Quarterly Reviews of Biophysics **36**, 2 (2003), pp. 229–256. © 2003 Cambridge University Press
DOI: 10.1017/S0033583503003883 Printed in the United Kingdom

Chaperonin-mediated protein folding: fate of substrate polypeptide

Wayne A. Fenton and Arthur L. Horwich[1,*]

Department of Genetics and [1] Howard Hughes Medical Institute, Yale University School of Medicine, 295 Congress Avenue, New Haven, CT 06510, USA

Abstract. Chaperonins are megadalton ring assemblies that mediate essential ATP-dependent assistance of protein folding to the native state in a variety of cellular compartments, including the mitochondrial matrix, the eukaryotic cytosol, and the bacterial cytoplasm. Structural studies of the bacterial chaperonin, GroEL, both alone and in complex with its co-chaperonin, GroES, have resolved the states of chaperonin that bind and fold non-native polypeptides. Functional studies have resolved the action of ATP binding and hydrolysis in driving the GroEL–GroES machine through its folding-active and binding-active states, respectively. Yet the exact fate of substrate polypeptide during these steps is only poorly understood. For example, while binding involves multivalent interactions between hydrophobic side-chains facing the central cavity of GroEL and exposed hydrophobic surfaces of the non-native protein, the structure of any polypeptide substrate while bound to GroEL remains unknown. It is also unclear whether binding to an open GroEL ring is accompanied by structural changes in the non-native substrate, in particular whether there is an unfolding action. As a polypeptide-bound ring becomes associated with GroES, do the large rigid-body movements of the GroEL apical domains serve as another source of a potential unfolding action? Regarding the encapsulated folding-active state, how does the central cavity itself influence the folding trajectory of a substrate? Finally, how do GroEL and GroES serve, as recently recognized, to assist the folding of substrates too large to be encapsulated inside the machine? Here, such questions are addressed with the findings available to date, and means of further resolving the states of chaperonin-associated polypeptide are discussed.

* Author to whom correspondence should be addressed. E-mail: horwich@csb.yale.edu

1. Chaperonin action – an overview

The final step in the transfer of information from coding DNA to active protein involves the folding of a polypeptide chain into its characteristic three-dimensional structure. The early work of Anfinsen and co-workers (Anfinsen, 1973, and references therein) made clear that the primary structure of a polypeptide chain contains all of the information necessary for reaching the native state, which usually lies at the energetic minimum. Studies of two-state folding proteins, which are generally small (fewer than 150 amino acids) and transit directly between unfolded and native states, have begun to elaborate the mechanisms by which specific primary sequences dictate folding into particular native structures. For example, φ-value mutational analysis (Matouschek *et al.* 1989; Oliveberg, 2001) and, more recently, computational and experimental methods for studying very fast folding proteins (Mayor *et al.* 2000, 2003), have led to the articulation of models for the transition states and folding pathways of a number of two-state proteins. Yet most cellular proteins are larger than 150 amino acids, have a complex domain structure, and do not follow simple two-state folding behavior. While the same folding mechanisms directed by primary structure appear to be operative, the energy landscapes for these proteins are more complicated. Instead of a smooth down-slope to the native state, the topology is rugged (Dill & Chan, 1997). Many, if not the majority, of molecules in an ensemble of non-native conformers of such proteins can land in false minima in this folding landscape, kinetic traps that effectively prevent the molecules from reaching the native state on a physiological timescale. Experimentally, whereas two-state proteins diluted from denaturant rapidly and efficiently fold to native form, these latter proteins will fail to reach the native state *in vitro* and, in many cases, produce an aggregate of insoluble protein. Yet *in vivo*, these same proteins are able to rapidly and efficiently reach the native state. This is the result of actions by a collective of cellular machines, known as molecular chaperones, that provide kinetic assistance to the folding process, keeping proteins out of kinetic traps and effectively serving to 'smooth' the energy landscape.

Studies of the past 15 years indicate that molecular chaperones function as a class by specifically binding non-native states of proteins through exposed hydrophobic surfaces that eventually become buried to the interior in the native state, effectively forestalling aggregation (e.g. Bukau & Horwich, 1998; Hartl & Hayer-Hartl, 2002). Bound proteins are then released, in many cases via the action of ATP, for another attempt at folding. In some cases, a net change of protein conformation attends the step of binding and/or release, whereas in others there is no observable change, but the overall action is rather one of protecting the protein from misfolding and aggregation until it can successfully proceed to a next step of biogenesis.

The particular families of chaperones exhibit different topologies of binding non-native forms (Bukau & Horwich, 1998). For example, the Hsp70 class binds non-native forms whose exposed hydrophobic side-chains are present in extended structures, with binding by the chaperone achieved through a hydrophobic arch-like structure (Mayer *et al.* 2002). Such binding plays a key role in preventing nascent polypeptides emerging from eukaryotic ribosomes from prematurely

folding, for example. It is also vital in assisting protein translocation into mitochondria and the ER by binding incoming segments of polypeptide at the *trans* side of membranes, providing a motor or ratchet action that enables forward translocation (Ryan & Pfanner, 2001). By contrast, in the case of the small heat-shock protein (sHsp) family, binding of non-native forms occurs at the surface of an oligomeric assembly of sHsp's, for example during heat shock (Van Montfort *et al.* 2002). Following the return of temperature to normal, the polypeptides are released. Differing further still from these two classes of chaperone, the Hsp60 'chaperonin' class of chaperone binds collapsed protein conformations in the central cavity of one of its characteristic rings, with exposed hydrophobic side-chains of the non-native species binding to the hydrophobic lining of the cavity. The multivalent nature of this binding could potentially contribute to reversing misfolded states.

However, it is the nature of release that particularly sets the chaperonin class apart from other chaperones. Instead of releasing the bound protein back into the bulk solution, as occurs with Hsp70 or with the sHsp's, for example, the chaperonin releases its substrate protein into an encapsulated central cavity where it has a chance to fold in an isolated environment. Studies of this mechanism of folding have particularly focused on the double-ring bacterial chaperonin, GroEL (for other recent reviews, see Sigler *et al.* 1998; Grantcharova *et al.* 2001; Thirumalai & Lorimer, 2001; Saibil *et al.* 2002; Horwich & Fenton, 2003).

GroEL is a tetradecamer of identical 57 kDa subunits, arranged in two heptameric rings stacked back-to-back with a central cavity in each ring (Fig. 1*a*). Each subunit comprises three domains: equatorial (dark shades), which binds ADP or ATP and provides the inter-ring contacts that hold the rings together; apical (medium shades), which contains the hydrophobic surface at the entrance to the cavity that binds polypeptide and interacts with the co-chaperonin, GroES; and intermediate (light shades), which links the other domains together and has flexible hinges at each end that permit the large-scale domain movements that occur during a folding cycle (Fig. 1*a*, compare left and right images, and see below). The co-chaperonin GroES is a single-ring heptamer of identical 10 kDa subunits that binds to GroEL in the presence of ADP or ATP, capping one cavity to form an asymmetric complex (Fig. 1*a*, right image). GroES is dynamic during the reaction cycle, binding to and releasing from one GroEL ring and then the other (Fig. 2).

The folding cycle is outlined in Fig. 2, starting with binding of a non-native polypeptide to an open GroEL ring (Fig. 2*a*). The binding of ATP and GroES to the same ring as polypeptide switches the GroEL ring from an open, binding-active state to a closed, folding-active state in which polypeptide has been ejected into a cavity that is now encapsulated by GroES (Fig. 2*b*) and whose surface has been converted by rigid-body movements from hydrophobic to hydrophilic (Fig. 1*b*). Folding to native form appears to be favored by at least three features of this so-called *cis* GroEL–GroES complex. One is the property of solitary confinement. Chaperonins generally bind only one substrate molecule at a time. Thus, a polypeptide folding in this encapsulated environment simply has no other species with which to form a multimolecular aggregate. The second property is the hydrophilicity of the cavity lining (Fig. 1*b*). Recent structural and biochemical studies indicate that the switch to completely hydrophilic character upon ATP/GroES binding is associated with ejection of substrate off the cavity wall. But the presence of a hydrophilic cavity wall during the folding process probably also serves to assist productive folding, insofar as it probably favors burial of hydrophobic surfaces and exposure of hydrophilic ones, properties of the native state, in the folding substrate protein. Finally, the situation of confinement within a cavity may limit the range of conformations that a non-native protein can explore.

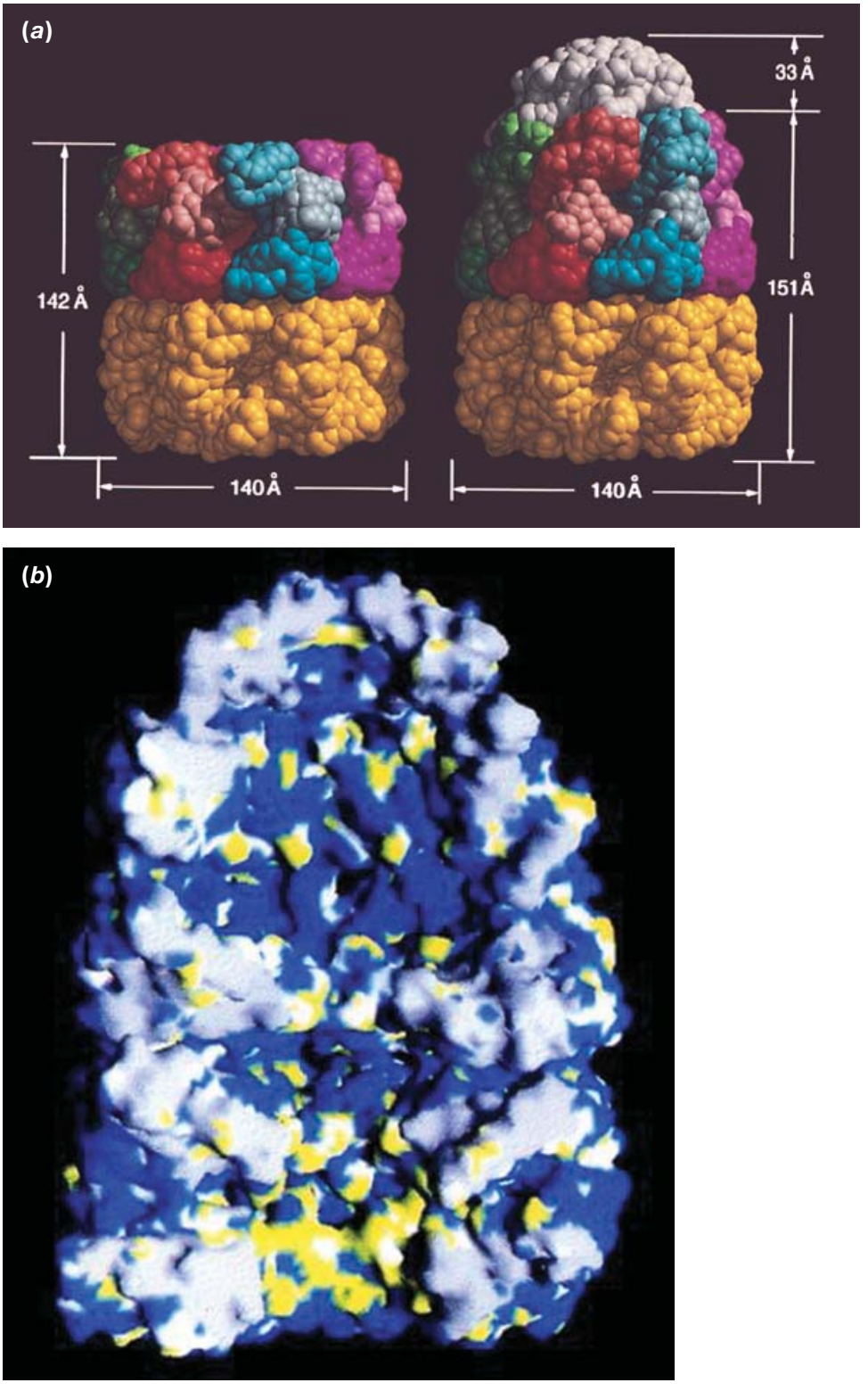

Fig. 1. For legend see opposite page.

https://doi.org/10.1017/S0033583503003883 Published online by Cambridge University Press

For example, fully extended conformations would be excluded from forming inside a cavity whose diameter measures maximally 60 Å. Thus these latter properties may serve to provide a novel and smoother shape to the energy landscape for a polypeptide folding inside a GroEL–GroES *cis* complex, as compared with the 'rugged' surface it would face in the bulk solution in the absence of chaperonin.

