Annu. Rev. Biochem. 1998. 67:581–608

STRUCTURE AND FUNCTION IN GroEL-MEDIATED PROTEIN FOLDING

Paul B. Sigler,[1,2] *Zhaohui Xu,*[1,2] *Hays S. Rye,*[2,3] *Steven G. Burston,*[3] *Wayne A. Fenton,*[3] *and Arthur L. Horwich*[2,3]
[1]Department of Molecular Biophysics and Biochemistry, [2]Howard Hughes Medical Institute, and [3]Department of Genetics, School of Medicine, Yale University, New Haven, Connecticut 06510; e-mail: sigler@csb.yale.edu

KEY WORDS: ATP, chaperonin, GroES, Hsp60

ABSTRACT

Recent structural and biochemical investigations have come together to allow a better understanding of the mechanism of chaperonin (GroEL, Hsp60)–mediated protein folding, the final step in the accurate expression of genetic information. Major, asymmetric conformational changes in the GroEL double toroid accompany binding of ATP and the cochaperonin GroES. When a nonnative polypeptide, bound to one of the GroEL rings, is encapsulated by GroES to form a *cis* ternary complex, these changes drive the polypeptide into the sequestered cavity and initiate its folding. ATP hydrolysis in the *cis* ring primes release of the products, and ATP binding in the *trans* ring then disrupts the *cis* complex. This process allows the polypeptide to achieve its final native state, if folding was completed, or to recycle to another chaperonin molecule, if the folding process did not result in a form committed to the native state.

CONTENTS

0066-4154/98/0701-0581$08.00

PURPOSE

The main purpose of this review is to put forth an integrated structure/function analysis of chaperonin-assisted protein folding that coordinates recent high-resolution crystallographic results with functional studies in solution. To provide an appropriate context, however, we first review where molecular chaperones fit in the scheme of biological catalysis and the special problems presented by the cell's need for quick and accurate expression of genetic information, ultimately in the form of properly folded functional proteins and their assemblies.

THE ROLE OF MOLECULAR CHAPERONES IN THE CELL

Biological Catalysts

Biological systems are necessarily metastable. They are created, modulated, and destroyed according to a temporal plan that meets the survival needs of the cell, organism, and species. This metastability applies over a wide time and spatial frame, ranging from the femtosecond and Ångstrom scale of molecular dynamics to the millennia lifetimes and 200-foot cellulose scaffold of the giant sequoia. Clearly, no biological system is close to true equilibrium or it would be dead—thus systems survive by consuming free energy and regulating rates. For example, a human turns over 40 kg of ATP daily, and 90% of the information in a genome encodes biological catalysts. To properly coordinate life processes, a system must have the capability to control rates of reactions independently; that is, the catalysts must be specific and carry out their tasks precisely. For interesting philosophical reasons beyond the scope of this review, one can justify the fact that biological catalysts almost always speed up reactions. Herein lies an important trade-off: The more quickly something is done, the more difficult it is to do it accurately. Accuracy and specificity often are sacrificed in the name of speed and vice versa. This balance is especially important in optimizing the rate of information transfer from the gene to the tertiary and quaternary structures of functional nucleic acids and proteins.

Biological catalysts can be categorized by the nature of the reactions they catalyze. *Enzymes* make and break covalent bonds. They address rates that alter the primary structure, where the components of the free energy of the activation barriers are both enthalpic (disruption of the electronic structure) and entropic (restricting the orientation and trajectory of the reacting compounds). *Channels and small-molecule transporters* facilitate the passage of ions and polar molecules through restrictive membrane barriers, and like enzymes that catalyze thermodynamically unfavorable reactions, they often couple free energy input to the movement of a substrate against a gradient.

Molecular chaperones address a problem on a different scale, that of achieving the correct tertiary structure (and in an indirect way quaternary structure) of proteins. Here the notion of an activation barrier cannot be viewed simply as a high free energy point along a linear reaction coordinate. Rather, a free energy hypersurface must be envisioned, defined by the polypeptide itself and its ultimate cellular destination, such as the bacterial cytosol or periplasm, a eukaryotic secretory vesicle, or an organelle. This folding hypersurface is replete with false minima, some of which are deep enough to entrap the wayward nonnative polypeptide for a period that can be considered irreversible on the cell's time scale (see Reference 1, for example). Just as enzymes guide chemical reactions along reaction paths with minimal free energy barriers, molecular chaperones catalyze productive folding by escorting nonnative polypeptides across the folding free energy surface, avoiding and, if necessary, reversing states that lead to the truly dead-end pitfalls of aggregation and/or proteolysis. Finally there are *scaffolding or assembly proteins*, which facilitate the efficient formation of quaternary structures. These are encountered, for example, in the assembly of a virion, the DNA replication machinery, or a preinitiation transcription complex.

The Entropic Barrier to Expressing Genes Accurately

Consider the likelihood that a gene will be replicated, transcribed, and translated nearly error free. If the cell completely lost its mechanisms for ensuring fidelity in replicating or transcribing the 600 nucleotides that encode the expressed component of a 200–amino acid protein, then the polymerases, in principle, might synthesize any one of 4^{600}-1 incorrect sequences. Moreover, translation of the polypeptide has a potential for 20^{200} errors. It would take about 10^{80} universes to accommodate such a collection of faulty proteins, assuming they would be completely folded and packed effectively. Even if the cell could tolerate polypeptide products with an average of one mistake per chain, the error rate in translation alone would have to be less than 10^{-4} at each step. If the process were slow and deliberate enough, such a low error rate might be achievable, but

aminoacylation and encoded ribosomal synthesis of polypeptides proceeds at very high speed; therefore, the likelihood of "noise-free" information transfer is negligible. The same considerations apply to replicating and transcribing the gene. How does the cell accomplish this seemingly impossible task? The answer is analogous to the process by which the writers of this review and the editors reduce errors in this chapter to a tolerable limit: They proofread and edit. Thus, the entropic activation barrier to prompt and accurate genetic expression at the level of enzymatic polymerization is overcome by the lavish use of nucleoside triphosphate to proofread and edit the polymer sequences. The same is true in the final steps of protein folding and assembly: Molecular chaperones, especially chaperonins, consume ATP, sometimes lavishly, to correct the inevitable and potentially irreversible mistakes in folding.

Categories of Molecular Chaperones

Studies of the past decade have focused intensively on a number of different types of conformational "editors"—families of molecular chaperones. Many of these were first identified by their increased expression during heat shock, where the workload of editing goes far beyond the "noise" of normal growth conditions. Each chaperone family appears to recognize specific nonnative protein conformations and acts on them in a characteristic way. However, all the families appear to share the ability to recognize hydrophobic surfaces exposed in nonnative proteins, surfaces that ultimately become buried in the interior of native proteins. The Hsp70 class appears to prefer hydrophobic regions in extended polypeptide chains (2–4), for example, those transiting cellular membranes (5). Through the action of ATP and cochaperones, polypeptides are released from the Hsp70 proteins, retaining their extended state in preparation for further steps of biogenesis. The Hsp20 class chaperones appear to act as small globular collectives like sponges, binding a multitude of nonnative species at their outer surface during heat shock (6, 7). On return to normal temperature, the substrate proteins are transferred to other chaperones and return to native form. The Hsp90 class chaperones appear to generally function in large multimeric complexes, recognizing a host of important signal transduction proteins in forms that may be near native and that in many cases await the appearance/binding of a ligand, such as a steroid hormone, for final conversion to an active conformation (8, 9).

But perhaps the most interesting of these is the chaperonin class (comprising the Hsp60/GroEL and TF55/CCT families) of ring-shaped complexes that recognize exposed hydrophobic surfaces of a wide range of globular nonnative conformations and bind them in the central cavity. Here, binding confers stability upon exposed, aggregation-prone, hydrophobic surfaces, preventing irreversible aggregation. Multivalent binding by the high density of

neighboring hydrophobic sites lining this machine's channel may also serve to untangle misfolded structures. Most dramatic, however, is that when a chaperonin binds ATP and, in some cases, a cochaperonin, it is converted to an active device that may further unfold a polypeptide and then releases it into a new, shielded environment—encapsulated, expanded in volume, and now hydrophilic—favorable to folding to the native state. This type of assistance probably represents an ultimate level of general conformational editing to be found in the cell. We describe below the structural and functional workings of this machine, focusing largely on the *Escherichia coli* chaperonin, GroEL (10–12).

CHAPERONINS

Chaperonin Architecture

From the outset, all chaperonins (cpn60), irrespective of their cellular or subcellular origin, were seen to be double toroids in the electron microscope (13, 14). The need for this architecture is now better understood and is discussed more fully below. The bacterial and organellar chaperonins are assisted by single-ring cochaperonins (cpn10) (15, 16). The *E. coli* cochaperonin, GroES, is bound as a cap at one end of GroEL to form a bulletlike structure (17–19) or, sometimes, at both ends to form a football-like structure (20, 21). As shown in Figure 1 (opposite p. 586), the asymmetric bullets represent well-defined, biochemically accessible states in the folding cycle. The football structures probably are transient states whose presence can be enriched by nonnatural magnesium ion analogs such as Mn^{2+} or other conditions (22, 23) but which are not obligatory participants in productive folding reactions (22). Whether they constitute an intermediate form that occurs in vivo remains unclear.

Early electron microscopy (EM) studies of both unlabeled (18) and gold-tagged (24) polypeptides showed nonnative polypeptide bound in the central channel of the double toroid. This finding has been confirmed by single-particle correlation reconstructions of cryoelectron micrographs (25) and low-angle neutron-scattering curves (26). The latter suggest that the bound peptide distributes itself in the shape of a champagne cork adhering to and protruding from the end of the central cavity.

The symmetry of the chaperonin assembly depends on the number and uniqueness of the subunits comprising the rings. Bacterial (GroEL) and mitochondrial (Hsp60) chaperonins are composed of only one type of subunit. Hence, the two rings can be considered identical and, extrapolating from the crystallographic studies of GroEL (and in agreement with earlier EM studies), they are all stacked back to back with twofold rotational symmetry. The

number of subunits in each ring varies. Those from eubacteria, mitochondria, and chloroplasts contain seven. Archaeal members of the TF55/CCT family have eight (27) or nine (28); the eight-membered archaeal thermosome contains two types of subunits that alternate within a ring to give fourfold rotational symmetry (27). Finally, the eukaryotic cytosolic chaperonin rings contain eight different gene products, which, cross-linking studies have indicated, are arranged in an explicit permutation (29). Unlike the bacterial chaperonins and their counterparts in endosymbiotic organelles, archaeal and eukaryal cytosolic chaperonins appear to function without cochaperonins (30–32). A possible structural basis for this is described below.

