GroEL−GroES-Mediated Protein Folding

Arthur L. Horwich,*[,†,‡,#] George W. Farr,[†,‡] and Wayne A. Fenton[†,#]

Department of Genetics and Howard Hughes Medical Institute, Yale School of Medicine, Boyer Center, 295 Congress Avenue, New Haven, Connecticut 06510, and The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received December 14, 2004

Contents

1. Introduction

The chaperonin-mediated folding reaction is an essential ATP-dependent reaction that provides kinetic assistance to the process of protein folding to the native state in a variety of cellular compartments. This reaction, carried out by a megadalton-sized double ring "machine," remains a fascination because it exhibits a multitude of interesting features, for example, allostery, with both positive and negative cooperative behavior with respect to nucleotide; local hydrophobic interactions between the chaperonin and the non-native polypeptide; and productive protein folding inside an encapsulated "privileged" chamber. We have recently reviewed particular aspects of the reaction, ATP action, and fate of polypeptide, in considerable depth (see refs 1 and 2), and so elect here to provide the reader with a description of the mechanism as we currently understand it, followed by a more detailed consideration of a step that has been of some interest, namely, the transition between polypeptide binding in an open ring and productive protein folding in a cochaperonin-encapsulated one. By way of introduction, because many of our colleagues are unacquainted with how this area of study emerged, we provide a brief historical overview, which should provide a useful cell physiologic background to the mechanistic considerations that follow.

* To whom correspondence should be addressed. Phone: 203-737-4431. Fax 203-737-1761. E-mail: horwich@csb.yale.edu.
† Department of Genetics, Yale School of Medicine.
‡ Howard Hughes Medical Institute, Yale School of Medicine.
# The Scripps Research Institute.

2. Chaperonins – Establishment of a Role in Mediating Protein Folding in the Cell

One of the least expected chemical reactions to be uncovered in the living cell is the chaperonin-mediated folding of polypeptide chains to their native state through the consumption of ATP. The work of Anfinsen and his colleagues in the late 1950s and early 1960s established that the primary sequence of a polypeptide contains all of the information required to direct it to the native state, typically at the energetic minimum.[3] Any need for assistance to the protein folding process was thus not immediately obvious. Yet, observations followed that indicated that there were factors in the cell that influenced the proper assembly of oligomeric structures. In hindsight, we know that this is a reflection of the action of specialized proteins, molecular chaperones, that govern the folding of monomeric polypeptides that, in many cases, are the component subunits of assemblies. The step of oligomeric assembly itself, by contrast, generally occurs in an unassisted manner.

The first observation came in 1972 when two groups of investigators, Georgopoulos and Kaiser in Utah[4] and Takano and Kakefuda in Japan,[5] studying biogenesis of bacteriophages T4 and λ, reported a host cell gene, named *GroE*, that led to a block of virus infection when defective. Consistent with a block of biogenesis, they observed aggregates of phage heads inside the infected cells. The name *GroE* was derived from the fact that aggregation of λ phage heads could be overcome by second site mutations in the phage gene, λE, encoding the major capsid protein, that lowered the level of its product, preventing it from aggregating. In 1980, a component inside chloroplasts that formed a physical association with newly translated subunits of the CO_2-fixing enzyme rubisco was identified by Barraclough and Ellis.[6] The component, however, was not associated with the final oligomeric structure of rubisco, which is composed of both the chloroplast-synthesized (large) subunits and the imported

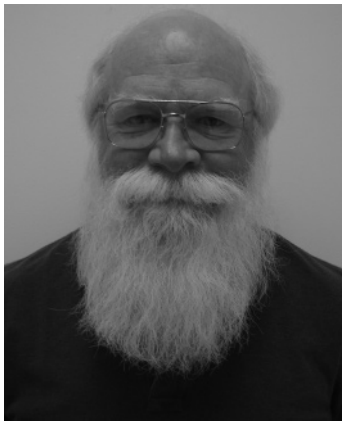
10.1021/cr040435v CCC: $59.00
Published on Web 04/20/2006

Arthur Horwich received A.B. and M.D. degrees from Brown University and trained in pediatric medicine at Yale University. He then pursued postdoctoral training in tumor virology at The Salk Institute with Walter Eckhart, studying polyoma T antigens, and in medical genetics at Yale University with Leon Rosenberg, studying targeting of a urea cycle enzyme to mitochondria. He subsequently joined the Yale faculty and is now Higgins Professor of Genetics, and Investigator in the Howard Hughes Medical Institute. His continuing interest in how proteins are trafficked to mitochondria led in 1987 to the discovery of Hsp60 as a folding machine in the mitochondrial matrix. A variety of methods were subsequently employed to examine how such chaperonins mediate protein folding. These focus most recently on the use of NMR methods, in collaboration with Kurt Wüthrich at The Scripps Research Institute, where Horwich is Visiting Investigator.