Given the assistance provided by the GroEL system through both the binding and the *cis*-folding steps, it is understandable that the outcome of folding of its normal substrate proteins in the absence of chaperonin is not a productive one. For example, in yeast with conditional deficiency of Hsp60, the mitochondrial homologue of GroEL, a large number of imported mitochondrial proteins fail to reach the native state and form insoluble aggregates (Cheng *et al.* 1989; Dubaquié *et al.* 1998). Similarly *in vitro*, when such proteins are diluted from denaturant in the absence of GroEL, they undergo wholesale aggregation. By contrast, in the presence of GroEL, they become quantitatively bound to it; and in the added presence of ATP and GroES, they are quantitatively folded to native form. Thus the action of kinetic assistance provided by chaperonins is essential, and genetic deletion experiments *in vivo* show that these components are required for cell viability under all temperature conditions (Cheng *et al.* 1989; Fayet *et al.* 1989).

While work of the past few years has been able to resolve some of the structural states of the GroEL–GroES machine and has elucidated how ATP drives the machine through its cycle (recently reviewed in Horwich & Fenton, 2003; see also Fig. 2 and legend), less understanding has been gained regarding the conformational state(s) of the substrate protein during the steps of binding to an open GroEL ring, during *cis* folding inside a GroEL–GroES complex, or following discharge from such a complex into the bulk solution. Moreover, although there are approximate understandings about the states occupied by the protein, the exact actions of the chaperonin on substrates are not well-understood. Indeed, substrate polypeptide remains the least resolvable component of the chaperonin system. This is because, of course, the non-native substrate lacks the structural order and symmetries of the GroEL–GroES machine itself, which have made it accessible to EM and X-ray studies. The most challenging aspect of studying substrates is that they almost certainly occupy ensembles of conformational states at all stages of folding up until

Fig. 1. Structures of GroEL and GroEL–GroES. (*a*) Space-filling models of GroEL (left) and GroEL–GroES–ADP (right), based on crystal structures (Braig *et al.* 1994; Xu *et al.* 1997, respectively). Individual subunits in the upper ring of both are colored in shades of green, red, blue, and pink. In each case, the darkest shade is the equatorial domain, the lightest the intermediate domain, and the medium shade the apical domain. The bottom ring is uniformly colored gold; GroES (right panel) is colored white. Note the changes in position and orientation of the apical domains, in particular, upon GroES and ADP binding (60° elevation and 90° clockwise twist; compare left and right panels), as well as in the dimensions of the upper ring. The GroES mobile loops, which directly contact the apical domains of GroEL, are inside the cavity in the GroEL–GroES–ADP complex and thus are not visible. (From Sigler *et al.* 1998, with permission.) (*b*) Hydrophobic and hydrophilic character of the surface inside the cavities of the GroEL–GroES–ADP_7 complex in a cutaway view based on the crystallographic model (Xu *et al.* 1997). The open *trans* (bottom) ring is similar in conformation to an unliganded GroEL ring and shows the hydrophobic character (yellow) of the walls of the central cavity, particularly in the apical domain near the opening of the cavity. The rigid-body movements of the apical domains upon GroES and nucleotide binding dislocate the hydrophobic side-chains from the cavity to form contacts with GroES and a new interface between apical domains, replacing the lining of the *cis* (upper) cavity with hydrophilic side-chains (blue). This is associated with the release of polypeptide from the apical binding surface into the cavity and the commencement of folding. This hydrophilic character of the folding-active cavity is likely to favor productive folding to the native state (see text). (From Xu *et al.* 1997, with permission.)

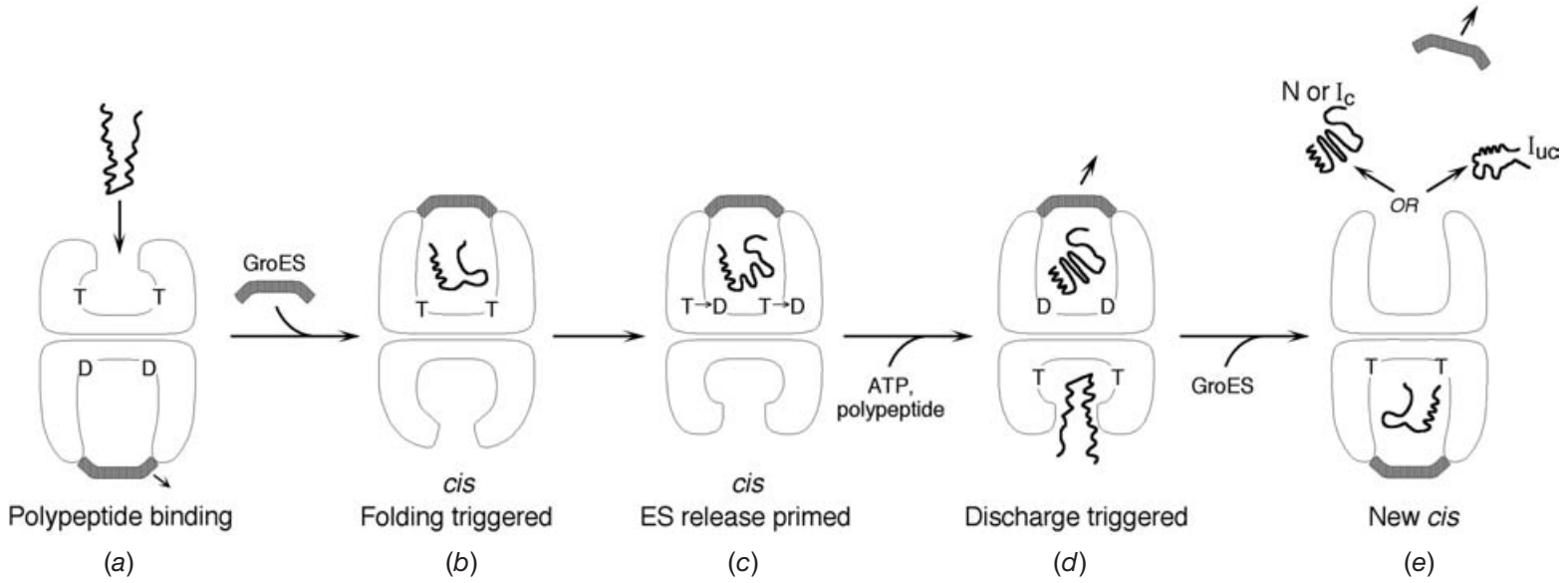

Fig. 2. The GroEL–GroES reaction pathway. (*a*) A non-native polypeptide (solid wavy line) binds to an asymmetric GroEL–GroES complex, the most likely acceptor state *in vivo*. It associates with the cavity walls of the open ring, specifically with exposed hydrophobic side-chains of the apical domains. (*b*) GroES (shaded image) binds to the same ring as the non-native polypeptide, in the presence of ATP, producing rigid-body movements in the apical domains that dislocate the hydrophobic surface away from the cavity and release the polypeptide into the sequestered space, triggering folding. Folding proceeds in the very stable *cis* ATP complex for 8–10 s. (at 20 °C), comprising the longest phase of the chaperonin cycle. This phase is ended by *cis* ATP hydrolysis. (*c*) Folding continues seamlessly in the encapsulated space of the *cis* ADP complex for its usually short lifetime (<1 s). Importantly, the affinity of GroEL for GroES is weakened in this complex relative to the ATP complex, thus priming the *cis* complex for release of its ligands. (*d*) ATP and another non-native polypeptide molecule bind to the *trans* ring, allosterically triggering dissociation of the ligands from the opposite ring, discharging the substrate into solution. (*e*) The released substrate either has reached the native state (N) or one committed to it (I_c) or is still in a non-native state (I_{uc}) that can bind to another GroEL molecule for a further attempt at folding. *In vivo*, I_{uc} can also partition to another chaperone or to the proteolysis machinery for further processing. With ordered binding of GroES to the opposite ring, a new, folding active *cis* complex is formed [see (*e*)]. ATP binding is positively cooperative within a ring, but negatively cooperative between rings, so that only one ring at a time is occupied by ATP. Thus, the two rings of GroEL alternate as folding-active, using the binding of seven ATPs in one ring to initiate folding there, while simultaneously discharging the products of folding from the other ring. While ATP/GroES *binding* is sufficient to trigger folding, ATP *hydrolysis* provides directionality to the cycle. Two opportunities exist in the cycle for further unfolding or rearrangement of a non-native polypeptide, one during binding to an open ring [see (*a*), (*d*)] and the other during the large movements of the apical domains attending ATP/GroES binding [see (*b*), (*e*)]. T denotes ATP, D denotes ADP, the solid wavy lines represent substrate polypeptide, and the shaded image represents GroES. (Redrawn from Grantcharova *et al.* 2001, copyright © 2001, with permission from Elsevier.)

achieving the native state. Multiple experimental approaches are likely be required for any full elucidation of the fate of substrate during the reaction sequence, and, as one considers our current understanding, it seems clear that newer, more penetrating methods that can deal with the ensemble behavior would be desirable. We review here both what is currently known about conformational states of substrates during the chaperonin reaction and what the prospects are for a better understanding of the action of the machine on them.

2. Polypeptide binding – an essential action

Early experiments made it clear that binding of non-native protein to GroEL could forestall misfolding and aggregation. For example, non-native subunits of ribulose bisphosphate carboxylase-oxygenase (Rubisco) from *Rhodospirillum rubrum*, diluted from denaturant into aqueous buffer devoid of chaperonin, quantitatively aggregated, whereas in the presence of GroEL a stoichiometric complex was formed that was fully productive of native Rubisco upon addition of ATP/GroES (Goloubinoff *et al.* 1989). Notably, addition of GroEL at later times could not reverse aggregation that had already occurred, but could forestall further aggregation. Thus, binding by GroEL directly competes with a pathway of misfolding and aggregation. Consistent with this, in intact cells conditionally deficient of GroEL or Hsp60, newly translated and newly translocated proteins, respectively, aggregated (Cheng *et al.* 1989; Horwich *et al.* 1993). More specifically, plasmids encoding GroELs with mutations in the polypeptide binding surface could not rescue growth of GroEL-deficient *E. coli*, despite formation of double-ring assemblies (Fenton *et al.* 1994). Thus the action of polypeptide binding is an essential one, within the more general observation that chaperonin components are essential to cells under all conditions.

Notably, the processes of translation and translocation themselves appear unaffected in the GroEL and Hsp60 conditional mutants, respectively, suggesting that chaperonin does not interact with nascent chains in these contexts, but rather with complete polypeptide chains that have been released from the ribosome or translocon, respectively. Correspondingly, physical association of GroEL/Hsp60 with translating/translocating polypeptides has not been observed (Gaitanaris *et al.* 1994). Thus, the ensemble of intermediates produced upon dilution of a substrate protein from denaturant probably resembles that exposed to GroEL or Hsp60 *in vivo* following release from a ribosome or from a mitochondrial import site, respectively.

The importance not only of initial binding of non-native polypeptide but also of rebinding the non-native forms that are released into solution along with native ones after any given cycle of binding and folding at GroEL is further emphasized by recent experiments. In one *in vitro* study, the rebinding of non-native substrates by GroEL was blocked by steric obstruction of the central cavity with streptavidin, which bound to biotinyl groups placed at the inlet to the cavity (Brinker *et al.* 2001). This obstruction was associated with an immediate halt to productive folding, indicating that released non-native polypeptides required further rounds of binding and folding at GroEL to achieve the native state. In a second study, the ability of GroEL to bind non-native forms was found to be sufficient to support productive folding even in the absence of *cis* complex formation (Farr *et al.* 2003). This was revealed by production of *trans*-only GroEL–GroES complexes that could bind and release polypeptide only from an open ring opposite one to which GroES was closely covalently tethered. Non-native polypeptide could be recursively bound by such complexes, which progress through a normal nucleotide cycle, but polypeptide could only be released by such complexes directly into the bulk solution, as opposed to into a *cis* cavity. Surprisingly, this was sufficient to allow refolding of Rubisco and another stringent,

GroEL–GroES-dependent substrate, mitochondrial malate dehydrogenase (MDH), in experiments *in vitro*. The rate of refolding of these substrates, however, was slower than that of a wild-type *cis* reaction, reflecting the apparent benefits to *cis* folding. Nevertheless, as with the *cis* reaction, the ability of the substrate molecules to proceed through multiple rounds of binding to an open ring, followed by a trial at folding probably after release into solution, is essential, as demonstrated by the ability of exogenously added GroEL trap molecules, which can bind but not release non-native substrate proteins, to arrest further refolding. Thus, recursive polypeptide binding is absolutely required for both *cis* and *trans* mechanisms.

3. Recognition of non-native polypeptide – role of hydrophobicity

Early experiments suggested a role for hydrophobicity in recognition by GroEL. For example, aggregation of a substrate protein, rhodanese, could be prevented by GroEL as well as by a detergent, lauryl maltoside, suggesting that the action of GroEL involved stabilizing hydrophobic surfaces of non-native intermediates (Mendoza *et al.* 1991). Structural and mutational studies supported such a conclusion. The crystal structure of unliganded GroEL revealed that the central cavity of an open GroEL ring is lined by a hydrophobic surface (Braig *et al.* 1994). In particular, each of the seven surrounding apical domains contains, at its cavity aspect, a tier of three structures, two α-helices and an underlying extended segment, on which are situated a set of cavity-exposed hydrophobic amino-acid side-chains. A ring of hydrophobic surface of $\sim$30 Å height is thus formed (see Fig. 1*b*, lower ring). Mutational change of any single amino acid within the hydrophobic collective to hydrophilic character (i.e. seven such changes in a ring) abolished polypeptide binding *in vitro*, and, as expected, such mutants could not rescue GroEL-deficient cells (Fenton *et al.* 1994).