GroEL and GroES Structures

In 1994, the first high-resolution structure of GroEL was determined to 2.8 Å (33). The structure is a porous, thick-walled cylinder that is slightly taller than it is wide and contains a substantial central cavity or channel (Figure 1). As indicated by the electron microscope, it is composed of two rings of seven subunits arranged with nearly exact sevenfold rotational symmetry. The rings are arranged back to back, contacting one another through an extensive, yet remarkably flat, equatorial interface that contains seven molecular dyads (so exact that one of them coincides with a crystallographic dyad).

The GroEL subunit (547 amino acids) folds into three distinctive domains (Figure 2):

1. A well-ordered, highly α-helical, "equatorial" domain that forms a solid foundation around the waist of the assembly. In addition to providing most of the intersubunit contacts as described above, the equatorial domain contributes most of the residues that constitute the ATPase site.

3. An "apical" domain that surrounds the opening at the ends of the central channel. The apical domain is considerably less well ordered and shows local flexibility within the domain as well as en bloc movements around a hinge that connects it to the intermediate domain.

3. A small, slender "intermediate" domain at the periphery of the cylinder that links the equatorial domain to the apical domain.

The crystal structures of isolated GroES (a heptamer of 10 kDa subunits) at 2.8 Å (34) and cpn10 from *Mycobacterium leprae* at 3.5 Å (35) show a sevenfold rotationally symmetric, dome-shaped architecture, about 75 Å in diameter and 30 Å high. Each of the seven subunits has a core β-barrel structure (Figure 2*C*) with two β-hairpin loops. One arches upward and inward from the top aspect, collectively forming the top of the dome. The other, at the lower lateral aspect, contains a structure that is disordered in all but one subunit, where

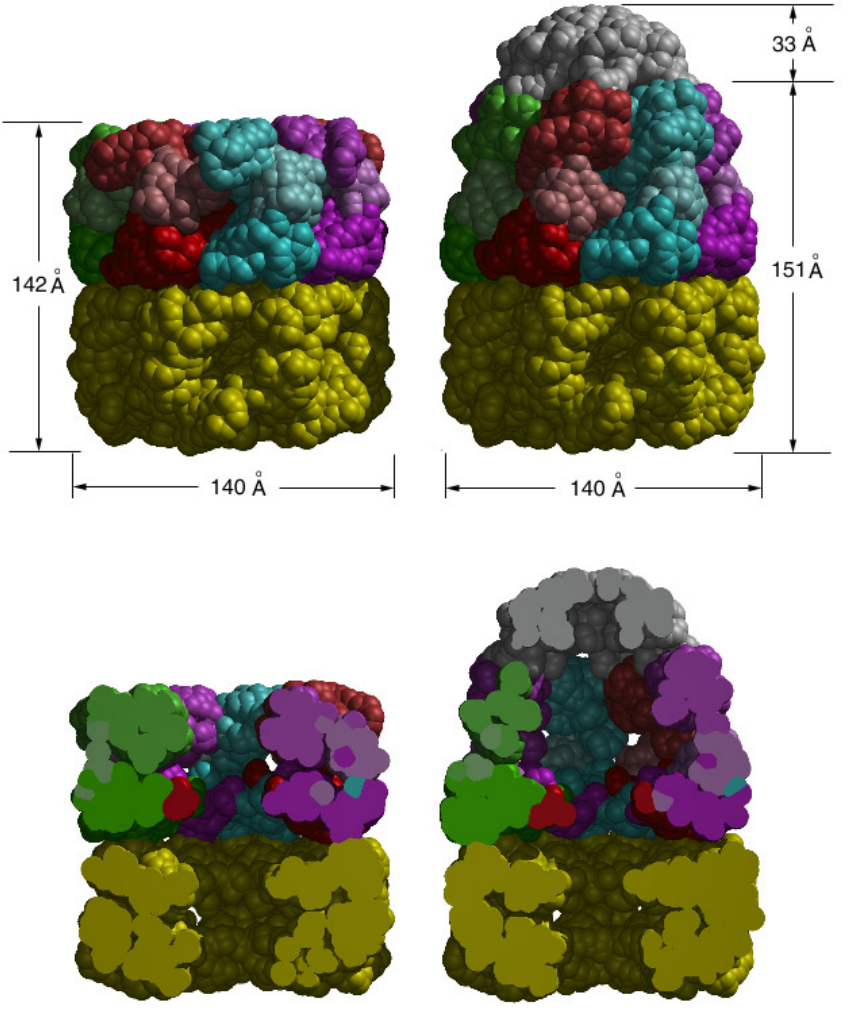


Figure 1 Overall architecture and dimensions of GroEL and GroEL-GroES-$(ADP)_7$. van der Waals space-filling models (6Å spheres around Cα) of GroEL (*left*) and GroEL-GroES-$(ADP)_7$ (*right*). *Upper panels* are outside views, showing outer dimensions; *lower panels* show the insides of the assemblies and were generated by slicing off the front half with a vertical plane that contains the cylindrical axis. Various colors are used to distinguish the subunits of GroEL in the upper ring. The domains are indicated by shading: equatorial, *dark hue*; apical, *medium hue*; intermediate, *light hue.* The lower GroEL ring is *uniformly yellow*. GroES is *uniformly gray.*

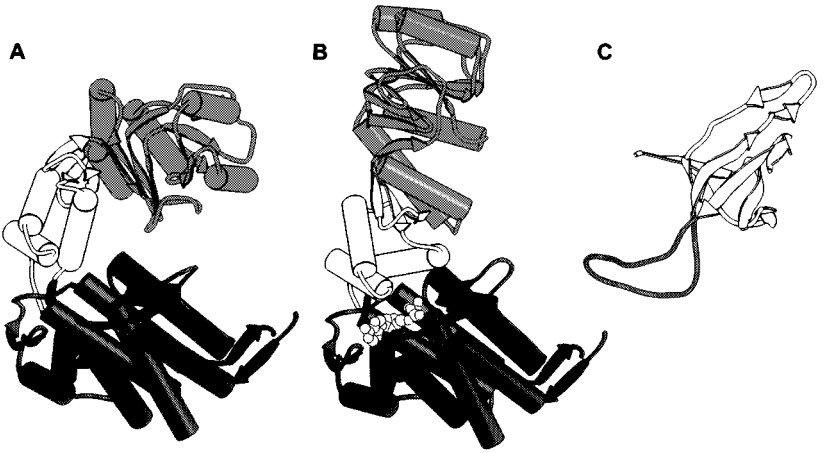


Figure 2 The subunit structures of GroEL and GroES. The subunit structure of GroEL as seen from within (*A*) an unliganded GroEL ring, as well as the *trans* ring of a GroEL-GroES-$(ADP)_7$ complex, and (*B*) the *cis* ring of a GroEL-GroES-$(ADP)_7$ complex. The *shading* corresponds to the three GroEL domains: equatorial, *dark gray*; intermediate, *white*; and apical, *medium gray*. Note the nucleotide (*white*) in the equatorial binding site in (*B*). (*C*) The subunit structure of GroES as seen in the GroEL-GroES-$(ADP)_7$ complex in the orientation required for interaction with the GroEL subunit in (*A*). The mobile loop is *dark gray*. The en bloc conformational changes relating the structures in (*A*) and (*B*) are shown diagrammatically in Figure 6.

it is stabilized by crystal contacts. Earlier proton nuclear magnetic resonance (NMR) studies indicated that this loop (Glu16 to Ala32) was mobile (hence the term mobile loop) but became better ordered on interaction with GroEL (36). These results have been confirmed in the GroEL-GroES-$(ADP)_7$ complex (see below).

The Chaperonin Reaction Cycle

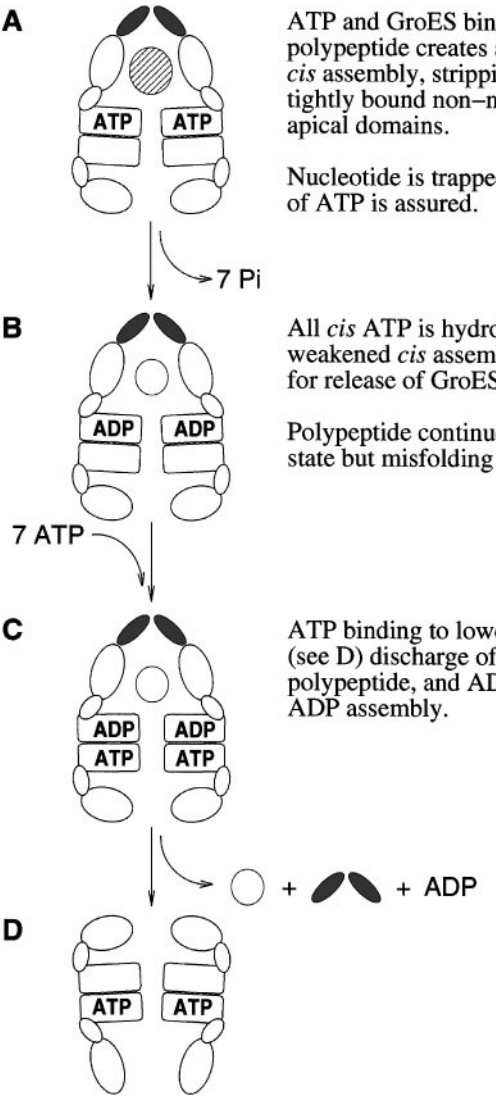
As noted above, the function of the GroEL-GroES chaperonin system as a biological catalyst is to facilitate the folding of proteins within the cellular environment—an enviroment that a variety of critical proteins, or at least some of their folding intermediates, find hostile during their journey to the native state. At its simplest, the chaperonin reaction consists of the cyclic binding and release of target polypeptides. When one considers the potentially tangled and aggregated confusion in which folding polypeptides might find themselves, it seems remarkable that such a simple mechanism can be so effective. Yet GroEL-GroES can routinely rescue greater than 80% of a denatured protein population that would generate no more than a few percent of native molecules without chaperonin (37). As the devil is always in the details, the secrets behind the

success of this deceptively simple mechanism lie in the remarkable structure of the GroEL-GroES protein complex, how it changes during the ATP-consuming reaction cycle, and exactly what effects the conformational states of the chaperonin have on the energetics of target polypeptides. Essentially, GroEL modulates its affinity for folding intermediates through the binding and hydrolysis of ATP. The highly coordinated act of binding and releasing substrate proteins then becomes sufficient to drive an otherwise dead-end folding reaction all the way to the native state. This reaction cycle can be operationally divided into four phases (Figure 3): (*a*) binding of polypeptide by GroEL (not shown); (*b*) release of polypeptide into the central chamber and initiation of folding, accomplished by binding GroES and ATP to form the high-energy *cis*-active state (Figure 3*A*); (*c*) decay of the high-energy *cis* state by hydrolysis of ATP, priming the *cis* assembly for the release of bound peptide, GroES, and ADP (Figure 3*B*); and (*d*) binding of ATP to the *trans* ring, providing the trigger to discharge GroES and entrapped polypeptide from the opposite side (Figure 3*C*).

