barrier crossing in a random energy model (with applications to protein folding). J Phys Chem 93(19):6902–6915
168. Goldstein RA, Luthey-Schulten ZA, Wolynes PG (1989) Protein tertiary structure recognition using optimized Hamiltonians with local interactions. Proc Natl Acad Sci USA 89(19):9029–9033
169. Shea J, Onuchic JN, Brooks CL III (2000) Energetic frustration and the nature of the transition state in protein folding. J Chem Phys 113(17):7663–7671
170. Nymeyer H, García AE, Onuchic JN (1998) Folding funnels and frustration in off-lattice minimalist protein landscapes. Proc Natl Acad Sci USA 95(11):5921–5928
171. Shea J-E, Nochomovitz YD, Guo Z, Brooks CL III (1998) Exploring the space of protein folding Hamiltonians: the balance of forces in a minimalist β-barrel model. J Chem Phys 109(7):2895–2903
172. Onuchic JN, Wolynes PG (2004) Theory of protein folding. Curr Opin Struct Biol 14(1):70–75
173. Mirny LA, Abkevich V, Shakhnovich EI (1996) Universality and diversity of the protein folding scenarios: a comprehensive analysis with the aid of a lattice model. Fold Des 1(2):103–116

174. Shea J, Onuchic JN, Brooks CL III (1999) Exploring the origins of topological frustration: design of a minimally frustrated model of fragment b of protein a. Proc Natl Acad Sci USA 96(22):12512–12517
175. Nymeyer H, Socci ND, Onuchic JN (2000) Landscape approaches for determining the ensemble of folding transition states: success and failure hinge on the degree of frustration. Proc Natl Acad Sci USA 97(2):634–639
176. Onuchic JN, Nymeyer H, García AE, Chahine J, Socci ND (2000) The energy landscape theory of protein folding: insights into folding mechanisms and scenarios. Adv Protein Chem 53:87–152
177. Clementi C, Nymeyer H, Onuchic JN (2000) Topological and energetic factors: what determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? An investigation for small globular proteins. J Mol Biol 298(5):937–953
178. Micheletti C, Banavar JR, Maritan A, Seno F (1999) Protein structures and optimal folding from a geometrical variational principle. Phys Rev Lett 82(16):3372–3375
179. Stan G, Brooks BR, Thirumalai D (2005) Probing the "annealing" mechanism of GroEL minichaperone using molecular dynamics simulations. J Mol Biol 350(4):817–829
180. Martin J, Hartl FU (1997) The effect of macromolecular crowding on chaperonein-mediated protein folding. Proc Natl Acad Sci USA 94(4):1107–1112
181. Burston SG, Weissman JS, Farr GW, Fenton WA, Horwich AL (1996) Release of both native and non-native proteins from a *cis*-only GroEL ternary complex. Nature 383:96–99
182. Elcock AH (2003) Atomic-level observation of macromolecular crowding effects: escape of a protein from the GroEL cage. Proc Natl Acad Sci USA 100(5):2340–2344
183. Zahn R, Buckle AM, Perrett S, Johnson CM, Corrales FJ, Golbik R, Fersht AR (1996) Chaperone activity and structure of monomeric polypeptide binding domains of GroEL. Proc Natl Acad Sci USA 93(26):15024–15029
184. Corrales FJ, Fersht AR (1995) The folding of GroEL-bound barnase as a model for chaperonin-mediated protein folding. Proc Natl Acad Sci USA 92(12):5326–5330
185. Makino Y, Taguchi H, Yoshida M (1993) Truncated GroEL monomer has the ability to promote folding of rhodanese without GroES and ATP. FEBS Lett 336(2):363–367
186. Zhou H, Dill K (2001) Stabilization of proteins in confined spaces. Biochemistry 40(38):11289–11293
187. Eggers D, Valentine J (2001) Crowding and hydration effects on protein conformation: a study with sol–gel encapsulated proteins. J Mol Biol 314(4):911–922
188. Zhou H–X (2008) Protein folding in confined and crowded environments. Arch Biochem Biophys 469(1):76–82
189. Rathore N, Knotts T, de Pablo J (2006) Confinement effects on the thermodynamics of protein folding: Monte Carlo simulations. Biophys J 90(5):1767–1773
190. Zhang S-Q, Cheung MS (2007) Manipulating biopolymer dynamics by anisotropic nanoconfinement. Nano Lett 7(11): 3438–3442
191. Cheung MS, Thirumalai D (2007) Effects of crowding and confinement on the structures of the transition state ensemble in proteins. J Phys Chem B 111(28):8250–8257
192. Chun SY, Strobel S, Bassford PJ, Randall LL (1993) Folding of maltose-binding protein. Evidence for the identity of the rate-determining step in vivo and in vitro. J Biol Chem 268:20855–20862
193. Wang JD, Michelitsch MD, Weissman JS (1998) GroEL–GroES-mediated protein folding requires an intact central cavity. Proc Natl Acad Sci USA 95:12163–12168
194. Silow M, Oliveberg M (1997) Transient aggregates in protein folding are easily mistaken for folding intermediates. Proc Natl Acad Sci USA 94:6084–6089
195. Nawrocki J, Chu R-A, Pannell LK, Bai Y (1999) Intermolecular aggregations are responsible for the slow kinetics observed in the folding of cytocrome *c* at neutral ph. J Mol Biol 293:991–995
196. Farr GW, Fenton WA, Horwich AL (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(13):5342–5347
197. Jewett AI, Shea J-E (2006) Folding on the chaperone: yield enhancement through loose binding. J Mol Biol 363(5):945–957
198. Horst R, Bertelsen EB, Fiaux J, Wider G, Horwich AL, Wüthrich K (2005) Direct NMR observation of a substrate protein bound to the chaperonin GroEL. Proc Natl Acad Sci USA 102(36):12748–12753
199. Boshoff A, Stephens LL, Blatch GL (2008) The *Agrobacterium tumefaciens* DnaK: ATPase cycle, oligomeric state and chaperone properties. Int J Biochem Cell Biol 40(4):804–812
200. Erbse A, Dougan DA, Bukau B (2003) A folding machine for many but a master of none. Nat Struct Biol 10(2):84–86
201. Wang JD, Herman C, Tipton KA, Gross CA, Weissman JS (2002) Directed evolution of substrate-optimized GroEL/S chaperonins. Cell 111:1027–1039