More recently, using GroEL rings produced as single continuous polypeptides, it has been possible to test effects of various numbers and arrangements of such polypeptide binding mutations (Farr *et al.* 2000). Three consecutive wild-type apical domain binding surfaces were required in order to support a full extent of binding of Rubisco or MDH *in vitro*. Similarly, in rescue of GroEL-deficient cells *in vivo*, this arrangement was required for cell viability. This suggests that a continuous hydrophobic binding surface extending $\sim 150^{\circ}$ around the inside of a ring is needed for efficient binding of non-native forms.

The indications from these functional experiments that polypeptide is bound simultaneously by multiple apical domains was directly evaluated by disulfide cross-linking studies (Farr *et al.* 2000). Disulfide bond formation was measured between bound non-native Rubisco and a fully functional variant GroEL containing a single cysteine at position 261 in the cavity face of each of its apical domains. Here, multiple oxidative cross-links, as a measure of multivalency of GroEL apical domain binding to a radiolabeled substrate protein, could be readily detected by autoradiography following non-reducing SDS–PAGE, with the size of the non-reduced species indicating the number of GroEL subunits bound (i.e. 50·5 kDa radiolabeled Rubisco + multiples of crosslinked 57 kDa GroEL). This revealed that generally three or four apical domains could react with bound Rubisco, which has five available cysteines fairly evenly spaced through its primary structure. 'Ladders' of species were observed at the expected positions of the adducts, probably reflecting different orders and arrangements of crosslinked GroELs and hence reflecting different bound conformations of the substrate.

The role of hydrophobicity in polypeptide binding has been supported not only from the standpoint of GroEL but also from that of the molecules that become bound to it. For example,

https://doi.org/10.1017/S0033583503003883 Published online by Cambridge University Press

one study of various substituted versions of the small protein, CI2, measuring GroEL binding by degree of retardation of the rate of its folding, observed less binding when hydrophobic residues were replaced with alanine or glycine (Itzhaki *et al.* 1995). Another study measured binding of a stably unfolded version of subtilisin by isothermal titration calorimetry, observing a negative heat capacity change, indicative of hydrophobic interaction (Lin *et al.* 1995). A third study examined peptides recovered after partial proteolytic treatment of a binary complex between the substrate protein rhodanese and GroEL, with the idea that regions of rhodanese immediately associated with the GroEL cavity wall would be resistant to such treatment and would remain stably associated (Hlodan *et al.* 1995). Two peptides of 7 and 11 kDa were recovered, which correspond in native rhodanese to amphiphilic α-helical regions, hydrophobic on one aspect of the α-helix and hydrophilic on the opposite one. These two helices contact each other through their hydrophobic aspects in the native state. Binding to GroEL was proposed to occur through these sites, stabilizing them until they could correctly interact after release from the cavity wall.

More recent studies of peptides have provided additional support for hydrophobic interaction between substrate protein and GroEL. Preuss *et al.* (1999) examined a series of 14-residue peptides with α-helical character in solution. Those peptides with amphiphilic character, i.e. with a continuous hydrophobic surface at one aspect, were bound with the greatest affinity. Likewise, when two 13-residue α-helical peptides with identical composition were compared, the one with amphiphilic character bound more strongly than the one with interspersed hydrophobic and hydrophilic residues (Wang *et al.* 1999). Thus, a maximum exposure of hydrophobic surface in non-native peptide favors its binding to GroEL. This appears likely to be the major determinant to substrate binding, with hydrophobic surface able to be recognized not only in the context of α-helix, as just described, but also in the context of extended conformation, as reviewed below.

4. Crystallographic analyses of peptide binding

Two groups have observed peptide–GroEL interactions at high resolution by X-ray crystallography. One study by Fersht and co-workers (Buckle *et al.* 1997) examined a lattice of isolated monomeric apical domains and observed that an N-terminal tag segment of one of the monomers was associated in an extended fashion with what would be the cavity-facing aspect of another monomer in the asymmetric unit. The extended segment lay in a groove between the two cavity-facing α-helices (H and I) (Fig. 3), and the hydrophobic amino-acid side-chains of the tag formed contacts with many of the hydrophobic residues that had been identified as involved with polypeptide binding by the earlier mutational study of GroEL. In a second study by Chen & Sigler (1999), an affinity-panning procedure binding 12-mer expressing phage to immobilized isolated apical domains was carried out to select peptides with the greatest affinity for the isolated GroEL apical domain. This identified one peptide that bound with an approximately micromolar affinity, with the sequence SWMTTPWGFLHP. This sequence is remarkable for its hydrophobicity and for its content of tryptophan. When this peptide was co-crystallized with isolated GroEL apical domains, a segment (amino acids 7–12) was found in an extended conformation lying in the inter-helical groove, again with numerous hydrophobic interactions observed between the hydrophobic side-chains in the peptide and the previously identified apical side-chains. The remaining N-terminal portion of the peptide formed a hairpin with the distal segment and did not contact the apical domains (Fig. 3). An identical topology was observed when this same peptide was co-crystallized with intact GroEL, with a peptide occupying each of the 14 GroEL apical domains.

https://doi.org/10.1017/S0033583503003883 Published online by Cambridge University Press

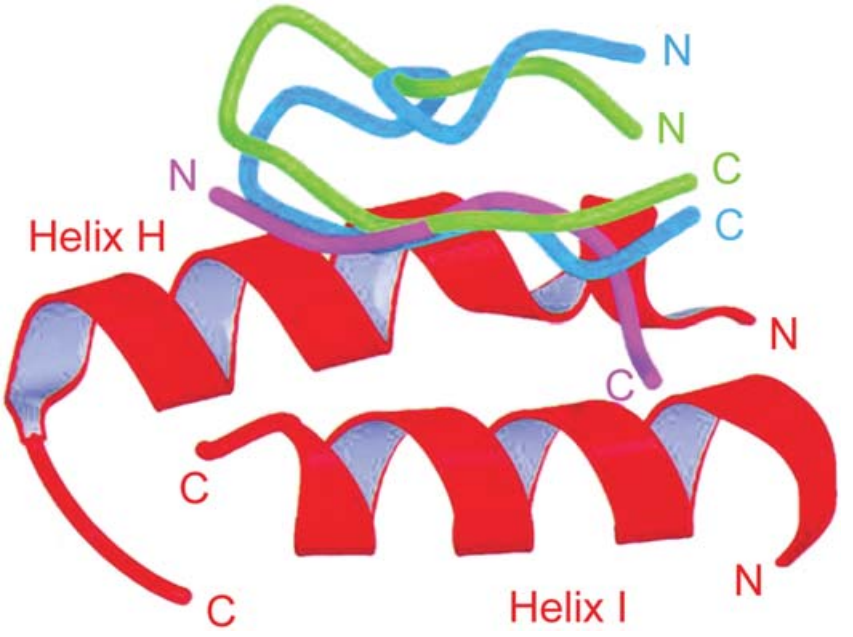

Fig. 3. Comparison of the structures of three peptides bound to helices H and I of the GroEL apical domain, viewed from the central cavity just below the level of the apical domain of a subunit. The Cα coordinates of the apical domains of three structures were superimposed (Chen & Sigler, 1999), and the backbones of the peptides are shown: blue, the GroES mobile loop (amino acids 18–30) from the GroEL–GroES–ADP_7 structure (Xu *et al.* 1997); magenta, the N-terminal extension of the apical domain reported in Buckle *et al.* (1997); green, the strong binding peptide (SBP) from Chen & Sigler (1999). Note that the bound segments of all three peptides lie in nearly the same position between the apical domain α-helices. (Redrawn from Chen & Sigler, 1999, copyright © 1999, with permission from Elsevier.)

Remarkably, this topology of an extended segment lying within the groove between apical α-helices H and I resembled that of the mobile loops of GroES bound 1:1 to the apical domains of GroEL in the GroEL–GroES–ADP_7 crystal structure (Xu *et al.* 1997). In this latter context, the apical domains of the GroES-bound GroEL ring occupy an elevated and twisted position relative to that of an open polypeptide-accepting ring, but the local topology is the same, with the mobile loop forming a β-hairpin structure that makes similar hydrophobic contacts with helices H and I through its distal segment, which contains a hydrophobic sequence, IVL (Fig. 3). Thus, this raises the question of whether the energetically favored position for binding short peptides in the H–I groove of the apical domain is simply a mimic of the normal binding of the GroES mobile loop (see Shewmaker *et al.* 2001) or whether this reflects a favored mode for binding intact polypeptide. In this regard, Ashcroft *et al.* (2002) examined the binding of two 20-mer peptides, one derived from the mobile loop of GroES and the other from helix D of rhodanese, to both isolated apical domains and wild-type GroEL. Based on changes in the fluorescence of dansylated peptides on binding, they concluded that these two peptides did not compete for the same binding site, suggesting that their sites of binding on the apical domain might be unique. Further structural characterization of these interactions would be helpful in identifying the binding sites. Notably, in all of the studies so far, the underlying extended segment of the apical domain, residues 199–209, containing additional hydrophobic residues essential for binding such substrates as Rubisco and MDH, has not been involved in the binding of small peptides.

https://doi.org/10.1017/S0033583503003883 Published online by Cambridge University Press

Moreover, the two peptide crystal structures examine only a mode of binding an extended segment, while a number of studies, as described above, indicate that other peptides either adopt or maintain α-helical structure upon association with GroEL. For example, both the early studies of Gierasch and co-workers (Landry & Gierasch, 1991; Landry *et al.* 1992) and more recent ones (Wang *et al.* 1999) observed that a 13-residue N-terminal peptide from rhodanese adopted an α-helical structure upon association with GroEL, as evaluated by NMR studies examining transferred NOEs. A further study by the Fersht group (Kobayashi *et al.* 1999) revisited this interaction using isolated apical domains whose resonances had been assigned, identifying chemical shift changes in helices H and I associated with peptide binding. But the structural topology of such α-helix binding remains to be seen, e.g. whether it is a coiled-coil type of interaction or otherwise. Finally, in the co-crystal structures, there is a failure of the bound peptide to cross beyond the immediate confines of a single apical domain to a neighboring one. Thus the structural course taken during multivalent binding of a natural substrate protein remains entirely unknown.

More generally, the visualization of GroEL-bound substrate proteins by X-ray crystallography poses formidable problems. To list some of the obstacles:

(1) Polypeptides appear to occupy an ensemble of states while bound to GroEL.
(2) The conformational states occupied by bound polypeptide appear to be 'loose', e.g. with low protection factors in hydrogen–deuterium (H–D) exchange studies (see below).
(3) Rotational symmetry, in which a polypeptide of any given conformation can be bound in seven identical ways, effectively reduces its appearance in any single rotational orientation within a crystal lattice to $\sim$15%.
(4) Negative cooperativity for polypeptide substrate binding between GroEL rings also affects occupancy; that is, when a substrate polypeptide binds to one GroEL ring, the opposite ring is effectively inhibited from binding a second molecule.

Substrate binding has recently been associated with small, allosterically directed adjustments of the apical domains of the unoccupied ring, visualized by cryo-EM (Falke *et al.* 2001). Thus, depending on whether such binding affects the outside aspect of a ring and thus the crystal packing of the binary complexes, the chaperonin complex might pack only one way, with all substrate-bound rings of GroEL occupying the same relative position in the lattice. Alternatively, there could be a random arrangement of occupied *versus* unoccupied rings, effectively reducing the presence of polypeptide in any given position by half.

Thus, there are large problems in seeing natural substrates crystallographically. These translate to specific practical questions, such as: Would it be possible to trap a single conformational state out of an ensemble of states of a bound non-native substrate? Could one fix polypeptide in the same rotational orientation in every GroEL ring in a crystal lattice, e.g. by using covalent rings bearing only a particular binding-proficient surface inside and a rotationally regularizing adduct on the outside? Clearly, the degrees of freedom of both substrate conformation and chaperonin will need to be restricted for any such structural analysis to be possible.

5. Topology and secondary and tertiary structure of bound substrate polypeptide – fluorescence, hydrogen exchange and NMR studies

In general, it appears that GroEL-bound proteins occupy collapsed, loosely packed conformations, with varying degrees of native secondary structure. In some experiments there has also

https://doi.org/10.1017/S0033583503003883 Published online by Cambridge University Press

been evidence for a native-like global topology (see below), but there is a need to confirm this by independent analyses. The overriding determinant of how much higher order structure is present and what particular structures are populated is likely to be the degree of exposure of hydrophobic surfaces (see below).