Phase I: Polypeptide Binding

The central channel of GroEL functions as two separate cavities, one in each ring, that are separated from each other by the confluence of the crystallographically disordered 24–amino acid C-terminal segments of the seven subunits. In electron micrographs and by small-angle neutron scattering, these segments appear to coalesce and block the central channel at the level of the equatorial domain (26, 38). Because polypeptide substrates cannot escape through the equatorial segments of single-ring GroEL mutants that are otherwise competent in refolding reactions (39, 40), the pair of cavities in double-ring GroEL do not appear able to exchange substrate across the equatorial plane. The total volume of each cavity is measured in the crystal structure to be ~85,000 $Å^3$, just large enough for a native protein with a molecular size of 70 kDa, assuming a perfect fit. The size of a more loosely packed, nonnative polypeptide that could fit completely inside the cavity would be much smaller. Nevertheless, because the channel is open at this stage of binding, polypeptides can protrude slightly

Figure 3 Model of the pathway of chaperonin-mediated protein folding. Phase I is not shown. *Panel A* represents phase II; *panel B*, phase III; *panel C*, phase IV; the *final panel* reflects current uncertainty about the details of the pathway by which the complex relaxes and recycles into a new peptide-binding state. The orientation of the intermediate and apical domains of the bottom ATP-bound ring simply reflects an ATP-driven change; neither the character nor the extent of this change has been established. See the text for details. GroES is *shaded gray*; folding polypeptide is the *cross-hatched circle*; misfolded or folded (or committed) polypeptide is shown as the *open circles*. The domains of GroEL are indicated by *rectangles* (equatorial), *small ovals* (intermediate), and *large ovals* (apical).



ATP and GroES binding to ring containing polypeptide creates a stable GroEL–GroES *cis* assembly, stripping off even the most tightly bound non–native polypeptide from apical domains.

Nucleotide is trapped; "complete" hydrolysis of ATP is assured.

All *cis* ATP is hydrolyzed, producing a weakened *cis* assembly that is "primed" for release of GroES.

Polypeptide continues to refold to a native state but misfolding can occur.

ATP binding to lower (*trans*) ring and causes (see D) discharge of GroES, product polypeptide, and ADP from the weakened *cis* ADP assembly.

from the cavity, as cryo-EM and small-angle neutron-scattering studies have shown.

Nine residues (located on helices H and I and a loop between strands 6 and 7) of the apical domains have been identified by mutagenesis as important for polypeptide binding (41). These residues, eight of which have nonpolar side chains (the ninth is serine), face the central cavity in the unliganded GroEL structure (33, 41), closely matching the position of bound polypeptides seen in EM images of GroEL-substrate binary complexes. Thus, in the polypeptide acceptor state, GroEL presents a ring of hydrophobic binding surface along the inner edge of its apical cavities, poised to interact with the excess hydrophobic surface typically presented by folding intermediates. This strong hydrophobic contribution to the binding of substrate polypeptides by GroEL has been examined and confirmed with a number of substrates (42, 43), although polar and ionic contributions may also play a role (44, 45).

Despite these general characteristics of substrate interaction with GroEL, the exact stereochemistry of substrate binding has not been well characterized. Because GroEL evolved to interact with a wide variety of folding intermediates encompassing a large diversity of sequence information, its promiscuity would seem to preclude a specific and universal binding interaction. For example, both α-helical and extended secondary structures have been observed associated with chaperonin. In NMR studies, a rhodanese peptide formed an α-helical structure when bound to GroEL (46). On the other hand, the recent crystal structure of an isolated GroEL apical domain containing a 17-residue N-terminal tag showed a well-resolved interaction between this N-terminal segment, in an extended conformation, and the apical binding surface of a neighbor in the crystal lattice (47). Seven of the nine mutagenically implicated GroEL residues were found in contact with the peptide segment, as were several other hydrophobic residues, most of which are located on helices H and I. The binding mode of small isolated peptides, especially those presented by a neighboring molecule of a crystal lattice, may be different from those of nonnative substrate proteins because the binding of the latter to a ring of seven apical domains is subject to restraints that do not apply to small peptides. Thus, it is difficult to extrapolate from current structural information to a general model of GroEL-substrate interaction, if indeed one exists.

As pointed out above, the vast conformational space available to a folding polypeptide can lead to significant local energetic minima, outside the global (or near-global) minimum for which the protein is searching. These kinetic traps can represent off-pathway misarrangements of the individual chain (misfolding) or potentially irreversible nonproductive interactions between chains (aggregation). Because GroEL must interact with individual folding intermediates and productively promote movement along a folding trajectory toward the

native state, one or more aspects of the interaction must be capable of driving the substrate polypeptide out of these kinetic traps. One way for GroEL to accomplish this goal is through unfolding, either globally or locally, the polypeptide that it binds (48). The potentially cooperative nature of binding a folding intermediate to the multiple, neighboring apical sites of the GroEL toroid make this an especially attractive mechanism. Once bound, the polypeptide may be further unfolded by the subsequent elevating and twisting movements of the apical domains that occur on binding of nucleotide and GroES and ultimately lead to its release into the central cavity (49). Thus, the energy required for unfolding could come from either or both of two sources, namely the energy produced by the interaction between the polypeptide and the hydrophobic channel face and the energy generated by binding ATP and GroES to form the *cis* folding-active complex.

The coupling of substrate binding to substrate unfolding on GroEL has been examined in a number of studies, either by characterizing the point at which GroEL interacts with a polypeptide along its folding pathway or by examining rates of hydrogen-deuterium exchange of bound substrates. In general, it appears that in the absence of GroES and ATP, GroEL interacts rapidly (10^6–10^7 M^{-1} s^{-1}) with relatively early, collapsed folding intermediates (48; MS Goldberg, unpublished results). Moreover, GroEL is capable of shifting the natural equilibrium between the native and unfolded states of several proteins toward the unfolded state by stably binding only the unfolded protein (50–52). These results do not in themselves constitute a demonstration that binding causes unfolding, however. Evidence for direct participation of GroEL in polypeptide unfolding has been more mixed. Studies with a small protein, barnase, seem to indicate that global unfolding of this polypeptide occurs concomitant with binding (48). Similar, but less certain, conclusions were reached by an examination of cyclophilin bound to GroEL (53). On the other hand, deuterium exchange experiments with three-disulfide α-lactalbumin intermediates bound to GroEL showed a low degree of overall protection, consistent with the retention of some secondary structure (54). Moreover, a pair of recent studies with dihydrofolate reductase (DHFR), using NMR analysis of hydrogen-deuterium exchange (55) and mass spectrometry (56), could find no evidence for large-scale global unfolding of the stably bound DHFR intermediates. Indeed, the protection afforded against hydrogen-deuterium exchange was consistent with a highly structured and native-like, albeit unstable, form of DHFR bound to GroEL.

Active unfolding of a substrate protein by GroEL needs to be invoked only if the kinetic traps are operating at the level of misfolding within individual polypeptide chains. The alternative off-pathway reaction that can prevent the acquisition of the native state is aggregation, which may be dominant in some

instances. Strong binding by GroEL of aggregation-prone folding intermediates would be sufficient to block this nonproductive reaction. Furthermore, provided there was timely and efficient release of the bound polypeptide, the overall result of this mechanism would be to catalyze the production of the native state without directly affecting the unfolding of the polypeptide itself. Ranson and colleagues have shown that the folding of mitochondrial malate dehydrogenase by GroEL involves such a mechanism (57). Here, even at stoichiometries as low as 1:20 relative to substrate, GroEL can catalyze the acquisition of the native state by increasing the proportion of monomeric folding intermediates available to the normal folding pathway. It is important to point out, however, that it is not known to what extent aggregation or intramolecular misfolding may contribute to the off-pathway fate of any substrate of GroEL, particularly in the context of the intact cell. Indeed, it seems likely that both mechanisms may be operative at different points on the folding pathway, and it may be a criticial feature of chaperonin-mediated folding that both can be dealt with by the same machinery.

Phase II: Nucleotide and GroES Binding

Since the first in vitro GroEL-assisted refolding experiments were performed, it has been noted that whereas GroEL alone inhibits refolding, GroEL in the presence of K^+ ions, Mg-ATP, and GroES promotes efficient folding to the native state (37, 58–61). This led investigators to propose the existence of at least two distinct conformations of the complex: one that binds unfolded polypeptides tightly and ATP weakly and another in which the binding properties are reversed (60, 62), suggesting that nucleotide-modulated conformational changes of the chaperonin are inherent to the protein folding cycle. Initial equilibrium measurements of ATP binding and hydrolysis by GroEL revealed a high degree of cooperativity, which was enhanced further upon the addition of GroES (63–66). Moreover, GroES binding depends on the binding of nucleotide. The order of these events and their rates were studied by following changes in the fluorescence of pyrene-labeled GroEL. Weak binding of ATP to GroEL triggers a rapid conformational change (half-time ~40 ms) (65), which precedes the fast association of GroES ($> 4 \times 10^7\ M^{-1}\ s^{-1}$) (67). Conformational changes have been observed directly upon the addition of ATP, using cryo-EM and three-dimensional single-particle reconstruction. They have been interpreted, assuming the en bloc domain movements demonstrated by the crystal structure, to show that the apical domains of one ring open out by about 5–10° relative to the equatorial axis and also twist, causing a slight elongation of the GroEL cylinder (25, 68). The relative effect of the various nucleotides upon the apical domain movements followed the order ATP > AMP-PNP > ADP, which is consistent with the molecular mechanism outlined below.

The allosteric behavior of GroEL with respect to its ATPase function does not fit well into a simple model of cooperativity. Rather, it has been quantitatively described in a nested model that combines positive allostery within each ring with a negative effect between the rings (69). This model provides a context within which to appreciate the molecular details of the ligand-induced conformational changes that underlie the allosteric effects, as revealed by the crystal structures (see below), and to rationalize the effects of various mutational changes that independently disrupt one or the other allosteric interaction. However, the remarkable structure/function relationships of the folding cycle are difficult to completely capture with the compact formalisms originally developed to describe the behavior of hemoglobin and the regulation of oligomeric enzymes. For example, GroEL-bound ATP is hydrolyzed on only one ring at a time, even in the presence of excess ATP (67), whereas the binding of seven ADP molecules to one ring of GroEL is virtually noncooperative and only marginally asymmetric between rings. In the presence of GroES, however, ADP binding is complete and essentially irreversible within one ring and nonexistent in the opposite one. It would appear that there is inherent positive and negative cooperativity within the GroEL tetradecamer that is driven by GroES into the discrete intermediate states that characterize the stages of the folding cycle.