A polypeptide substrate forms its association with GroEL via multivalent hydrophobic contacts inside the central cavity of a ring, but there is a limited volume available inside a ring for encompassing polypeptide, totaling $\sim 60\,000$ Å^3. This estimate of available volume takes into consideration the collective of seven flexible C-terminal tails of the GroEL subunits, which collectively occupy the equatorial zone of the cavity. Assuming a partial specific volume of protein of 1·7 Å^3/Da, and an expansion of volume of a collapsed, loosely packed polypeptide substrate of $\sim 30\%$ beyond the volume of the native state, a protein of only $\sim$20–30 kDa can be fully accommodated inside the cavity of an open ring. Thus for many proteins, at least a portion of the non-native conformer must lie outside the central cavity in contact with the bulk solution. This has been observed in a neutron scattering experiment examining rhodanese (33 kDa), which bound in part within the GroEL ring, but also substantially outside of it, with a topology suggested to resemble that of a champagne cork (Thiyagarajan *et al.* 1996). The question thus arises as to whether this topology is a characteristic one, with the same sets of amino acids of rhodanese always inside *versus* outside the cavity, or whether it reflects an ensemble of topologies in which rhodanese can be positioned in the cavity in a variety of orientations. A number of different experiments may be able to address this issue.

The extent of secondary and tertiary structure formation in bound substrate polypeptides has been examined by a variety of experimental approaches. Early experiments recognized that bound substrate proteins were very susceptible to digestion by added protease, supporting the presence of a loosely packed, accessible structure (e.g. Martin *et al.* 1991). Similarly, tryptophan fluorescence measurements of bound rhodanese and dihydrofolate reductase, availing of the lack of tryptophan in GroEL, indicated that substrate tryptophan residues occupied environments intermediate in polarity between native and fully unfolded states.

Concerning the amount of native structure present in bound forms, the results have varied considerably. When thermally unfolded β-lactamase became bound to GroEL, it produced a native-like tryptophan fluorescence spectrum, and, in H–D exchange/MS analysis, the bound protein exhibited a degree of protection indicating that considerable secondary structure was present (Gervasoni *et al.* 1996). By contrast, when a thermophilic lactate dehydrogenase was diluted from denaturant and incubated to allow varying amounts of structure to form before GroEL addition, only early conformational states could be bound, whereas a later molten-globule-like one could not (Badcoe *et al.* 1991; Staniforth *et al.* 1994). In another variation (Goldberg *et al.* 1997), fluorescence studies of human DHFR showed that it could be bound as an early burst phase intermediate (i.e. within the first 10 ms of dilution from denaturant), a time when initial collapse of the structure was occurring, but that it could also be rapidly and efficiently bound when GroEL was added at later times (e.g. after 30 s or 1 min) when final secondary and tertiary contacts of the native state were forming. Thus a wide range of DHFR conformers is recognizable by GroEL.

Attempts have been made to characterize the GroEL-bound conformations of a number of proteins using H–D exchange of backbone amide protons of a substrate in substrate–GroEL binary complexes. In all of these experiments, with the exception of the thermally unfolded β-lactamase mentioned above, only a low degree of exchange protection has been observed, reflecting that the secondary structure in the bound protein must be unstable. For example, with

scrambled 3-disulfide intermediates of α-lactalbumin, where exchanged binary complexes were directly examined by electrospray–MS, which dissociated the complex, protection factors in α-lactalbumin of the order of 2–10 were observed (Robinson *et al.* 1994). With MDH, a similarly low amount of protection was observed, here able to be mapped to a particular region of the protein (Chen *et al.* 2001). In particular, after binary complexes were exchanged, the MDH was separated under quenching conditions, and peptic peptides, prepared at low pH, were examined by HPLC–MS. Two strands and an α-helix within a Rossmann fold of the N-terminal NAD-binding domain exhibited protection factors of $\sim$100. In a further exchange study with human DHFR, NMR analysis of the exchanged protein was carried out to obtain site-specific information about exchange (Goldberg *et al.* 1997). Binary complex was exchanged for varying times, and quenching of further exchange was accomplished by ATP/GroES-mediated conversion to native form. The extrapolated rates of exchange for individual amide protons revealed that, while bound to GroEL, DHFR exhibited the most significant degree of exchange protection in what corresponds in the native state to its central parallel β-sheet, with protection factors ranging from 5 to 50. This suggests that a native-like topology, albeit an unstable one, may be sampled while the protein is bound to GroEL.

To better resolve non-native states while bound to GroEL, it would be desirable to be able to examine them directly. Recently, NMR techniques for examining large molecules have been developed, particularly through the work of Wüthrich and colleagues (Pervushin *et al.* 1997; Riek *et al.* 2002). These techniques effectively slow the fast relaxation of magnetization that occurs for large, slowly tumbling molecules in solution, preventing the broadening of resonance lines beyond detection. Such techniques, in particular TROSY (transverse relaxation-optimized spectroscopy) and CRIPT (cross relaxation-induced polarization transfer) (see Riek *et al.* 2000, for review), have recently been applied together to the GroEL system, enabling spectra to be collected both on $^{15}N^{2}H$-labeled GroEL alone and on complexes in which $^{15}N^{2}H$-GroES was bound to GroEL (Fiaux *et al.* 2002). The latter was particularly informative because it had already been possible, using TROSY triple-resonance analyses, to assign GroES as a stand-alone complex. The use of CRIPT/TROSY allowed the observation of resonances from nearly all of the residues of GroES while complexed with GroEL (in ADP). No changes in chemical shift were detected for most of the amide protons of GroES, with the exception of amino acids 17–32, which showed large chemical shift changes upon association of GroES with GroEL. These amino acids map to the mobile loop region, the region of GroES that has been shown through earlier NMR (Landry *et al.* 1993) and crystallographic work (Hunt *et al.* 1996; Xu *et al.* 1997) to convert from a flexible, mobile region in stand-alone GroES to one that is stably physically associated with the GroEL apical domains, as a β-hairpin, upon GroEL–GroES complex formation.

The ability to observe GroEL–GroES complexes by NMR raised the question of whether isotopically labeled polypeptide substrates might also be detectable while complexed with GroEL or its smaller, single-ring version. In the case of polypeptide, of course, instead of seven signals from each of seven identical GroES subunits bound in rotationally symmetric fashion to a GroEL ring, there is only one polypeptide molecule bound per GroEL ring, potentially reducing the signal beyond detection. Initial tests with binary complexes of $^{15}N^{2}H$-human DHFR and GroEL appear promising, however, producing detectable resonances and suggesting the feasibility of directly examining non-native proteins while present at GroEL (E. Bertelsen, G. Wider, R. Horst, A. Horwich & K. Wüthrich, unpublished observations). Of course, such studies will be examining the ensemble of bound forms, however heterogeneous this may be, and may already

https://doi.org/10.1017/S0033583503003883 Published online by Cambridge University Press

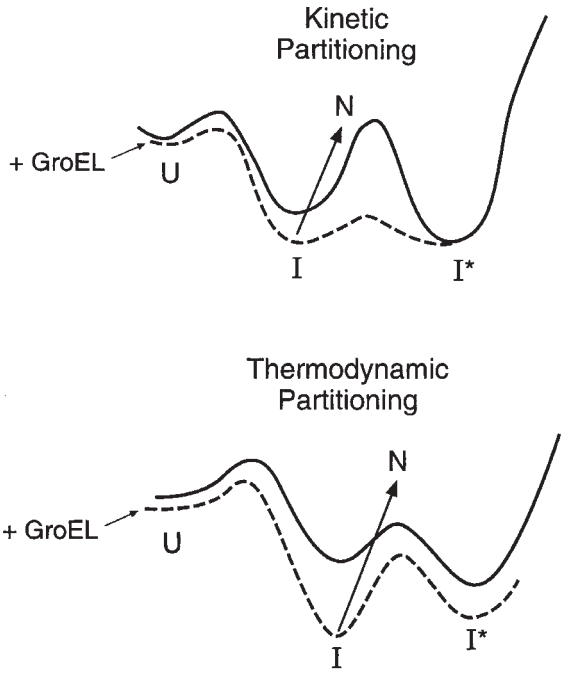

Fig. 4. Kinetic *versus* thermodynamic partitioning. Two reaction coordinate diagrams are shown for a model folding reaction with kinetic partitioning (top) and thermodynamic partitioning (bottom), the solid lines denote the absence of GroEL, the dashed lines its presence. In both cases, U denotes the unfolded state of the substrate, I an intermediate folding state, I* a misfolded state, and N the native protein. The angled arrow from I to N is meant to indicate folding along another coordinate in three-dimensional space in the presence of GroEL. In the kinetic partitioning model, the primary effect of GroEL is to decrease the barrier between I* and I, increasing the rate of interconversion of these states and making it possible for a substrate to escape from a folding trap. In the thermodynamic partitioning model, interaction with GroEL increases the stability of the I state and biases the I/I* equilibrium in favor of the productive state. (From Fenton & Horwich, 1997, with permission.)

provide some idea of whether there is a predominant species in this population. It could also inform as to whether there are particular regions of DHFR that associate with the cavity wall, and thus exhibit the same dynamics as GroEL, *versus* regions that are moving freely inside the central cavity.

6. Binding by GroEL associated with a putative unfolding action

The observation that polypeptide binding by GroEL could prevent non-native forms from irreversible misfolding and aggregation suggested early on that there could be an action of unfolding or unscrambling associated with binding (e.g. Jackson, 1993). Two scenarios could be envisaged (Fig. 4). In one, referred to as thermodynamic partitioning (Zahn *et al.* 1994; Zahn & Plückthun, 1994), GroEL would preferentially bind certain less-folded conformations from the ensemble of non-native states. Assuming a facile equilibrium between conformations in the ensemble, mass action would lead all of the non-native molecules to adopt the forms most likely to bind, resulting in a net unfolding of the ensemble. In the second, called kinetic partitioning, binding itself would bring about an unfolding by reducing the kinetic barrier between the folded

https://doi.org/10.1017/S0033583503003883 Published online by Cambridge University Press

and unfolded states (Itzhaki *et al.* 1995; Ranson *et al.* 1995). The potential for a non-native polypeptide to achieve multivalent interaction with the GroEL apical domains, a situation shown to be critical for stringent substrates both *in vitro* and *in vivo* (Farr *et al.* 2000), might drive such unfolding. These possibilities are not mutually exclusive, of course, and both could be operative for any given substrate polypeptide.

The occurrence of an unfolding action has been examined with several polypeptides. For example, both human DHFR in the absence of ligands (Viitanen *et al.* 1991) and pre-β-lactamase (Laminet *et al.* 1990; Zahn *et al.* 1994) lose activity in the presence of chaperonin, binding tightly to GroEL. The addition of GroES and ATP refolds these proteins to their native, active states. In contrast, barnase, a small RNAse, does not bind stably to GroEL when unfolded, although it does interact transiently during its spontaneous refolding (Gray & Fersht, 1993; Corrales & Fersht, 1995). In this case, incubation of barnase with a sub-stoichiometric amount of GroEL increased the H–D exchange rate of even the core amide protons of its folding intermediate, suggesting a transient global unfolding catalyzed by chaperonin (Zahn *et al.* 1996b). In another study, rhodanese was translated *in vitro* from a construct lacking a stop codon (Reid & Flynn, 1996). The stalled chain was not recognized by GroEL, even though unfolded mature rhodanese is a GroEL substrate. Instead, the chain folded at the ribosome to produce a protease-protected 17 kDa N-terminal domain. Upon puromycin release of the stalled chain, however, the rhodanese was recognized and bound by GroEL, and the N-terminal domain became protease-sensitive, indicating that it was now unfolded.

Kinetic studies of MDH refolding by GroEL also examined the possibility of GroEL-driven unfolding. In one study, MDH, unfolded over a range of temperatures, occupied a metastable, kinetically trapped state (Peralta *et al.* 1994). This could be rescued and the native state produced by addition of GroEL, GroES, and ATP. In a second study, rates and yields of spontaneous and chaperonin-assisted folding of MDH diluted from denaturant were used to produce a kinetic model in which GroEL binding redirected reversible low-order aggregates of MDH away from irreversible higher-order aggregation and towards productive monomer folding (Ranson *et al.* 1995). An argument for an action of reversing aggregation, as opposed to simply blocking it, was based on the ability of GroEL (with GroES/ATP) to be productive even at sub-stoichiometric concentration. The suggestion was made that GroEL might exert a disrupting action, breaking protein–protein contacts of the low-order aggregates, but no direct evidence has been presented for such an action upon binding.

Which of the mechanisms discussed above is actually operating in each of these cases has been difficult to establish. For human DHFR in the absence of ligands and pre-β-lactamase, the native states are known to be unstable relative to inactive, less-folded ones, so it seems possible that GroEL simply binds the latter forms, particularly if they expose significantly more hydrophobic surface. This would shift the equilibrium towards unfolding without any active participation by the chaperonin, a thermodynamic partitioning effect (Zahn *et al.* 1994). Similarly in the case of MDH, given the topology of binding in the GroEL cavity, which would favor binding of monomeric species, it seems possible that there is a facile equilibrium between the low-order oligomers and binding-accessible monomers and that the presence of GroEL shifts that equilibrium towards the monomer (and eventually, in the presence of GroES and ATP, the refolded state) (Ranson *et al.* 1995).

In another example, thermodynamic partitioning has been invoked to account for the effects of GroEL on the refolding of thioredoxin, a small protein whose folding is slowed in the presence of GroEL (Bhutani & Udgaonkar, 2001). In this system, the three major forms that

comprise the unfolded ensemble bind differentially to GroEL. Because only one of these can fold productively in the presence of GroEL, all of the unfolded molecules eventually refold by passing through the productive state and do so with kinetics that are identical to those shown by this state in the absence of GroEL. These data are consistent with a mass-action effect in directing all of the thioredoxin molecules in the unfolded ensemble through the one bound state that leads to native enzyme. Finally, in a direct demonstration of a thermodynamic partitioning model, Walter *et al.* (1996) studied a chemically modified RNAse T1 that could not fold completely to its native state, but was trapped as a metastable state in equilibrium with a folding intermediate, which could bind to GroEL. Because the presence of GroEL did not affect the rates of conversion between these states, but only the overall equilibrium mixture of states, it was clear that no direct action of unfolding had been exerted by the chaperonin.

On the other hand, a direct unfolding action seems the more likely explanation in certain cases, for example in the case of rhodanese released from the ribosome (see above; Reid & Flynn, 1996). Moreover, for barnase, further experiments indicated that GroEL accelerated by >1000-fold the unfolding of native barnase to a known intermediate state, a kinetic partitioning effect. In contrast, the presence of GroEL inhibited both the further unfolding of the intermediate to a fully unfolded state and the refolding of this most unfolded state, the latter effect accounting for the increased global exchange seen in the presence of GroEL (Zahn *et al.* 1996a). Even for MDH, if low-order aggregates (i.e. misfolded dimers or trimers) can bind in the GroEL cavity, perhaps via one component subunit or domain (perhaps resembling the binding of a large protein such as aconitase; see below), then an unfolding action on one of the constituent monomers might release it from the others and drive the aggregation reaction back towards non-native monomers that can refold productively with the further action of GroEL, GroES and ATP.

The possible importance of an unfolding behavior was further emphasized when it became clear that non-native forms were being released along with native ones after each trial of folding at GroEL and that rebinding was permitting a further attempt at productive folding (Todd *et al.* 1994; Weissman *et al.* 1994; Smith & Fisher, 1995). Thus, there was the implication that rebinding would return a polypeptide to its original GroEL-bound state, perhaps via an action of unfolding. Indeed, when the partial protease protection patterns of bound non-native rhodanese were compared between an initial binary complex and one recovered at later times during a folding reaction, they were the same, consistent with similar structures in the bound non-native states (Weissman *et al.* 1994). In a higher resolution study, Groß *et al.* (1996) used MS to examine H–D exchange protection in human DHFR bound to GroEL, both in the initial binary complex obtained by diluting denatured DHFR into a GroEL-containing solution and in a complex recovered after several minutes of ATP-driven refolding. Remarkably, the kinetics and extent of hydrogen exchange were the same in the two samples, suggesting that both had a similar core of residues somewhat protected from exchange and implying that both had the same basic structure. In contrast, it was also reported in this study that differences were observed between such samples when two other measures of structure were considered, namely iodide quenching of tryptophan fluorescence (reflecting residue solvent exposure) and ANS binding (reflecting amount of exposed hydrophobic surface). Because the changes in these parameters were relatively small, the authors concluded that, although the overall structure of bound DHFR was the same throughout the course of a chaperonin-mediated refolding reaction, local or minor structural rearrangements might be occurring. To date, however, there is no evidence that

progressively more-folded states of a substrate polypeptide bind to GroEL during the course of a folding reaction. Rather, folding appears to be an all-or-none process.

7. A potential action of substrate unfolding driven by ATP/GroES binding

Although it is probable that unfolding can accompany the association of a polypeptide with an open GroEL ring, it also seems possible that an unfolding action could be exerted on the polypeptide during the course of a folding cycle as a consequence of the major domain movements that take place in the chaperonin itself. Because non-native polypeptide binds multivalently to the apical domains of GroEL and because it is these domains that undergo the largest movements with respect to each other during the chaperonin cycle, associated with ATP/GroES binding, it has been postulated that such movements could stretch and effectively unfold a bound polypeptide in the brief period before it is released into the GroES-enclosed *cis* cavity. Even if initial binding to a GroEL ring had failed to rescue a misfolded polypeptide, such movements during the chaperonin cycle could serve to pull a polypeptide out of a structural kinetic trap and might result in all molecules in an ensemble beginning to refold at the same (high) point on the energy landscape at each round of folding. This mechanism has been termed 'iterative annealing' (Corrales & Fersht, 1996; Todd *et al.* 1996). The latter group in particular emphasized that the energy of ATP and GroES binding would be directly employed for unfolding in a cycling reaction.

Two experiments have attempted to test this hypothesis and have arrived at somewhat different conclusions. Shtilerman *et al.* (1999) examined hydrogen–tritium exchange of tritiated Rubisco diluted from denaturant into protic buffer, under conditions in which Rubisco did not refold spontaneously. There were 10–12 tritiums that were highly protected from exchange (retained after 10 min or more at pH 8 and 22 °C) in both the absence and presence of a stoichiometric amount of GroEL. When GroES and ATP or AMPPNP were added to the GroEL-bound sample, all but 2–3 of these remaining tritiums exchanged in less than 5 s, i.e. in less than the half-time for a chaperonin cycle under these conditions (13 s). When carried out with a sub-stoichiometric (1:20) amount of GroEL, exchange required GroES and ATP, but now did not occur with AMPPNP. In this case, the exchange kinetics were consistent with substrate unfolding upon interaction with GroEL and GroES/ATP, with released, exchanged, non-native Rubisco competing with the labeled non-native Rubisco in solution for rebinding to GroEL. Two conclusions were drawn from these data. First, the initial binding of the non-native Rubisco ensemble to the chaperonin did not further unfold whatever local structure (or structures) was represented by the slowly exchanging tritiums. Secondly, the addition of GroES and ATP (or AMPPNP) to the stable binary complex rapidly caused most of this local structure to unfold, allowing these residues to exchange. Thus, these data appear to support the proposed model, in which the movement of the GroEL apical domains upon GroES and ATP (or AMPPNP) binding causes a stretching or disruption of local structure (or structures), resulting in the exposure of its backbone amide protons to solvent and consequent rapid exchange.

Several points should be noted regarding these findings. First, ADP plus GroES did not lead to exchange. Thus, the large-scale elevations and rotations of the apical domains that accompany ADP and GroES binding, including significant movements of the substrate-binding domains relative to each other, were insufficient to activate exchange. This is consistent with other observations that ADP with GroES does not release substrate protein from the cavity walls

https://doi.org/10.1017/S0033583503003883 Published online by Cambridge University Press

(Weissman *et al.* 1996; G. W. Farr, unpublished observations). How the γ-phosphate of ATP triggers productive release is under structural study. Either the kinetics or extent of apical domain movement produced by ATP/GroES *versus* ADP/GroES could account for its effectiveness. Secondly, the timescale of the observed exchange of Rubisco tritiums is consistent with an earlier fluorescence study (Rye *et al.* 1997). Stopped-flow examination of changes in tryptophan anisotropy of Rubisco during its refolding showed a rapid initial decrease ($t_{\frac{1}{2}} < 1$ s), followed by a slower increase with a rate constant comparable to that of the recovery of Rubisco activity.

On the other hand, questions concerning the topology of the GroEL–Rubisco complex require further evaluation of the conclusion that the observed exchange resulted from unfolding. Because these experiments were carried out with wild-type, double-ring GroEL, binding of GroES should have occurred randomly with respect to the bound Rubisco, resulting in both *cis* and *trans* ternary complexes in roughly 1:1 proportion (Weissman *et al.* 1995). Only the *cis* complex would be expected to produce an unfolding action with ATP, because it is the one in which the substrate polypeptide is being affected by apical domains that are elevating and twisting. Thus, only half of the initially protected tritiums should be subject to exchange in a single turnover, not the 80% or so observed within 5 s when the complex had been formed with stoichiometric amounts of GroEL. Alternatively, it is possible that Rubisco bound in *trans* is also affected as a result of ATP/GroES-directed release of Rubisco from the *trans* ring (see below). Thus, this suggests caution in interpreting these exchange data as reflecting unfolding rather than, potentially, the act of polypeptide release.

A different result was obtained in an experiment by Chen *et al.* (2001), examining hydrogen exchange in non-native MDH bound to GroEL and after release in *cis* by GroES and ATP. Here, pulsed deuterium exchange of backbone amides was evaluated by MS of peptic peptides generated from the MDH. The single-ring version of GroEL, SR1, was used to avoid the above-mentioned difficulties in interpretation due to the mixture of *cis* and *trans* ternary complexes that would form with double-ring, wild-type chaperonin. SR1 has been shown to be as efficient in folding as wild-type GroEL, and, because there is only one chamber to be capped by GroES, it obligatorily forms only a *cis* complex (Weissman *et al.* 1996). When the MDH–SR1 binary complex was exposed briefly (1 s) to deuterated buffer, a substantial fraction of the amide protons were exchanged. Four peptides showed significant protection from exchange, however, with 90% exchange requiring 10–100 s in deuterated buffer, and these mapped to three α-helices and two β-strands in the second Rossmann fold of the NAD-binding domain. Their relative protection implied that some secondary structure persisted in this part of the non-native MDH when bound to GroEL. Notably, the protection observed here for non-native MDH appeared to be much less than that reported for the persistent tritiums in non-native Rubisco (Shtilerman *et al.* 1999). Importantly, a 1-s pulse of deuterium carried out 1 s after GroES and ATP were added to the binary complex showed little change in the pattern of protection of these peptides, even though by this time the MDH had been released into the SR1–GroES central cavity. Some additional exchange was noted in the weakly protected peptides scattered throughout the MDH structure, but the better-protected peptides lost none or at most one additional proton. Thus, in contrast to the Rubisco experiment, there was no evidence here for an unfolding of the residual structure present in bound MDH upon GroES and ATP binding. Rather, a slight, broadly distributed additional exchange was observed that was interpreted as the loss of local protection afforded by interaction between segments of the non-native protein and the apical binding sites. Consistent with these results, a stopped-flow fluorescence anisotropy experiment with tryptophan-bearing variants of MDH showed no early, rapid change of anisotropy upon

GroES and ATP binding. Only a single exponential increase with kinetics consistent with the MDH refolding rate was observed (W. A. Fenton, unpublished observations). This leaves open the possibility that the behavior of Rubisco, which exhibited an early drop in anisotropy, might be different from that of MDH. Pulsed H–D exchange studies of Rubisco during and after ATP/GroES binding would be desirable to address this question.

8. Folding in the *cis* cavity

Whatever the effects on polypeptide of binding to and release from the apical domains, the mechanism by which the *cis* cavity supports productive folding remains to be addressed. Is it an 'Anfinsen cage', a space where a single protein molecule can fold without interference or aggregation, essentially at infinite dilution (e.g. Ellis, 1994)? Or is it a special environment, where spatial confinement and the hydrophilic chamber walls serve to smooth the energy landscape for folding? Or is the chamber needed at all, if a non-native protein can be returned to a productive folding pathway directly by the process of binding and release? These questions have been difficult to answer, particularly for substrates that are dependent on the complete GroEL–GroES–ATP system for efficient folding. Size considerations alone suggest that the simplest Anfinsen cage model, a box of infinite dilution, cannot apply to many GroEL substrates. Globular substrate proteins of moderate size, such as rhodanese (33 kDa) and the MDH subunit (33 kDa) already are a close fit in the GroEL–GroES cavity, based on their native X-ray structures. Rubisco (50·5 kDa), closer to the upper limit of substrate size, is even tighter. If the non-native forms of these proteins released into the cavity are similar to molten-globule-folding intermediates, then their volumes are likely to be increased by 30% or more, leading to even less room, although increased flexibility might compensate somewhat. Further restriction might be imposed by the shells of hydration and other solvent molecules associated with both the cavity walls and the exterior of the protein. The combined effects of some of these factors has been observed for GFP, allowed to fold to its native, fluorescent state inside the stable SR1–GroES chamber (Weissman *et al.* 1996). When its fluorescence anisotropy decay was measured, it yielded a rotational correlation time of 54 ns, much slower than the 13·2 ns correlation time measured for GFP free in solution. This implied that the 27 kDa native GFP was significantly constrained in its rotation when sequestered in the SR1–GroES cavity.