The structure of nucleotide in its binding site at the top of the equatorial domain, facing the central cavity, was first observed in the crystal structure of a variant GroEL (R13G/A126V) fully complexed with 14 ATPγS molecules, solved to a resolution of 2.4 Å (70) (Figure 4*A*). The site contains residues 87–91 (DGTTT), which interact with the β- and γ-phosphates of ATP and lie in a loop region (between helices C and D) that is highly conserved among chaperonins (71). Two metal ions are present in the nucleotide binding site. One is a magnesium ion, which chelates an oxygen from each of the three nucleotide phosphates, the carboxylate of Asp87, and two water molecules. Surprisingly, the overall architecture of the ATPγS complex was largely superimposable over the X-ray structure of unliganded GroEL (33, 72), in contrast to the significant domain movements seen in the cryo-EM image reconstructions. It was suggested that the discrepancy between the cryo-EM and crystal structures might be due in part to a decreased negative cooperativity observed for the variant protein (73), allowing the binding of ATPγS to all 14 subunits and creating a highly symmetrical molecule that is able to form a well-defined crystal lattice. This appears not to be the case, however, because further crystallographic studies with wild-type GroEL and ATPγS also reveal an isomorphous structure (D Boisvert, unpublished observations), leaving open the possibility that lattice forces have prevailed over movements of a flexible apical domain. In contrast to the crystalline specimens, which demand a uniform state, the noncrystalline complexes in the cryo-EM studies are not subject to such constraints on

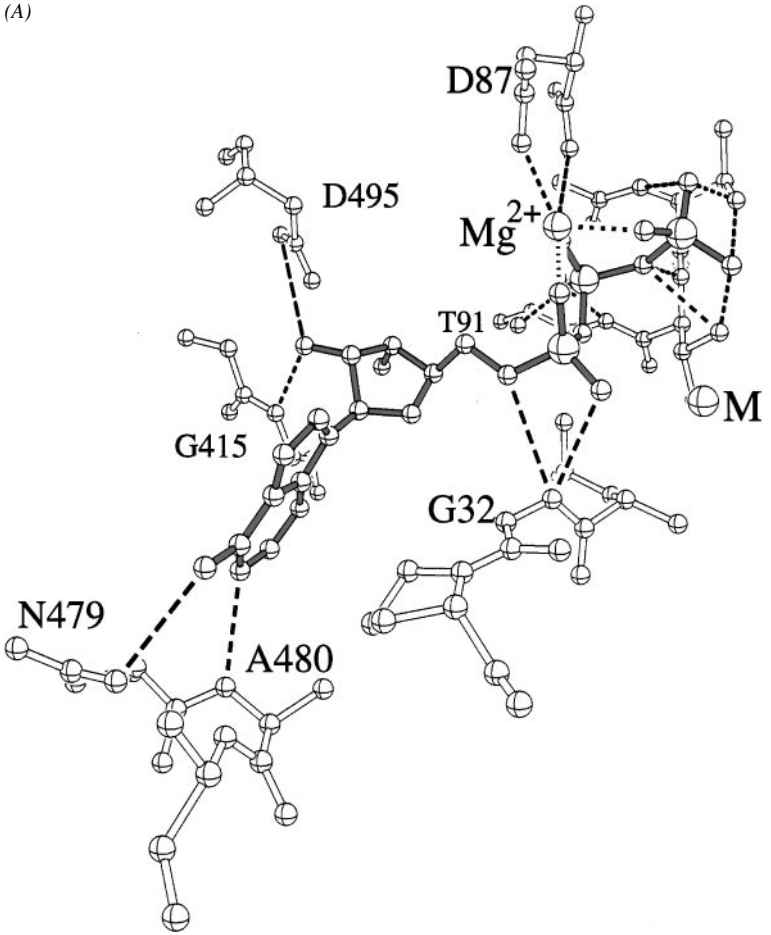


Figure 4 Skeletal representations of the nucleotide binding sites. The view is similar to that in Figure 2. The protein has *unshaded* bonds, and the nucleotide has *gray* bonds. The metal ions are *large spheres*. (*A*) ATPγS bound to GroEL in the GroEL-(ATPγS)$_{14}$ structure. (*B*) ADP bound to GroEL in the GroEL-GroES-(ADP)$_7$ structure. Two residues from the M helix of the intermediate domain have *black* bonds.

symmetry. Nonetheless, two significant local substructure shifts were observed in the ATPγS crystal structure: a large axial translation of helix C and a movement of a stem loop (Lys34 to Asp52), whose antiparallel stem forms an essential parallel β-contact with the neighboring subunit within the ring through a β-strand near the C terminus (strand 19) (Figure 5*A*). The importance of these two substructure movements is now evident in light of the recent GroEL-GroES-(ADP)$_7$ complex structure (74).

(B)

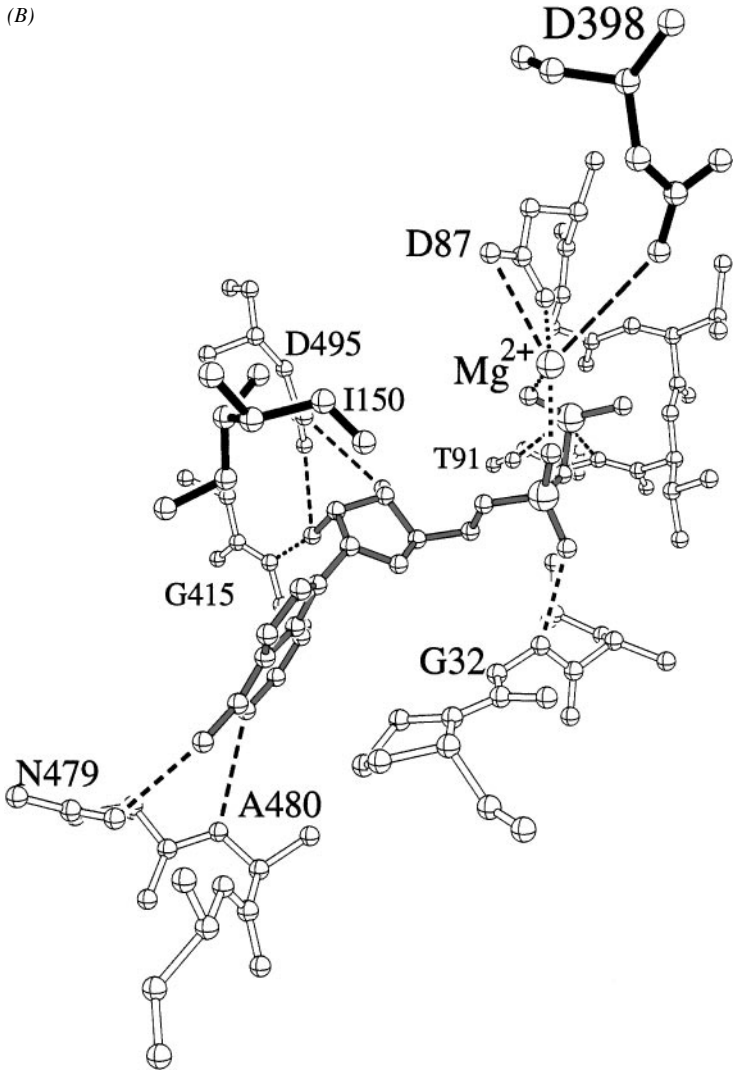


Figure 4 (*Continued*)

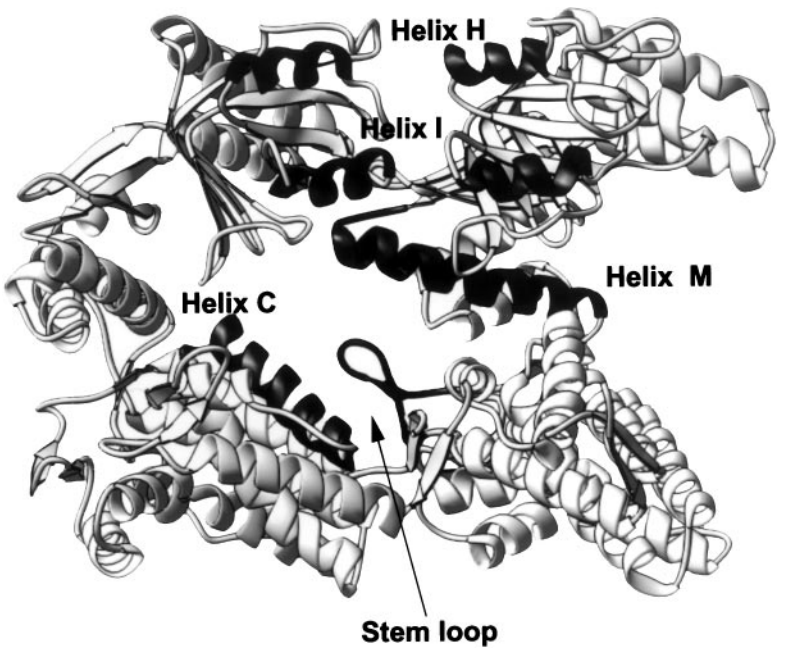


Figure 5 Substructures involved in stabilizing the *cis* assembly. Ribbon drawings for two adjacent subunits in *trans* (*A*) and *cis* (*B*) GroEL rings of the GroEL-GroES-$(ADP)_7$ complex, viewed from the inside of the ring. Both are oriented similarly with respect to the equatorial interface (at the bottom). The substructures that form the new interface between equatorial and intermediate domains (helix C, helix M, and the stem loop), as well as those that form the GroEL-GroES interface (helix H and helix I), are *dark gray*. Helix C and the stem loop are shifted by nucleotide to provide new stabilizing contacts for the reoriented helix M, as shown in *panel B*.

The Asymmetrical GroEL-GroES-ADP Complex Structure

OVERALL STRUCTURE AND DOMAIN MOVEMENTS The overall structure of the GroEL-GroES-$(ADP)_7$ asymmetrical complex is the expected bullet-shaped image (Figure 1, opposite p. 586), resulting from the smoothness of the union between GroEL and GroES (Figure 5*B*). GroES caps one end of GroEL, which is elongated and tapered toward the GroEL-GroES interface. The change in the shape of GroEL is due mostly to the change in the *cis* GroEL ring. Large en bloc movements of the apical and intermediate domains in the *cis* ring widen and elongate the *cis* cavity. The GroES ring, assembled as in its stand-alone structure, caps the apical surface of the *cis* ring, anchoring the elevated orientation of the apical domains and closing off the end of the cavity. The net result is a dome-shaped chamber that has the elevated apical domains as its

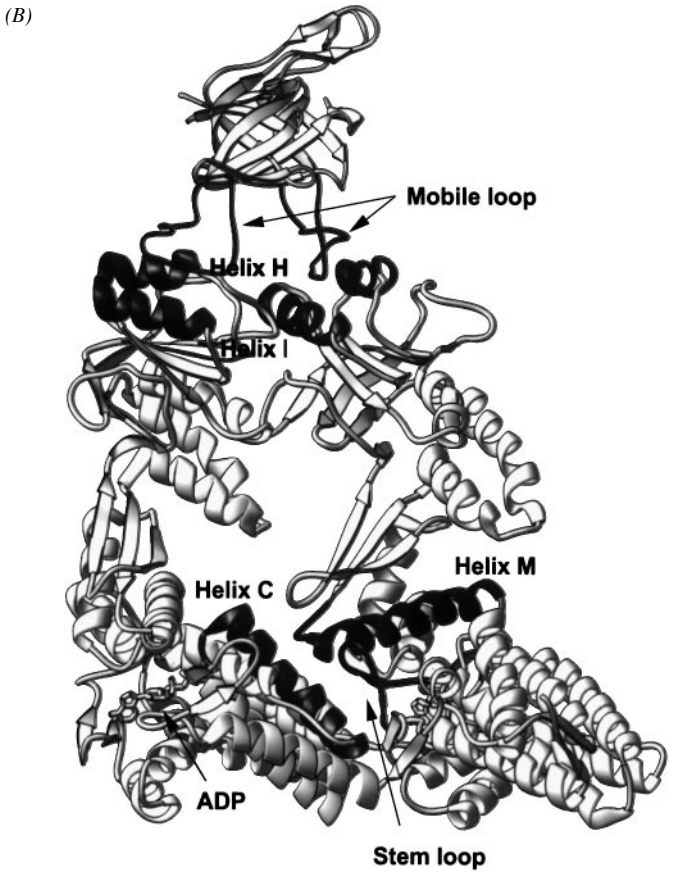


Figure 5 (*Continued*)

walls and the GroES cap as its roof. The GroEL and GroES rings share one nearly exact sevenfold rotational axis (Figure 1, opposite p. 586). In contrast to the dramatic changes in the *cis* ring, the *trans* ring (the empty ring) closely resembles that of unliganded GroEL.

The dramatic reshaping of the *cis* ring is due to rearrangements involving both intermediate and apical domains (Figures 4, 5, and 6). First, the intermediate domain swings down toward the equatorial domain and the central channel, pivoting approximately 25° around Pro137 and Gly410. The movement closes the occupied nucleotide binding site, located on the top inner surface of the equatorial domain, and generates numerous new interactions with the bound

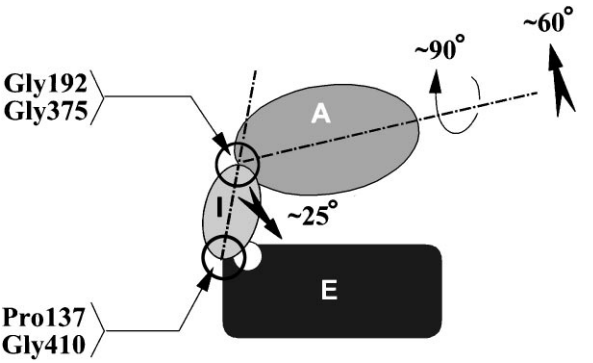


Figure 6 A schematic drawing showing direction and magnitude of the domain movement within the *cis* ring. The *shading* corresponds to the three GroEL domains: equatorial, *dark gray*; intermediate, *white*; and apical, *medium gray*. The *small disc* on the top of the equatorial domain represents the nucleotide binding site.

nucleotide and the equatorial domain. Second, the apical domain swings up 60° relative to the equator and twists around the long axis of the domain by about 90°, forming new interfaces with neighboring apical domains and leading to an interaction with the mobile loop of GroES. Both intermediate and apical domain movements are largely en bloc. The *cis* equatorial domains also show an en bloc movement that is small in magnitude compared to those of the intermediate and apical domain but every bit as important. The equatorial domains of the *cis* assembly tilt inward toward the cylindrical axis by 4° (Figure 9). Because the strong interface between the rings is conserved, there is a complementary outward tilt of the equatorial domains in the *trans* ring. This imposes a further asymmetry between the two rings and underlies the negative allostery that relates them (see below).

NUCLEOTIDE SITE The asymmetry of the GroEL-GroES-$(ADP)_7$ complex is most obvious in its ligand binding: GroES binds to one end of GroEL, and the seven *cis* nucleotide binding sites are fully occupied, while the seven *trans* sites are completely empty. Except for the absence of a second metal ion–mediated interaction with the α-phosphate of ATPγS, the specific interactions of ADP with the equatorial domain largely mirror those in the GroEL-ATPγS structure (70) (Figure 4*B*). In the GroEL-ATPγS structure (as well as in unliganded GroEL), the nucleotide binding site is largely open (compare panels *A* and *B* in both Figure 2 and 5), so the nucleotide can enter and exit without much steric hindrance. Upon GroES binding, however, residues of the helices F and M of the reoriented intermediate domain clamp onto the nucleotide, the Mg^{2+} cofactor, and residues of the equatorial domain, thereby closing the nucleotide

binding site. Thus, nucleotide is trapped in the *cis* ring and will remain there until the *cis* complex is disassembled, even at nucleotide concentrations low enough to completely empty the *trans* ring binding sites.

Some of the interactions clarify previously unexplained observations. For example, Ile150 of helix F forms a van der Waals interaction with the sugar moiety of the ADP, which is probably why mutation of Ile150 is lethal (41). Helix M contributes the carboxylate oxygen of Asp398 to the Mg^{2+} ion coordination cage, explaining why the D398A mutant GroEL retains only 2% of the wild-type ATPase activity, even though its affinity for ATP is unaffected (40). Furthermore, if Asp398 is prevented from assuming this new active-site position through restriction of domain movements by covalent cross-linking, GroEL is unable to hydrolyze bound ATP (75). These observations reaffirm the conclusion drawn from the fully liganded ATPγS-bound structure; namely, that binding of ATP to GroEL does not require shifts of the intermediate or apical domains but that subsequent hydrolysis of ATP does.

GroES BINDING The binding of ATP or ADP supports the binding of GroES (16), albeit with different rates of GroES association to the different GroEL-nucleotide binary complexes. For example, GroES binds rapidly ($> 4 \times 10^7$ M^{-1} s^{-1}) to GroEL-$(ATP)_7$ after the ATP-induced conformational change (67), whereas GroES associates more slowly (1×10^5 M^{-1} s^{-1}) with the GroEL-$(ADP)_7$ state (65). The stereochemical bases for the interdependence of GroES and nucleotide binding are threefold. First, the equatorial domains of the *cis* ring of the GroEL-GroES-$(ADP)_7$ complex show the same nucleotide-induced substructural shifts seen in the fully saturated ATPγS structure: an axial shift of helix C and displacement of the Lys34-Asp52 stem loop. These shifts provide stabilizing contacts for the radically reoriented intermediate domain (Figure 5). Second, the Mg^{2+} complex of the bound nucleotide provides additional direct stabilizing contacts for the reoriented intermediate domain. These interactions are the same as those that trap nucleotide in the *cis* assembly. Third, the nucleotide-shifted intermediate domain has now repositioned the hinge connecting it to the apical domain, so that intermediate domain/apical domain contacts between subunits of the unliganded structure are disrupted and the ensuing stabilizing interactions with GroES are sterically feasible. Thus, GroES binding is enabled by the structural transition initiated by nucleotide binding.

The extent of the ligand-induced conformational changes follows the order ATP > AMPPNP > ADP, implying highly stereo-explicit interactions with the β-γ phosphoanhydride of the triphosphate moiety. The same order of functional effectiveness is also observed in the apparent rates and affinities of GroES binding to the different GroEL-nucleotide complexes and in the increased stability of the GroEL-GroES-$(ATP)_7$ complex over its ADP counterpart (40), suggesting a similar dependence on the terminal diphosphate. However, the specific

interactions that account for this ranking cannot be derived with confidence from the current structures.

The movement of Ile150 and Asp398 into the ATPase active site is locked into place by the binding of GroES, thereby sequestering the nucleotide and precluding the loss of ATP (or the hydrolysis product ADP) or its exchange with free ligand. Because of its matching sevenfold rotational symmetry, GroES imposes the same restriction on all subunits of the *cis* ring simultaneously. Therefore, only in the presence of GroES are the seven bound ATP nucleotides of the *cis* ring committed to hydrolysis as a unit. Thus, GroES markedly increases the positive cooperativity of ATP hydrolysis within one ring, producing the quantized behavior of the ATPase activity of GroEL, as described by Viitanen and coworkers (76) and quantitated by Burston and colleagues (67).

In contrast to the major structural changes occurring in GroEL, the GroES heptamer ring in the GroEL-GroES complex is similar to that in the stand-alone structure (34) except for the mobile-loop residues, which now become structured. As expected, the mobile loop forms the interface with the elevated and twisted H and I helices of the GroEL apical domain through small aliphatic side chains, including Ile25, Val26, and Leu27 of GroES.

Several observations indicate that certain aspects of the role of the cochaperonin in the folding reaction remain incompletely defined. First, archaeal and eukaryal cytosolic chaperonins (TF55 and CCT, respectively) do not have a cochaperonin. Instead the thermosome appears to have an extended loop in its apical domain that serves to cap the *cis* folding chamber (76a). Perhaps the need to assist the folding of multidomain substrate proteins requires a less restricted apical hole. Second, mutations in the hinge regions of GroEL (V174F, V190I, and G375S) suppress the disruption of GroES binding caused by certain mutations in the GroES mobile loop (77). Current crystal structures do not offer a straightforward explanation for this observation. Third, GroES binding appears to limit the size of polypeptides whose folding can be assisted by the GroEL-GroES complex. For example, although T4 bacteriophage utilizes the host chaperonin, it encodes its own version of cochaperonin, known as gp31. The recent crystal structure of gp31 (78) suggests that it can form a larger folding cavity in the *cis* complex (Anfinsen cage) to accommodate the >50 kDa phage head protein (gp23), which may not fit comfortably under GroES.