Other experiments have also taken advantage of the stable SR1–GroES chamber. Because SR1 lacks a *trans* ring to relay the allosteric signals of *trans* ATP binding necessary for GroES release, SR1–GroES complexes are stable and do not cycle. Notably, only a single ATP turnover occurs in SR1 complexes, producing an SR1–GroES–ADP–polypeptide ternary complex that remains folding-active for an extended period (Weissman *et al.* 1996). Like GFP, rhodanese, a monomer, folds to its native, active state within SR1 when ATP and GroES are added, and it can be assayed while still encapsulated. The kinetics of the regain of rhodanese activity within this chamber are identical to those observed with wild-type GroEL and GroES during a cycling reaction when the substrate polypeptide is repeatedly released and rebound (see Fig. 2) (Weissman *et al.* 1996). Likewise, dimeric MDH and Rubisco regain native form in the SR1–GroES cavity similarly to the reaction with wild-type GroEL, as measured by activity assays after releasing the GroES with EDTA at 4 °C and allowing a brief time for dimerization. Other measures of refolding have produced a similar conclusion, namely that a single event of release into the *cis* cavity is sufficient for productive folding. For example, changes in fluorescence anisotropy of tryptophans in Rubisco are identical when ATP and GroES are added to binary

https://doi.org/10.1017/S0033583503003883 Published online by Cambridge University Press

complexes of Rubisco with either wild-type GroEL or SR1 (Rye *et al.* 1997). A H–D exchange study of MDH refolding in the SR1–GroES cavity led to corresponding results (Chen *et al.* 2001); that is, pulsed exchange at points throughout the course of the refolding reaction showed that monomeric species with near-native exchange characteristics accumulated in parallel with MDH activity, measured after GroES release and dimerization. Importantly, no differently protected species appeared to be present at times corresponding to one or two turnovers of the normal cycle, suggesting that folding confined to the chamber did not involve unique species not observed in a cycling reaction. Collectively, the foregoing data indicate that a single round of release into the *cis* cavity is sufficient to permit efficient folding of essentially all of the input substrate molecules, without any requirement for rebinding-associated unfolding to remove the substrate from kinetic traps.

The question remains, however, as to how the *cis* cavity supports refolding. Does it provide an action beyond sequestering a non-native protein from the aggregation that would occur in the bulk solution? A recent experiment supports a role beyond simple sequestration in supporting productive folding of Rubisco (Brinker *et al.* 2001). GroEL was biotinylated on a cysteine residue substituted into the inlet to the central cavity. This modified GroEL was fully active until streptavidin was added, leading to rapid blocking of the GroEL central cavity ($t_{\frac{1}{2}} \sim 200$ ms). When this addition was carried out during a Rubisco refolding reaction under stringent conditions (i.e. where a substrate protein cannot fold spontaneously and aggregates), released non-native protein was prevented from rebinding to the chaperonin and refolding immediately stopped. This indicated that Rubisco released from the *cis* chamber did not have the capacity in solution to partition to a form capable of reaching the native state. The situation with rhodanese was somewhat different, with a small increment (5–10%) of additional folding occurring in solution after streptavidin addition had blocked the GroEL cavity. This suggests that rhodanese can, to a limited extent, partition in the bulk solution to a state committed to refolding before aggregation supervenes.

Similar experiments were also carried out under so-called permissive conditions, i.e. where spontaneous folding can occur and aggregation is absent or sharply reduced. Under such conditions, rates of folding inside the *cis* cavity and outside in the bulk solution can be directly compared. For rhodanese, dilution alone to less than 20 nM in the refolding reaction, i.e. at 25-fold lower concentration than used in stringent conditions, was sufficient to permit spontaneous refolding without aggregation. Addition of GroEL alone halted refolding, consistent with its ability to rapidly bind the non-native protein. Adding GroES and ATP along with the GroEL allowed folding to proceed at the same rate as spontaneous folding, and blocking the biotinyl–GroEL cavity with streptavidin was without effect. These data indicate that, for rhodanese, the GroEL cavity may indeed simply be a box of infinite dilution, with no other special contribution from the *cis* cavity. For Rubisco, however, a different result was obtained. Here, to achieve permissive conditions, a combination of both increased dilution and reduced temperature was necessary to prevent aggregation and allow spontaneous folding to the native state (van der Vies *et al.* 1992; Schmidt *et al.* 1994). As with rhodanese, addition of GroEL alone prevented refolding. But in contrast, adding GroES and ATP resulted in a refolding rate that was about four-fold faster than spontaneous. Addition of streptavidin to this reaction immediately returned the reaction to the slower rate, indicating that folding within the GroES–GroEL cavity was responsible for the rate enhancement. Repeated binding and potential unfolding was not required for this acceleration, because the same increased rates were observed with SR1 complexes. Thus, for Rubisco, the GroEL chamber appears to reduce one or more kinetic barriers to

folding, smoothing the energy landscape or altering it completely by preventing the formation of trapped intermediates. Whether the difference between rhodanese and Rubisco in this regard reflects their size difference or the unique requirements of their respective folding pathways remains to be determined. Moreover, how the cavity accomplishes this modification of the folding landscape, whether by confinement or by rearrangement of solvent, remains an important question for further experimentation.

9. GroEL–GroES-mediated folding of larger substrate proteins by a *trans* mechanism

Early studies of polypeptide binding recognized that, when total *E. coli* proteins were diluted from denaturant into a mixture with GroEL, a large percentage of the species associated with the chaperonin, including many with a molecular size likely to exceed the capacity of the *cis* cavity (~60 kDa) formed when GroES binds (Viitanen *et al.* 1992). More recent studies isolated polypeptide–GroEL binary complexes from metabolically labeled cells using immunoprecipitation with anti-GroEL antibodies and likewise observed both smaller proteins (~20–60 kDa) and larger ones (>60 kDa) (Ewalt *et al.* 1997; Houry *et al.* 1999). The question remained, however, as to whether these larger species could be productively folded in the presence of ATP or ATP/GroES. An approach to this problem was taken by Rospert and collaborators (Dubaquié *et al.* 1998), who carried out an experiment importing *in vitro* translation products programmed from total yeast mRNA into mitochondria isolated from either Hsp60- or Hsp10-deficient yeast. In two-dimensional gel analyses, they observed a number of proteins that were imported but became insoluble inside both populations of mutant mitochondria, whereas they were fully soluble upon import into wild-type mitochondria. One of these was an 82 kDa protein, identified by NH_2-terminal sequencing as aconitase, a large monomeric iron–sulfur cluster (Fe_4S_4)-containing enzyme of the Krebs cycle. Not only chaperonin but also co-chaperonin were apparently required *in vivo* for productive folding of this large enzyme.

Subsequently, *in vitro* reconstitution studies with GroEL and GroES indicated that, indeed, productive refolding of acid-denatured aconitase could be accomplished only in the presence of both chaperonin components and ATP (Chaudhuri *et al.* 2001). Surprisingly, the renatured apo form of the enzyme remained bound to GroEL but could be subsequently released as active holoenzyme upon chemical formation of the Fe_4S_4 cluster. To elucidate the role of the GroES co-chaperonin in refolding of aconitase, studies were carried out with single- and mixed-ring complexes. SR1, an obligatorily *cis*-acting chaperonin, was unable to mediate refolding in the presence of GroES and ATP. Consistent with this, aconitase could not be protected from exogenous protease by GroES when a binary complex of aconitase–GroEL was incubated with ADP and GroES; that is, a *cis* ternary complex could not be formed. Thus, the requirement for GroES was postulated to involve its binding in *trans*, to the ring opposite that bound by polypeptide (Fig. 5). In agreement with this, a mixed double-ring complex containing a wild-type ring and one unable to bind polypeptide or GroES failed to refold aconitase. The step requiring GroES binding in *trans* was conjectured to drive polypeptide release. This was demonstrated in an experiment in which a biotinylated version of GroEL was mixed with a preformed binary complex between radiolabeled aconitase and wild-type GroEL and then incubated with ATP and GroES. Aconitase molecules released from the wild-type GroEL were detected by their association with the biotinylated GroEL, recovered upon incubation with avidin beads. Whereas more than half of the initially bound aconitase was recovered with the avidin beads after adding

https://doi.org/10.1017/S0033583503003883 Published online by Cambridge University Press

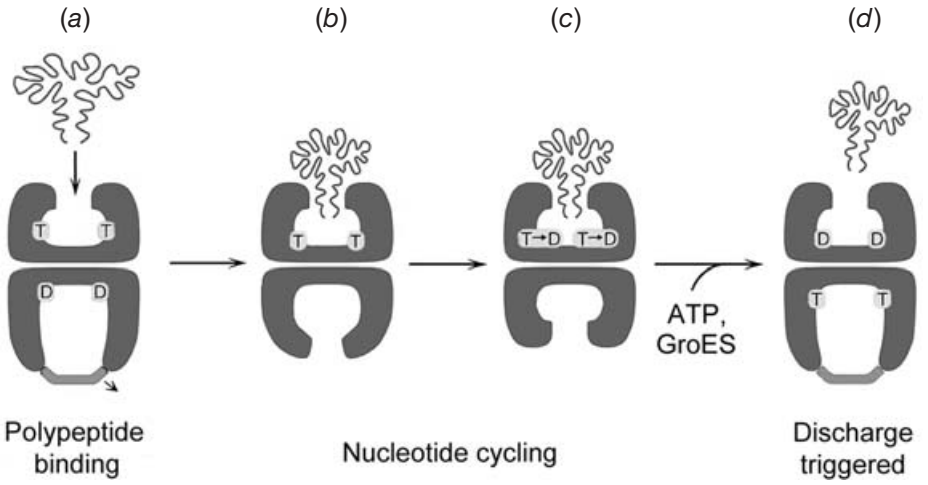

Fig. 5. Reaction pathway for folding in *trans* by GroEL–GroES. (*a*) Non-native polypeptide binds to the open *trans* ring of an asymmetric complex. (*b*, *c*) Because of steric interference by the large substrate polypeptide, GroES cannot bind to the ATP-containing ring, but the ATP hydrolysis cycle proceeds. When the *trans* ring binds ATP and GroES (*d*), this allosterically drives release of non-native polypeptide into the bulk solution, where it has an opportunity to fold or to be rebound by GroEL for another attempt at folding. Compared to the *cis* cycle outlined in Fig. 2, the *trans* cycle does not involve encapsulation and cannot use the potential advantages accorded by the *cis* cavity (see text for details).

ATP/GroES, only a small fraction was recovered after adding ATP alone. This contrasts with the discharge of *cis* ternary GroEL–GroES–polypeptide complexes, where binding of ATP alone in *trans* is sufficient to trigger release of the *cis* ligands.

As with the *cis* reaction, the refolding of aconitase was observed to be associated with multiple cycles of release and rebinding of non-native forms. In particular, when aconitase refolding was confined to a single round by using the hydrolysis-defective D398A mutant and ATP/GroES, only ∼10–15% of aconitase was refolded. Thus, a relatively small fraction reaches native form in one round of binding and release, and multiple cycles are apparently needed in order to recover all of the input molecules in native form. Each round probably represents an 'all-or-none' trial at reaching native form, with those molecules that fail to reach the native state becoming rebound in a conformation similar to that occupied before the trial.

A major difference between the *cis* and *trans* reaction topologies (cf. Figs 2 and 5) is that polypeptide released in *trans* is ejected directly into the bulk solution instead of into a sequestered, hydrophilic *cis* cavity. Because aconitase does not fold efficiently in the bulk solution when directly diluted from denaturant, the observation that a *trans* mechanism can be productive in the presence of chaperonins and ATP suggests that the reaction may involve unfolding actions, most likely associated with rebinding, in addition to simple release.

A further question raised by the study of aconitase refolding concerns whether the non-native forms bound by GroEL occupy a characteristic topology. Because the polypeptide even in its native state (for which there is an X-ray model) is far too large to be accommodated inside the open ring of GroEL, presumably only one domain, or a portion of a domain, can have access to the binding sites. It is possible that this would be a domain or surface of non-native aconitase that exhibits particular difficulty in achieving native form and thus may preferentially present a hydrophobic surface to the GroEL cavity. This remains to be addressed, along with the related question concerning topology of binding of smaller substrates.

https://doi.org/10.1017/S0033583503003883 Published online by Cambridge University Press

More recent studies have further probed the *trans* mechanism (Fig. 5) using *trans*-only GroEL–GroES complexes, in which GroES has been closely tethered to one of the two GroEL rings, preventing polypeptide from binding to the tethered GroEL ring but allowing normal nucleotide-driven association of the tethered GroES (Farr *et al.* 2003). Such complexes, in the absence of free GroES, only allow binding and release of non-native polypeptide in *trans*. Because the *trans*-only complexes can proceed through a normal nucleotide cycle, this allows recursive cycles of polypeptide binding and release in *trans*. When such complexes were tested *in vitro* with aconitase, they mediated refolding at a rate and extent identical to that of wild-type GroEL and GroES, as expected if the same mechanism was operative.