Phase III: Polypeptide Release and Folding

The crystal structure of the GroEL-GroES complex suggests how the substrate polypeptide can be stripped from its binding sites on the channel walls and released into the cavity of the *cis* assembly (Figure 7, opposite pp. 600 and 601). Helices H and I of the apical domain, bearing peptide-binding residues Leu234, Leu237, Val259, Leu263, and Val264, move to the very top of the GroEL

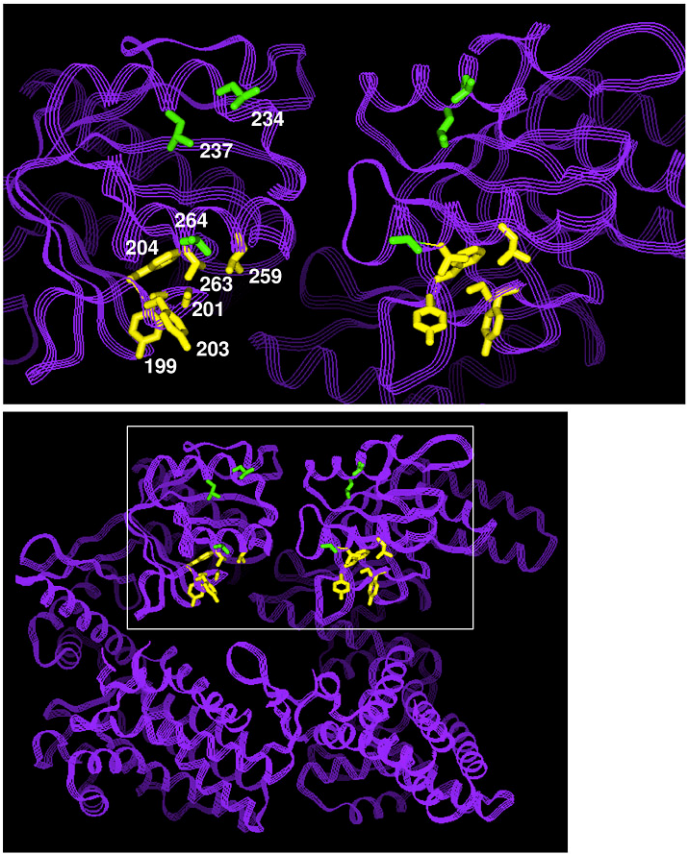


Figure 7A The change in the central cavity. Coiled-line ribbon drawing of two neighboring subunits of the *trans* ring viewed from the central cavity, oriented with the equatorial plane at the bottom; a magnified view of the polypeptide-interacting region (*rectangular area*) is shown at the top of the panel. Skeletal side chains denote residues involved in polypeptide binding, derived from mutagenesis studies; these residues, with the exception of S201, have hydrophobic side chains.

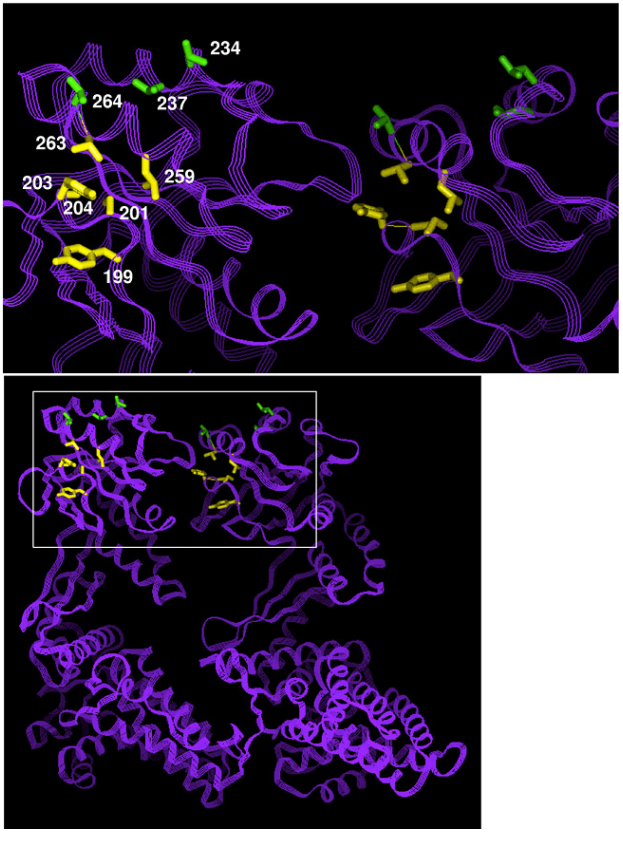


Figure 7B Two neighboring subunits of the *cis* ring, viewed in the same orientation and highlighted as in Figure 7*A*. The polypeptide-binding residues have moved away from the cavity surface, either to form the GroES interface [L234, L237, V264 (*green*)] or to contribute to the new interfaces between the apical domains of the *cis* GroEL ring [Y199, S201, Y203, F204, L259, V263 (*yellow*)].

cavity to form part of the GroEL-GroES interface. The loop between strands 6 and 7, bearing peptide-binding residues Tyr199, Ser201, Phe203, and Tyr204, is elevated and rotated into the *cis* ring's newly formed interface between the reoriented apical domains. Thus, the binding of GroES and nucleotide deprives the substrate polypeptide of its binding elements, which are now involved either in binding GroES directly (helices H and I) or in supporting GroES binding indirectly by stabilizing the interface between elevated and rotated apical domains. That the nine polypeptide-binding residues support the folding-active *cis* assembly explains the observation that a drastic mutagenic change of just one residue can prevent GroES binding and, indeed, be lethal to the mutant strain (41). As the apical domains move into their new positions, the corresponding peptide binding elements on the apical domains separate, possibly putting the substrate "on the rack" and helping to unfold it, either locally or globally (49). This might be sufficient to pull a bound polypeptide out of a conformation that is in a kinetic folding trap and place it back on the folding landscape just prior to or concomitant with its release into the cavity. Hydrophobic residues, which originally bound the nonnative polypeptide (presumably through hydrophobic interaction) in the cavity of the otherwise unliganded ring, are now buried in the *cis* ring assembly and have been replaced on the cavity walls with mostly polar residues [Figure 8 (opposite p. 602), compare *cis* and *trans* rings]. The released polypeptide is now free to reinitiate folding as an isolated molecule in a much-enlarged cavity whose hydrophilic lining is conducive to burial of hydrophobic residues and folding into a native structure.

Phase IV: Protein Folding and Release of Ligands

COMMUNICATION TO THE *TRANS* RING AND NEGATIVE COOPERATIVITY Superimposition of the equatorial domains of the *cis* GroEL ring on those of unliganded GroEL shows that the plane of the *cis* ring is slightly deformed (Figure 9). In the *cis* ring, each subunit tilts about 4° toward the cylinder axis, so that the inside of the ring is 3 Å lower than the original plane and the outside is 5 Å higher. Some of the largest shifts are observed for residues that are involved in cross-ring interactions: For example, the Cα of Glu434 moves 4.9 Å and the Cα of Ala109 moves 3.8 Å away from equitorial plane. Despite these shifts, the chemical details of the interface are maintained. To preserve the inter-ring interface, the *trans* ring must shift in a complementary direction. This causes each *trans* subunit to tilt in the opposite direction, that is, away from the central axis, by about 2°. Thus, the formation of the *cis* GroEL-GroES assembly favors a structural change in the opposite ring that is opposed to the formation of a second *cis* assembly (a symmetrical football). Binding events in the *cis* ring compete against similar events in the *trans* ring, explaining the transmission of negative allosteric effects across the equatorial plane. Like the positive effects

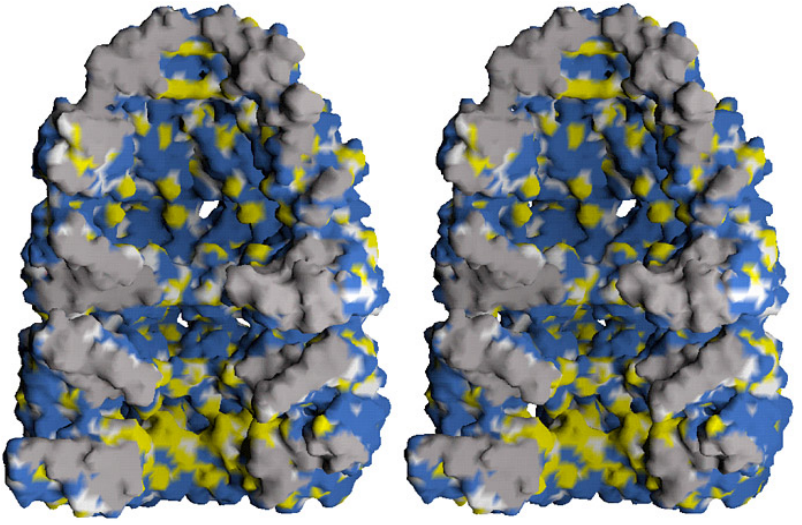


Figure 8 Stereo view of the solvent-accessible surface of the central cavity of the GroEL-GroES-$(ADP)_7$ complex. The three subunits from each of the rings nearest the reader were removed to show the inside of the assembly. Colors represent the type of surface: all backbone atoms, *white*; all hydrophobic side-chain atoms (A, V, L, I, M, F, P, Y), *yellow*; all polar and charged side-chain atoms (S, T, H, C, N, Q, K, R, D, E), *blue*; all solvent-excluded surface at subunit interfaces, *gray*. Most of the *yellow* hydrophobic patches on the surface of the *trans* GroEL cavity are replaced by *blue* polar patches on the surface of *cis* GroEL cavity.

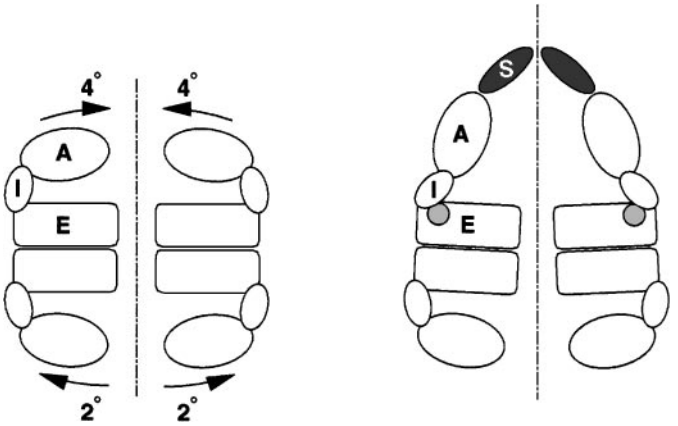


Figure 9 Negative cooperativity between rings. The presence of nucleotide (ATP or ADP) (*shaded circle*) and GroES causes the top of the *cis* equatorial domains to tilt inward by about 4°. Preservation of the equatorial interface forces the top of the *trans* equatorial domains to splay outward 2°, in a direction opposing the formation of a second GroES/nucleotide assembly on this ring. The *shading* corresponds to the three GroEL domains: equatorial, *dark gray*; intermediate, *white*; and apical, *medium gray*. S indicates GroES.

within a ring, this transmission is primarily through en bloc movements rather than through conformational shifts within the domains. Unlike the mechanisms in most other allosteric systems, however, this expression of negative allostery depends on the preservation, rather than alteration, of the quaternary interfacial contacts across the equatorial plane.

The winner in the competition between the two rings is decided by the type of adenine nucleotide it binds. Functional experiments show that ATP is dominant over ADP because it provides a stronger stabilizing force for the complex. First, binding of GroES with ATP, but not ADP, will release the most tightly bound nonnative polypeptides into the domed folding cavity of the *cis* ring, permitting them to initiate and, if sufficient time elapses, complete folding to a native protein (40). Second, bound ATP, but not ADP, causes an ATPase-deficient assembly (containing the D398A mutation) to maintain its structural integrity when challenged by low temperature and 0.5 M guanidinium HCl (40). Finally, ATPase-deficient rings of D398A maintain their domed *cis* assembly and will not release nucleotide, GroES, and folded polypeptide upon exposure of the *trans* ring to ATP, unless the bound ATP in the *cis* ring is permitted to hydrolyze to ADP. As noted earlier, it appears that the β-γ diphosphate/Mg^{2+} portion of the nucleoside triphosphate contributes additional strong contacts beyond those available from ADP that ultimately stabilize the changes that underlie the formation of a domed *cis* assembly. Thus, the *cis* ring association

of GroES is weakened when the nucleotide loses its γ-phosphate by hydrolysis, and the complex becomes susceptible to disassembly (and release of GroES, polypeptide, and ADP) when subjected to the outward tilting stress imposed by the binding of ATP on the opposite ring (Figure 3). This model is fully consistent with the role of nucleotides in driving other molecular mechanical systems such as G-proteins and contractile systems. Moreover, it explains why the GroEL-assisted dynamic protein folding cycle requires a double-toroid structure.

RECYCLING CHAPERONIN AND NONNATIVE POLYPEPTIDE The timing of the formation and dissociation of the folding-active GroEL-GroES intermediates can be gauged from several experiments (at 23°C). Following binding of GroES and ATP to a GroEL-polypeptide binary complex to form a *cis* ATP complex, polypeptide is released into the central channel within a second, as revealed by fluorescence anisotropy measurements (39, 40). Interestingly, during this initial second, there is a sharp drop of anisotropy, reflecting increased mobility of the reporting groups and possibly the process of precipitous unfolding mentioned above, as the polypeptide is initially stretched upon the mobilizing apical binding sites. After polypeptide release into the channel, there is a period of ~6–8 s before ATP hydrolysis in the *cis* complex converts it to the primed ADP state (67). Productive folding appears to continue during this time and after formation of the *cis* ADP state, as evidenced both by the continuous anisotropy changes and by ongoing accretion of biological activity of the folding substrate (39, 40). After ~5–10 s in the *cis* ADP state, discharge of GroES and polypeptide is prompted by ATP binding in the *trans* ring (40). The dissolution of the complex appears likely to be a fast step, as suggested by recent dynamic fluorescence experiments (HS Rye, unpublished results), consistent with the quantized cooperativeness within a ring being coupled to negative cooperativeness between the rings, as described above. What likely follows for the former *cis* ring is that its apical domains relax and move back down to the conformation with high affinity for polypeptide, as seen in the unliganded GroEL structures. What happens to the former *trans* ring remains poorly understood. For example, it has not been clearly established whether nonnative polypeptide or GroES becomes bound on the *trans* ring during folding and/or ATP-triggered discharge of the *cis* complex. In a study with substrate bound in a *trans* ternary ADP complex, the addition of excess GroES and ATP was associated with nonproductive release instead of productive folding, suggesting that binding a second GroES is disfavored (WA Fenton, unpublished data). Thus, although it is attractive to consider that the chaperonin machine could simultaneously operate two different productive reactions within opposite rings, shifted 180° in the cycle from each other, this may not be its working mode. It remains possible that polypeptide and GroES must both be completely

cleared from GroEL before a next round of productive binding (to rings in a fully relaxed state) can ensue. In this regard, a study with a mixed-ring GroEL assembly, able to bind polypeptide and GroES on only one of its rings, exhibited kinetics of refolding of rhodanese that were identical to those of refolding by wild-type GroEL (79). This supports the idea that, at least for rhodanese, GroEL progresses through one entire cycle before commencing the next; otherwise, one would have expected the kinetics of the wild-type chaperonin, with two competent cavities, to be faster. Further studies may resolve how the machine cycles between active states.

Given the schedule of events on GroEL dictated by ATP binding and hydrolysis, it becomes clear that nonnative proteins encapsulated in the *cis* ternary complex have a set amount of time, ~15 s, in which to reach the native state prior to disruption of the folding-active chaperonin complex. It appears that only a certain fraction of bound, nonnative polypeptide molecules can reach the native state or a conformation that is committed to the native state (that is, no longer recognizable by GroEL) during this time. But the remaining fraction fails to reach native form and requires either another round of interaction with GroEL or, in the cell, interaction with other chaperones or with protease. The latter possibility seems particularly important for removal of damaged proteins to prevent the clogging of the chaperone machinery. This nonnative fraction appears in large part to be discharged from GroEL, as revealed by a number of experiments, in particular those using a "trap" molecule—a mutant or modified form of GroEL that can bind but not release nonnative proteins. Addition of such molecules to a folding reaction has reproducibly revealed a rate of release of rhodanese, for example, that is approximately 10 times its rate of refolding. Thus, a substantial portion of the nonnative forms, if not all, are discharged from chaperonin during a round of dissociation of the *cis* complex (76, 79–83). The question has been raised as to whether rebinding of such forms occurs in a conformation that has progressed toward the native state or whether this is an all-or-none process, with each round of folding starting from the same or similar nonnative conformation. The evidence at hand would argue for the latter, as the conformation of nonnative protein bound after transfer to other GroEL molecules appears to correspond to that initially bound to chaperonin (56, 80).

CONCLUSIONS AND PROSPECTS

Certainty Versus Uncertainty in the Stereochemistry of the Folding Cycle

We now understand in molecular detail how the binding of nucleotides potentiates the large conformational changes in the *cis* ring that lead to the release

of bound nonnative polypeptide into the shielded folding chamber of the *cis* assembly. We can also speculate with some confidence on how these large domain shifts might further unfold the polypeptide prior to or just at the time of its release. We also know that the mechanism of this release is coupled to the stabilization of the *cis* complex, as residues of the apical domain implicated in peptide binding form interfaces that support both the walls and the roof of the *cis* assembly. Bound GroES, acting as a multivalent keystone of the dome, stabilizes these enormous shifts. In doing so, it not only holds all the nucleotides of the *cis* complex in place but also maintains the catalytic machinery necessary to hydrolyze ATP effectively. This latter property may exemplify the general theme of extreme induced fit, wherein large domain shifts, triggered by nucleoside triphosphate binding, are stabilized by specific interactions with partner proteins that promote hydrolysis and productive return of the nucleotide-driven cycle to its ground state. Here, the partner protein is GroES; in the myosin headpiece, it is actin; and in G proteins, they are GTPase activating proteins (GAPs) or regulators of G-protein signaling (RGS) proteins. We also appreciate that for the cycle to proceed with productive release of polypeptide, along with GroES and ADP, ATP hydrolysis in *cis* (priming disruption) must be followed by ATP binding in *trans* (triggering disassembly). This mechanism underlies the need for a double toroid for efficient folding.

Although we understand the cooperative, quantized action of the *cis* system in convincing stereochemical detail, the mechanism of negative cooperativity between rings is less well understood. The inward tilting of the *cis* equatorial domains, coupled to the preservation of the interface between rings, causes outward tilting in the *trans* ring. This presents an attractive mechanism for negative cooperativity, in which formation of a *cis* complex on one ring opposes a similar event on the opposite ring. Thus, ATP binding and generation of a *cis* assembly on what was originally the *trans* ring could lead to the disruption of the original, but now weakened, ADP-primed *cis* complex, with productive release of GroES, polypeptide, and ADP. However, this opposing-tilt mechanism is derived from a structure in which bound GroES stabilizes a fully developed *cis* assembly, and there is no evidence that GroES is required, in addition to bound ATP, to discharge the products sequestered in the opposing ring. Indeed, a variety of experiments indicate that ATP binding alone will suffice. It will be necessary to visualize—in the absence of GroES—the mechanism by which ATP binding to one ring disrupts a *cis* assembly on the other before these questions can be answered.

The puzzle of the negative allosteric interplay between rings extends to the binding of nonnative polypeptides to otherwise unliganded GroEL. When presented with an excess of nonnative polypeptide, the chaperonin rarely binds polypeptide to both rings. Indeed, it is still not clear whether a nonnative

polypeptide substrate can bind to the empty *trans* ring during the interval when protected folding and ATP hydrolysis is taking place in the *cis* complex. Prudence would dictate that we retreat—at least until there are supporting data—from the tempting two-cylinder metaphor, in which the expulsion of products from one ring is coupled to the intake of reactants in the other. Rather, at this stage, we must restrict ourselves to the more conservative, albeit potentially less efficient, model shown in Figure 3.

While sorting out the exact sequence of events in the folding cycle, it will also be necessary to establish the molecular basis of the unique capacity of ATP to drive the cycle by binding to the catalytic center of the equatorial domain. This has been a central focus of research on the other nucleotide-driven molecular machines mentioned above. We have only inferential hints derived from the structures of the symmetrical GroEL$(ATP\gamma S)_{14}$ complex and the asymmetrical GroEL-GroES-$(ADP)_7$ one. Direct visualization of an ATP complex may be necessary to address this question.

A special feature of chaperonins is the wide range of nonnative folding intermediates that can be captured in the central cavity, even though specificity is apparent, at least in vivo. This raises several questions that require further study. To what extent are polypeptides bound by a unique surface of the apical domains, as exemplified in the binding of peptides to Hsp90 and the antigen presentation system? Are multiple alternative binding modes used to grasp nonnative polypeptides? And if so, to what extent and by what mechanism do they share common residues on the apical surface? What are the principal determinants of affinity in the polypeptide-GroEL interface? What fraction of subunits in a ring are needed to capture and stably bind a full-length nonnative polypeptide? Despite the analysis of folding dynamics from deuterium exchange and fluorescence studies, we still have only glimpses of the structural basis for these events.

Finally, can we extrapolate from the extensive studies of the GroEL-GroES folding cycle to the mechanisms of other chaperonin-assisted folding pathways? There are clear architectural differences between GroEL-GroES and the chaperonins of chloroplasts and, in particular, those of archaea and the eukaryotic cytosol that may require different mechanisms. This is especially true for events involving GroES, because there is no cochaperonin in archaea or the eukaryotic cytosol. On the other hand, sequence conservation in the nucleotide-binding segments of the equatorial domain suggests that the initial nucleotide-driven allosteric events in all chaperonins may be well represented by those seen in GroEL. The symmetry and subunit uniformity of GroEL and GroES, coupled to the amenity of *E. coli* to mutagenesis and recombinant technologies, have made the GroE system the ideal starting point for these inquiries into structure/function relationships in chaperonins. Even if we were to understand the

GroE system in detail, we would still be only near the end of the introductory chapter of the long and essential story of the final steps in the accurate expression of the genetic code.