The *trans*-only complexes were also tested with smaller protein substrates that are able to fold in *cis*, because even such smaller species were excluded from the tethered ring. Thus, it was possible to directly examine whether such substrates could be folded by a *trans* mechanism. Remarkably, a substantial recovery of activity was observed for such stringent, GroEL–GroES-dependent substrates as Rubisco, MDH, and rhodanese. The rates of recovery, however, were 25–50% that of a *cis* reaction; that is, 2–4 times as many trials of folding were required to recover the same percentage of molecules in the native state. Because these substrates cannot fold in the bulk solution, such inefficiency may reflect that only a small percentage of the ensemble of conformations bound in the *trans* ring represents committed forms that can proceed to native form when released. For Rubisco, this was estimated to be $\sim$0·5% per cycle. The inability of a 25-fold excess of chaperonin over substrate to affect the kinetics of refolding supports the likelihood that commitment to reaching native form has already been determined while the substrate is bound at chaperonin; i.e. among the ensemble of bound conformations, a small fraction occupies a state (or states) that is committed to reaching native form upon release. This contrasts with the *cis* reaction, where commitment can also occur inside the *cis* chamber during the 10-s lifetime of this folding-active species. The extreme of this case is represented by SR1–polypeptide–GroES complexes, where, following a single round of release into the cavity of stable *cis* complexes, all of the substrate molecules ultimately commit to the native state.

Not only did the *trans* reaction support folding of stringent, GroEL–GroES-dependent substrates *in vitro*, but, when a plasmid encoding a *trans*-only complex was expressed in GroEL–GroES-deficient *E. coli*, it rescued the growth of this strain, although colony size was substantially smaller than with rescue by a wild-type plasmid. This indicated that the *trans* mechanism was able to mediate folding of all of the essential substrates of GroEL–GroES. Presumably, the relatively slower colony growth reflected that folding of smaller substrates, normally folded in *cis*, proceeded less efficiently. Nevertheless, this mechanism, involving cycles of binding in an open ring followed by release into the bulk solution, can apparently function universally for both large and small substrates. This further underscores the importance of the ability of an open GroEL ring to recursively bind non-native proteins.

10. Prospects for resolving the conformations and fate of polypeptide in the chaperonin reaction

The presence of both an ensemble of substrate protein conformations and of a megadalton machine acting on them presents a technical challenge to the better understanding of the folding trajectory of substrate proteins in the chaperonin system. Ideally, one would like to be able to follow a single molecule through its conformational trajectory, from binding to an open GroEL ring, through *cis* complex formation, during folding in the *cis* cavity, and after ejection into the

https://doi.org/10.1017/S0033583503003883 Published online by Cambridge University Press

bulk solution, followed by its rebinding to a GroEL ring. In particular, one would like to know what conformational changes distinguish a molecule that succeeds in reaching native form in a given folding cycle from those that fail and have to rebind for another trial. Are there single-molecule tools that can accomplish these aims? At present, this seems beyond immediate reach. While single GroEL molecules can be attached to surfaces and still function, and while single GFP molecules associated with them have been monitored by single-molecule fluorescence for reaching the native state (Taguchi *et al.* 2001), tools for following GFP or other proteins prior to reaching native form are lacking. Conceivably, distance analyses using single-molecule FRET between two heterologous probes attached to the same substrate protein could be informative (Schuler *et al.* 2002), but the sizes and properties of fluorophores commonly employed in single-molecule studies, such as fluorescein and Texas Red, make it likely that these reporters will themselves perturb the system. For example, Texas Red is $\sim$15 Å across, nearly 25% the diameter of the GroEL cavity. This presents a potential steric interference with the system by size considerations alone. In addition, the aromatic character of many such probes may recruit them directly to the hydrophobic apical binding surface, disturbing the normal physiology of the segment of substrate polypeptide to which they are attached. Assuming these perturbation issues could be solved with a new generation of reporter molecules, a multiplex system would be desirable, where multiple points on a single polypeptide chain could be related at the same time, instead of only two. The quest to develop single-molecule X-ray techniques (Hajdu, 2000; Neutze *et al.* 2000) could ultimately result in the observation of single GroEL molecules, perhaps including resolution of a bound substrate, but such techniques are in their infancy, currently being tested on small molecules. Clearly, the chaperonin system, as well as many others, would be ripe for testing as additional single-molecule methodologies evolve.

In ensemble studies, both EM techniques and NMR methods are becoming ever more powerful, offering the chance to see ensemble-averaged structures of non-native states. EPR spectroscopy might also be used informatively (for example, see Persson *et al.* 1999), but here also, only two-point distance behavior of an ensemble is observable at present. In initial efforts to use this technique to probe the secondary structure of a substrate bound to GroEL, we made individual cysteine substitutions at six consecutive positions along an α-helix in DHFR. Although the substitutions were tolerated in terms of allowing proper folding, once nitroxide spin probes were attached, for 5 of the 6 positions the modified DHFR molecules could no longer fold to the native state (A. Erbse, E. Bertelsen & A. L. Horwich, unpublished observations). Thus, the perturbation of conformation of substrate proteins by the various attached probes presents a significant problem. Although it may be feasible in some cases to use present reagents and methods to address substrate behavior in well-controlled experiments, new tools that allow specific labeling of substrate proteins and their observation while at the chaperonin are clearly desirable.

11. References

Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science* **181**, 223–230.

Ashcroft, A. E., Brinker, A., Coyle, J. E., Weber, F., Kaiser, M., Moroder, L., Parsons, M. R., Jager, J., Hartl, U. F., Hayer-Hartl, M. & Radford, S. E. (2002). Structural plasticity and noncovalent substrate binding in the GroEL apical domain. A study using electrospay ionization mass spectrometry and fluorescence binding studies. *Journal of Biological Chemistry* **277**, 33115–33126.

Badcoecoe, I. G., Smith, C. J., Wood, S., Halsall, D. J., Holbrook, J. J., Lund, P. & Clarke, A. R. (1991). Binding of a chaperonin to the folding intermediates of lactate dehydrogenase. *Biochemistry* **30**, 9195–9200.

Bhutani, N. & Udgaonkar, J. B. (2001). GroEL channels the folding of thioredoxin along one kinetic route. *Journal of Molecular Biology* **314**, 1167–1179.

Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L. & Sigler, P. B. (1994). The crystal structure of the bacterial chaperonin GroEL at 2·8 Å. *Nature* **371**, 578–586.

Brinker, A., Pfeifer, G., Kerner, M. J., Naylor, D. J., Hartl, F. U. & Hayer-Hartl, M. (2001). Dual function of protein confinement in chaperonin-assisted protein folding. *Cell* **107**, 223–233.

Buckle, A. M., Zahn, R. & Fersht, A. R. (1997). A structural model for GroEL-polypeptide recognition. *Proceedings of the National Academy of Sciences USA* **94**, 3571–3575.

Bukau, B. & Horwich, A. L. (1998). The Hsp70 and Hsp60 chaperone machines. *Cell* **92**, 351–366.

Chaudhuri, T. K., Farr, G. W., Fenton, W. A., Rospert, S. & Horwich, A. L. (2001). GroEL/GroES-mediated folding of a protein too large to be encapsulated. *Cell* **107**, 223–233.

Chen, J., Walter, S., Horwich, A. L. & Smith, D. L. (2001). Folding of malate dehydrogenase inside the GroEL–GroES cavity. *Nature Structural Biology* **8**, 721–728.

Chen, L. & Sigler, P. B. (1999). The crystal structure of a GroEL/peptide complex: plasticity as a basis for substrate diversity. *Cell* **99**, 757–768.

Cheng, M. Y., Hartl, F. U., Martin, J., Pollock, R. A., Kalousek, F., Neupert, W., Hallberg, E. M., Hallberg, R. L. & Horwich, A. L. (1989). Mitochondrial heat-shock protein hsp60 is essential for assembly of proteins imported into yeast mitochondria. *Nature* **337**, 620–625.

Corrales, F. J. & Fersht, A. R. (1995). The folding of GroEL-bound barnase as a model for chaperonin-mediated protein folding. *Proceedings of the National Academy of Sciences USA* **92**, 5326–5330.

Corrales, F. J. & Fersht, A. R. (1996). Towards a mechanism for GroEL–GroES chaperone activity: an ATPase-gated and -pulsed folding and annealing cage. *Proceedings of the National Academy of Sciences USA* **93**, 4509–4512.

Dill, K. A. & Chan, H. S. (1997). From Levinthal to pathways to funnels. *Nature Structural Biology* **4**, 10–19.

Dubaquié, Y., Looser, R., Funfschilling, U., Jeno, P. & Rospert, S. (1998). Identification of *in vivo* substrates of the yeast mitochondrial chaperonins reveals overlapping but non-identical requirements for hsp60 and hsp10. *EMBO Journal* **17**, 5868–5876.

Ellis, R. J. (1994). Molecular chaperones. Opening and closing the Anfinsen cage. *Current Biology* **4**, 633–635.

Ewalt, K. L., Hendrick, J. P., Houry, W. A. & Hartl, F. U. (1997). *In vivo* observation of polypeptide flux through the bacterial chaperonin system. *Cell* **90**, 491–500.

Falke, S., Fisher, M. T. & Gogol, E. P. (2001). Structural changes in GroEL effected by binding a denatured protein substrate. *Journal of Molecular Biology* **308**, 569–577.

Farr, G. W., Fenton, W. A., Chaudhuri, T. K., Clarke, D. K., Saibil, H. R. & Horwich, A. L. (2003). Folding with and without encapsulation by *cis* and *trans*-only GroEL–GroES complexes. *EMBO Journal* **22**, 3220–3230.

Farr, G. W., Furtak, K., Rowland, M. B., Ranson, N. A., Saibil, H. R., Kirchhausen, T. & Horwich, A. L. (2000). Multivalent binding of nonnative substrate proteins by the chaperonin GroEL. *Cell* **100**, 561–573.

Fayet, O., Ziegelhoffer, T. & Georgopoulos, C. (1989). The GroES and GroEL heat shock gene products of *Escherichia coli* are essential for bacterial growth at all temperatures. *Journal of Bacteriology* **171**, 1379–1385.

Fenton, W. A. & Horwich, A. L. (1997). GroEL-mediated protein folding. *Protein Science* **6**, 743–760.

Fenton, W. A., Kashi, Y., Furtak, K. & Horwich, A. L. (1994). Residues in chaperonin GroEL required for polypeptide binding and release. *Nature* **371**, 614–619.

Fiaux, J., Bertelsen, E. B., Horwich, A. L. & Wüthrich, K. (2002). NMR analysis of a 900K GroEL–GroES complex. *Nature* **418**, 207–211.

Gaitanaris, G. A., Vysokanov, A., Hung, S. C., Gottesman, M. E. & Gragerov, A. (1994). Successive action of *Escherichia coli* chaperones *in vivo*. *Molecular Microbiology* **14**, 861–869.

Gervasoni, P., Staudenmann, W., James, P., Gehrig, P. & Plückthun, A. (1996). β-Lactamase binds to GroEL in a conformation highly protected against hydrogen/deuterium exchange. *Proceedings of the National Academy of Sciences USA* **93**, 12139–12194.

Goldberg, M. S., Zhang, J., Sondek, S., Matthews, C. R., Fox, R. O. & Horwich, A. L. (1997). Native-like structure of a protein-folding intermediate bound to the chaperonin GroEL. *Proceedings of the National Academy of Sciences USA* **94**, 1080–1085.

Goloubinoff, P., Gatenby, A. A. & Lorimer, G. H. (1989). GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose bisphosphate carboxylase oligomers in *Escherichia coli*. *Nature* **337**, 44–47.

Grantcharova, V., Alm, E. J., Baker, D. & Horwich, A. L. (2001). Mechanisms of protein folding. *Current Opinion in Structural Biology* **11**, 70–82.

Gray, T. E. & Fersht, A. R. (1993). Refolding of barnase in the presence of GroE. *Journal of Molecular Biology* **232**, 1197–1207.

Groß, M., Robinson, C. V., Mayhew, M., Hartl, F. U. & Radford, S. E. (1996). Significant hydrogen exchange protection in GroEL-bound DHFR is maintained during iterative rounds of substrate cycling. *Protein Science* **5**, 2506–2513.

Hajdu, J. (2000). Single-molecule X-ray diffraction. *Current Opinion in Structural Biology* **10**, 569–573.

Hartl, F. U. & Hayer-Hartl, M. (2002). Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* **295**, 1852–1858.

Hlodan, R., Tempst, P. & Hartl, F. U. (1995). Binding of defined regions of a polypeptide to GroEL and its implications for chaperonin-mediated protein folding. *Nature Structural Biology* **2**, 587–595.

Horwich, A. L. & Fenton, W. A. (2003). The role of ATP in directing chaperonin-mediated protein folding. In *The Enzymes*, vol. 23 (ed. F. Tamanoi). New York: Academic Press.

Horwich, A. L., Low, K. B., Fenton, W. A., Hirshfield, I. N. & Furtak, K. (1993). Folding *in vivo* of bacterial cytoplasmic proteins. *Cell* **74**, 909–917.