ACKNOWLEDGMENTS

We thank Dr. Patrick Fleming for invaluable assistance with the illustrations.

Literature Cited

1. Dill KA, Chan HS. 1997. *Nat. Struct. Biol.* 4:10–19
2. Flynn GC, Chappell TG, Rothman JE. 1989. *Science* 245:385–90
3. Zhu XT, Zhao X, Burkholder WF, Gragerov A, Ogata CM, et al. 1996. *Science* 272:1606–14
4. Rudiger S, Germeroth L, Schneider-Mergener J, Bukau B. 1997. *EMBO J.* 16: 1501–7
5. Rudiger S, Buchberger A, Bukau B. 1997. *Nat. Struct. Biol.* 4:342–49
6. Ehrnsperger M, Gräber S, Gaestel M, Buchner J. 1997. *EMBO J.* 16:221–29
7. Lee GJ, Roseman AM, Saibil HR, Vierling E. 1997. *EMBO J.* 16:659–71
8. Sullivan W, Stensgard B, Caucutt G, Bartha B, McMahon N, et al. 1997. *J. Biol. Chem.* 272:8007–12
9. Bohen SP, Kralli A, Yamamoto KR. 1995. *Science* 268:1303–4
10. Fenton WA, Horwich AL. 1997. *Protein Sci.* 6:743–60
11. Ellis RJ, ed. 1996. *The Chaperonins.* San Diego, CA: Academic
12. Hartl F-U. 1996. *Nature* 381:571–80
13. Hendrix RW. 1979. *J. Mol. Biol.* 129:375–92
14. Hohn T, Hohn B, Engel A, Wurtz M, Smith PR. 1979. *J. Mol. Biol.* 129:359–73
15. Tilly K, Murialdo H, Georgopoulos C. 1981. *Proc. Natl. Acad. Sci. USA* 78: 1629–33
16. Chandrasekhar GN, Tilly K, Woolford C, Hendrix R, Georgopoulos C. 1986. *J. Biol. Chem.* 261:12414–19
17. Saibil H, Dong Z, Wood S, auf der Mauer A. 1991. *Nature* 353:25–26
18. Langer T, Pfeifer G, Martin J, Baumeister W, Hartl F-U. 1992. *EMBO J.* 11:4757–65
19. Ishii N, Taguchi H, Sumi M, Yoshida M. 1992. *FEBS Lett.* 299:169–74
20. Azem A, Kessel M, Goloubinoff P. 1994. *Science* 265:653–56
21. Schmidt M, Rutkat K, Rachel R, Pfeifer G, Jaenicke R, et al. 1994. *Science* 265: 656–59
22. Engel A, Hayer-Hartl MK, Goldie KN, Pfeifer G, Hegerl R, et al. 1995. *Science* 269:832–36
23. Azem A, Diamant S, Kessel M, Weiss C, Goloubinoff P. 1995. *Proc. Natl. Acad. Sci. USA* 92:12021–25
24. Braig K, Simon M, Furaya F, Hainfeld JF, Horwich AL. 1993. *Proc. Natl. Acad. Sci. USA* 90:3978–82
25. Chen S, Roseman AM, Hunter AS, Wood SP, Burston SG, et al. 1994. *Nature* 371: 261–64
26. Thiyagarajan P, Henderson SJ, Joachimiak A. 1996. *Structure* 4:79–88
27. Phipps BM, Hoffmann A, Stetter KO, Baumeister W. 1991. *EMBO J.* 10:1711–22
28. Trent JD, Nimmesgern E, Wall JS, Hartl F-U, Horwich AL. 1991. *Nature* 354:490–93
29. Liou AKF, Willison KR. 1997. *EMBO J.* 16:4311–16
30. Horwich AL, Willison KR. 1993. *Philos. Trans. R. Soc. London Ser. B* 339:313–26
31. Kubota H, Hynes G, Carne A, Ashworth A, Willison K. 1994. *Curr. Biol.* 4:89–99
32. Lewis SA, Tian G, Vainberg IE, Cowan NJ. 1996. *J. Cell Biol.* 132:1–4
33. Braig K, Otwinowski Z, Hegde R, Boisvert DC, Joachimiak A, et al. 1994. *Nature* 371:578–86
34. Hunt JF, Weaver AJ, Landry SJ, Gierasch L, Deisenhofer J. 1996. *Nature* 379:37–45
35. Mande SC, Mehra V, Bloom BR, Hol WGJ. 1996. *Science* 271:203–7
36. Landry SJ, Zeilstra-Ryalls J, Fayet O, Georgopoulos C, Gierasch LM. 1993. *Nature* 364:255–58

37. Goloubinoff P, Christeller JT, Gatenby AA, Lorimer GH. 1989. *Nature* 342:884–89
38. Saibil HR, Zheng D, Roseman AM, Hunter AS, Watson GMF, et al. 1993. *Curr. Biol.* 3:265–73
39. Weissman JS, Rye HS, Fenton WA, Beechem JM, Horwich AL. 1996. *Cell* 84:481–90
40. Rye HS, Burston SG, Fenton WA, Beechem JM, Xu Z, et al. 1997. *Nature* 388: 792–98
41. Fenton WA, Kashi Y, Furtak K, Horwich AL. 1994. *Nature* 371:614–19
42. Lin ZL, Schwarz FP, Eisenstein E. 1995. *J. Biol. Chem.* 270:1011–14
43. Itzhaki LS, Otzen DE, Fersht AR. 1995. *Biochemistry* 34:14581–87
44. Katsumata K, Okazaki A, Tsurupa GP, Kuwajima K. 1996. *J. Mol. Biol.* 264: 643–49
45. Perrett S, Zahn R, Sternberg G, Fersht AR. 1997. *J. Mol. Biol.* 269:892–901
46. Landry SJ, Gierasch LM. 1991. *Biochemistry* 30:7359–62
47. Buckle AM, Zahn R, Fersht AR. 1997. *Proc. Natl. Acad. Sci. USA* 94:3571–75
48. Zahn R, Perrett S, Stenberg G, Fersht AR. 1996. *Science* 271:642–45
49. Lorimer G. 1997. *Nature* 388:720–23
50. Zahn R, Pluckthun A. 1994. *J. Mol. Biol.* 242:165–74
51. Walter S, Lorimer GH, Schmid FX. 1996. *Proc. Natl. Acad. Sci. USA* 93:9425–30
52. Viitanen PV, Donaldson GK, Lorimer GH, Lubben TH, Gatenby AA. 1991. *Biochemistry* 30:9716–23
53. Zahn R, Spitzfaden C, Ottiger M, Wuthrich K, Pluckthun A. 1994. *Nature* 368: 261–65
54. Robinson CV, Groß M, Eyles SJ, Ewbank JJ, Mayhew M, et al. 1994. *Nature* 372:646–51
55. Goldberg MS, Zhang J, Sondek S, Matthews CR, Fox RO, Horwich AL. 1997. *Proc. Natl. Acad. Sci. USA* 94:1080–85
56. Groß M, Robinson CV, Mayhew M, Hartl F-U, Radford SE. 1996. *Protein Sci.* 5: 2506–13
57. Ranson NA, Dunster NJ, Burston SG, Clarke AR. 1995. *J. Mol. Biol.* 250:581–86
58. Laminet AA, Ziegelhoffer T, Georgopoulos C, Pluckthun A. 1990. *EMBO J.* 9: 2315–19
59. Viitanen PV, Lubben TH, Reed J, Goloubinoff P, O'Keefe DP, Lorimer GH. 1990. *Biochemistry* 29:5665–71
60. Martin J, Langer T, Boteva R, Schramel A, Horwich AL, Hartl F-U. 1991. *Nature* 352:36–42
61. Buchner J, Schmidt M, Fuchs M, Jaenicke R, Rudolph R, et al. 1991. *Biochemistry* 30:1586–91
62. Badcoe IG, Smith CJ, Wood S, Halsall DJ, Holbrook JJ, et al. 1991. *Biochemistry* 30:9195–9200
63. Gray TE, Fersht AR. 1991. *FEBS Lett.* 292:254–58
64. Bochkareva ES, Lissin NM, Flynn GC, Rothman JE, Girshovich AS. 1992. *J. Biol. Chem.* 267:6796–800
65. Jackson GS, Staniforth RA, Halsall DJ, Atkinson T, Holbrook JJ, et al. 1993. *Biochemistry* 32:2554–63
66. Todd MJ, Viitanen PV, Lorimer GH. 1993. *Biochemistry* 32:8560–67
67. Burston SG, Ranson NA, Clarke AR. 1995. *J. Mol. Biol.* 249:138–52
68. Roseman AM, Chen SX, White H, Braig K, Saibil HR. 1996. *Cell* 87:241–51
69. Yifrach O, Horovitz A. 1995. *Biochemistry* 34:9716–23
70. Boisvert DC, Wang JM, Otwinowski Z, Horwich AL, Sigler PB. 1996. *Nat. Struct. Biol.* 3:170–77
71. Kim S, Willison KR, Horwich AL. 1994. *Trends Biochem. Sci.* 19:543–48
72. Braig K, Adams PD, Brunger AT. 1995. *Nat. Struct. Biol.* 2:1083–94
73. Aharoni A, Horovitz A. 1996. *J. Mol. Biol.* 258:732–35
74. Xu Z, Horwich AL, Sigler PB. 1997. *Nature* 388:741–50
75. Murai N, Makino Y, Yoshida M. 1996. *J. Biol. Chem.* 271:28229–34
76. Todd MJ, Viitanen PV, Lorimer GH. 1994. *Science* 265:659–66
76a. Klumpp M, Baumeister W, Essen L-O. 1977. *Cell* 91:263–70
77. Zeilstra-Ryalls J, Fayet O, Georgopoulos C. 1996. *FASEB J.* 10:148–52
78. Hunt JF, van der Vies SM, Henry L, Deisenhofer J. 1997. *Cell* 90:361–71
79. Burston SG, Weissman JS, Farr GW, Fenton WA, Horwich AL. 1996. *Nature* 383:96–99
80. Weissman JS, Kashi Y, Fenton WA, Horwich AL. 1994. *Cell* 78:693–702
81. Smith KE, Fisher MT. 1995. *J. Biol. Chem.* 270:21517–23
82. Taguchi H, Yoshida M. 1995. *FEBS Lett.* 359:195–98
83. Mayhew M, da Silva ACR, Martin J, Erdjument-Bromage H, Tempst P, Hartl F-U. 1996. *Nature* 379:420–26

Annual Review of Biochemistry
Volume 67, 1998

CONTENTS