Houry, W. A., Frishman, D., Eckerskorn, C., Lottspeich, F. & Hartl, F. U. (1999). Identification of *in vivo* substrates of the chaperonin GroEL. *Nature* **402**, 147–154.

Hunt, J. F., Weaver, A. J., Landry, S. J., Gierasch, L. & Deisenhofer, J. (1996). The crystal structure of the GroES co-chaperonin at 2·8 Å resolution. *Nature* **379**, 37–45.

Itzhaki, L. S., Otzen, D. E. & Fersht, A. R. (1995). Nature and consequences of GroEL-protein interactions. *Biochemistry* **34**, 14581–14587.

Jackson, G. S., Staniforth, R. A., Halsall, D. J., Atkinson, T., Holbrook, J. J., Clarke, A. R. & Burston, S. G. (1993). Binding and hydrolysis of nucleotides in the chaperonin catalytic cycle: implications for the mechanism of assisted protein folding. *Biochemistry* **32**, 2554–2563.

Kobayashi, N., Freund, S. M., Chatellier, J., Zahn, R. & Fersht, A. R. (1999). NMR analysis of the binding of a rhodanese peptide to a minichaperone in solution. *Journal of Molecular Biology* **292**, 181–190.

Laminet, A. A., Ziegelhoffer, T., Georgopoulos, C. & Pluckthun, A. (1990). The *Escherichia coli* heat shock proteins GroEL and GroES modulate the folding of the β-lactamase precursor. *EMBO Journal* **9**, 2315–2319.

Landry, S. J. & Gierasch, L. M. (1991). The chaperonin GroEL binds a polypeptide in an alpha-helical conformation. *Biochemistry* **30**, 7359–7362.

Landry, S. J., Jordan, R., McMacken, R. & Gierasch, L. M. (1992). Different conformations for the same polypeptide bound to chaperones DnaK and GroEL. *Nature* **355**, 455–457.

Landry, S. J., Zeilstra-Ryalls, J., Fayet, O., Georgopoulos, C. & Gierasch, L. M. (1993). Characterization of a functionally important mobile domain of GroES. *Nature* **364**, 255–258.

Lin, S., Schwarz, F. P. & Eisenstein, E. (1995). The hydrophobic nature of GroEL-substrate binding. *Journal of Biological Chemistry* **270**, 1011–1014.

Martin, J., Langer, T., Boteva, R., Schramel, A., Horwich, A. L. & Hartl, F. U. (1991). Chaperonin-mediated protein folding at the surface of groEL through a 'molten globule'-like intermediate. *Nature* **352**, 36–42.

Matouschek, A., Kellis, J. T. Jr., Serrano, L. & Fersht, A. R. (1989). Mapping the transition state and pathway of protein folding by protein engineering. *Nature* **340**, 122–126.

Mayer, M. P., Brehmer, D., Gässler, C. S. & Bukau, B. (2002). Hsp70 chaperone machines. *Advances in Protein Chemistry* **59**, 1–44.

Mayor, U., Guydosh, N. R., Johnson, C. M., Grossman, J. G., Sato, S., Jas, G. S., Freund, S. M. V., Alonso, D. O. V., Daggett, V. & Fersht, A. R. (2003). The complete folding pathway of a protein form nanoseconds to microseconds. *Nature* **421**, 863–867.

Mayor, U., Johnson, C. M., Daggett, V. & Fersht, A. R. (2000). Protein folding and unfolding in microseconds to nanoseconds by experiment and simulation. *Proceedings of the National Academy of Sciences USA* **97**, 13518–13522.

Mendoza, J. A., Rogers, E., Lorimer, G. H. & Horowitz, P. M. (1991). Chaperonins facilitate the *in vitro* folding of monomeric mitochondrial rhodanese. *Journal of Biological Chemistry* **266**, 13044–13049.

Neutze, R., Wouts, R., Van Der Spoel, D., Weckert, E. & Hajdu, J. (2000). Potential for biomolecular imaging with femtosecond X-ray pulses. *Nature* **406**, 752–757.

Oliveberg, M. (2001). Characterization of the transition states for protein folding: towards a new level of mechanistic detail in protein engineering analysis. *Current Opinion in Structural Biology* **11**, 94–100.

Peralta, F., Hartman, D. J., Hoogenraad, N. J. & Høj, P. B. (1994). Generation of a stable folding intermediate which can be rescued by the chaperonins GroEL and GroES. *FEBS Letters* **339**, 45–49.

Persson, M., Hammarström, P., Lindgren, M., Jonsson, B.-H., Svensson, M. & Carlsson, U. (1999). EPR mapping of interactions between spin-labeled variants of human carbonic anhydrase II and GroEL: evidence for increased flexibility of the hydrophobic core by the interaction. *Biochemistry* **38**, 432–441.

Pervushin, K., Riek, R., Wider, G. & Wüthrich, K. (1997). Attenuated T_2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proceedings of the National Academy of Sciences USA* **94**, 12366–12371.

Preuss, M., Hutchinson, J. P. & Miller, A. D. (1999). Secondary structure forming propensity coupled with amphiphilicity is an optimal motif in a peptide or protein for association with chaperonin 60 (GroEL). *Biochemistry* **38**, 10272–10286.

Ranson, N. A., Dunster, N. J., Burston, S. G. & Clarke, A. R. (1995). Chaperonins can catalyse the reversal of early aggregation steps when a protein misfolds. *Journal of Molecular Biology* **250**, 581–586.

Reid, B. G. & Flynn, G. C. (1996). GroEL binds to and unfolds rhodanese posttranslationally. *Journal of Biological Chemistry* **271**, 7212–7217.

Riek, R., Fiaux, J., Bertelsen, E. B., Horwich, A. L. & Wüthrich, K. (2002). Solution NMR techniques for large molecular and supramolecular structures. *Journal of the American Chemical Society* **124**, 12144–12153.

Riek, R., Pervushin, K. & Wüthrich, K. (2000). TROSY and CRINEPT: NMR with large molecular and supramolecular structures in solution. *Trends in Biochemical Sciences* **25**, 462–468.

Robinson, C. V., Gross, M., Eyles, S. J., Ewbank, J. J., Mayhew, M., Hartl, F. U., Dobson, C. M. & Radford, S. E. (1994). Conformation of GroEL-bound alpha-lactalbumin probed by mass spectrometry. *Nature* **372**, 646–651.

Ryan, M. T. & Pfanner, N. (2001). Hsp70 proteins in protein translocation. *Advances in Protein Chemistry* **59**, 223–242.

Rye, H. S., Burston, S. G., Fenton, W. A., Beechem, J. M., Xu, Z., Sigler, P. B. & Horwich, A. L. (1997). Distinct actions of *cis* and *trans* ATP within the double ring of the chaperonin GroEL. *Nature* **388**, 792–798.

Saibil, H. R., Horwich, A. L. & Fenton, W. A. (2002). Allostery and protein substrate conformational change during GroEL/GroES-mediated protein folding. *Advances in Protein Chemistry* **59**, 45–72.

Schmidt, M., Buchner, J., Todd, M. J., Lorimer, G. H. & Viitanen, P. V. (1994). On the role of GroES in the chaperonin-assisted folding reaction. Three case studies. *Journal of Biological Chemistry* **269**, 10304–10311.

Schuler, B., Lipman, E. A. & Eaton, W. A. (2002). Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature* **419**, 743–747.

Shewmaker, F., Maskos, K., Simmerling, C. & Landry, S. J. (2001). The disordered mobile loop of GroES folds into a defined beta-hairpin upon binding GroEL. *Journal of Biological Chemistry* **276**, 31257–31264.

Shtilerman, M., Lorimer, G. H. & Englander, S. W. (1999). Chaperonin function: folding by forced unfolding. *Science* **284**, 822–825.

Sigler, P. B., Xu, Z., Rye, H. S., Burston, S. G., Fenton, W. A. & Horwich, A. L. (1998). Structure and function in GroEL-mediated protein folding. *Annual Review of Biochemistry* **67**, 581–608.

Smith, K. E. & Fisher, M. T. (1995). Interactions between the GroE chaperonins and rhodanese: multiple intermediates and release and rebinding. *Journal of Biological Chemistry* **270**, 21517–21523.

Staniforth, R. A., Burston, S. G., Atkinson, T. & Clarke, A. R. (1994). Affinity of chaperonin-60 for a protein substrate and its modulation by nucleotides and chaperonin-10. *Biochemical Journal* **300**, 651–658.

Taguchi, H., Ueno, T., Tadakuma, H., Yoshida, M. & Funatsu, T. (2001). Single-molecule observation of protein-protein interactions in the chaperonin system. *Nature Biotechnology* **19**, 861–865.

Thirumalai, D. & Lorimer, G. H. (2001). Chaperonin-mediated protein folding. *Annual Review of Biophysics and Biomolecular Structure* **30**, 245–269.

Thiyagarajan, P., Henderson, S. J. & Joachimiak, A. (1996). Solution structures of GroEL and its complex with rhodanese from small-angle neutron scattering. *Structure* **4**, 79–88.

Todd, M. J., Lorimer, G. H. & Thirumalai, D. (1996). Chaperonin-facilitated protein folding: optimization of rate and yield by an iterative annealing mechanism. *Proceedings of the National Academy of Sciences USA* **93**, 4030–4035.

Todd, M. J., Viitanen, P. V. & Lorimer, G. H. (1994). Dynamics of the chaperonin ATPase cycle: implications for facilitated protein folding. *Science* **265**, 659–666.

Van Der Vies, S. M., Viitanen, P. V., Gatenby, A. A., Lorimer, G. H. & Jaenicke, R. (1992). Conformational states of ribulosebisphosphate carboxylase and their interaction with chaperonin 60. *Biochemistry* **31**, 3635–3644.

Van Montfort, R., Slingsby, C. & Vierling, E. (2002). Structure and function of the small heat shock protein/α-crystallin family of molecular chaperones. *Advances in Protein Chemistry* **59**, 105–156.

Viitanen, P. V., Donaldson, G. K., Lorimer, G. H., Lubben, T. H. & Gatenby, A. A. (1991). Complex interactions between the chaperonin 60 molecular chaperone and dihydrofolate reductase. *Biochemistry* **30**, 9716–9723.

Viitanen, P. V., Gatenby, A. A. & Lorimer, G. H. (1992). Purified chaperonin 60 (groEL) interacts with the nonnative states of a multitude of *Escherichia coli* proteins. *Protein Science* **1**, 363–369.

Walter, S., Lorimer, G. H. & Schmid, F. X. (1996). A thermodynamic coupling mechanism for GroEL-mediated folding. *Proceedings of the National Academy of Sciences USA* **93**, 9425–9430.

Wang, Z., Feng, H., Landry, S. J., Maxwell, J. & Gierasch, L. M. (1999). Basis of substrate binding by the chaperonin GroEL. *Biochemistry* **38**, 12537–12546.

Weissman, J. S., Kashi, Y., Fenton, W. A. & Horwich, A. L. (1994). GroEL-mediated protein folding proceeds by multiple rounds of binding and release of nonnative forms. *Cell* **78**, 693–702.

Weissman, J. S., Hohl, C. M., Kovalenko, O., Kashi, Y., Chen, S. X., Braig, K., Saibil, H. R., Fenton, W. A. & Horwich, A. L. (1995). Mechanism of GroEL action – productive release of polypeptide from a sequestered position under GroES. *Cell* **83**, 577–587.

Weissman, J. S., Rye, H. S., Fenton, W. A., Beechem, J. M. & Horwich, A. L. (1996). Characterization of the active intermediate of a GroEL–GroES-mediated protein folding reaction. *Cell* **84**, 481–490.

Xu, Z. H., Horwich, A. L. & Sigler, P. B. (1997). The crystal structure of the asymmetric GroEL–GroES–$(ADP)_7$ chaperonin complex. *Nature* **388**, 741–749.

Zahn, R., Axmann, S. E., Rucknagel, K. P., Jaeger, E., Laminet, A. A. & Pluckthun, A. (1994). Thermodynamic partitioning model for hydrophobic binding of polypeptides by GroEL. I. GroEL recognizes the signal sequences of β-lactamase precursor. *Journal of Molecular Biology* **242**, 150–164.

Zahn, R., Perrett, S. & Fersht, A. R. (1996a). Conformational states bound by the molecular chaperones GroEL and SecB: a hidden unfolding (annealing) activity. *Journal of Molecular Biology* **261**, 43–61.

Zahn, R., Perrett, S., Stenberg, G. & Fersht, A. R. (1996b). Catalysis of amide protein exchange by the molecular chaperones GroEL and SecB. *Science* **271**, 642–645.

Zahn, R. & Plückthun, A. (1994). Thermodynamic partitioning model for hydrophobic binding of polypeptides by GroEL. II. GroEL recognizes thermally unfolded mature β-lactamase. *Journal of Molecular Biology* **242**, 165–174.