George W. Farr received a B.S. in chemistry from John Carroll University in 1985 and a Ph.D. in pharmacology from Case Western Reserve University in 1993 where he worked on tubulin biogenesis with Himan Sternlicht. He carried out postdoctoral work with Arthur Horwich, studying the mechanism of action of the eukaryotic cytosolic chaperonin (CCT) and currently holds a position at Yale University as Jr. Research Scientist. His research interests continue to focus on structure and mechanism in chaperonin-assisted protein folding.

(small) subunits. This implicated this so-called rubisco binding protein in the oligomeric assembly of rubisco.

In 1988, a relation between these two seemingly disconnected components in bacteria and chloroplasts came together when the sequences of the two genes were analyzed and compared by a collective of investigators including Georgopoulos and Ellis.[7] Both of the sequences were found to encode a 58 kDa protein, with greater than 60% amino acid identity along the length of the two primary structures. Moreover, it was clear in both cases that the subunit was part of a larger oligomeric assembly composed of two back-to-back rings, each with seven radially arranged subunits.[8−10] These components were termed chaperonins. Their functional role was the subject of some speculation, but an idea consistent with the functional studies that had gone before



Wayne A. Fenton studied biochemistry in the laboratory of Robert Abeles at Brandeis University (United States), where he received his Ph.D. in 1974 for research on the chemistry of vitamin B_{12} model reactions. His postdoctoral work with Dr. Leon Rosenberg in the Genetics department at Yale University School of Medicine concerned the biochemistry and physiology of vitamin B_{12} in mammalian cells, particularly involving inherited metabolic diseases related to coenzyme forms of this vitamin. He remained at Yale Medical School and currently is a Research Scientist in Genetics. He also holds the position of Senior Staff Scientist at the Scripps Research Institute. His current research interests focus on protein folding and unfolding in the cell, with particular emphasis on GroEL (folding) and ClpA/ClpP (unfolding).

was that they mediated the final folding and assembly of oligomeric proteins. This seemed entirely consistent with the work of Anfinsen, which appeared able to account for folding to the native or near-native state as a spontaneous process, followed by this assisted step of assembly.

Studies in mitochondria, however, published in 1989, challenged this model of chaperonin action in protein biogenesis.[11,12] In the field of mitochondrial biogenesis, it had been known for several years that, for cytosolically translated mitochondrial protein precursors to be posttranslationally imported across the mitochondrial membranes, an unfolded conformation was required. In particular, an early study of Eilers and Schatz[13] had shown that a fusion protein composed of a mitochondrial targeting peptide and dihydrofolate reductase (DHFR) could not be imported into isolated mitochondria in the presence of methotrexate, a ligand that stabilizes the DHFR moiety in a folded native form. By contrast, if the protein was first unfolded with denaturant and then incubated with isolated organelles, it was efficiently imported.[13]

In the face of this observation that precursors were translocated into mitochondria as unfolded monomeric chains, one could ask, did the imported chains fold spontaneously to their native forms at the other side of the membranes? Or did they require assistance to reach native form? This question was addressed in a genetic screen of a library of conditional lethal yeast mutants, inspecting for one in which a reporter protein precursor (of the urea cycle enzyme ornithine transcarbamylase, OTC) would be imported under nonpermissive conditions, would have its signal peptide proteolytically removed by a matrix processing enzyme, but would fail to reach its enzymatically active native form inside the matrix compartment.[11] Indeed, such a mutant was found, in which the OTC subunits were localized inside the matrix compartment in their mature size, but no enzymatic activity was detected. Endogenous yeast mitochondrial proteins were likewise affected − they could enter the organelles, their signal peptides were cleaved, but

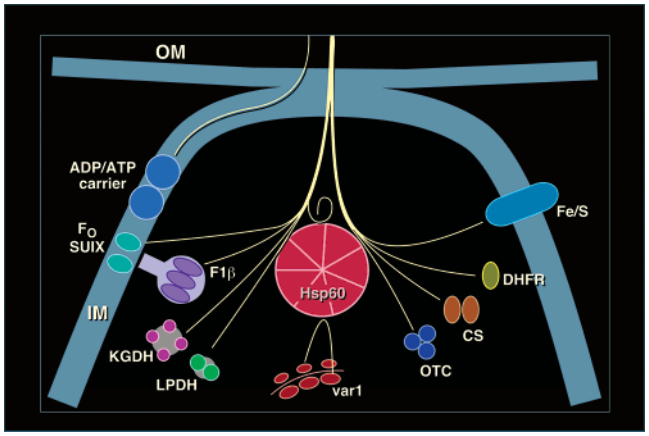


Figure 1. Involvement of the mitochondrial chaperonin Hsp60 with folding newly imported proteins. The biogenesis of the illustrated proteins was affected in yeast cells carrying a defective version of hsp60.[11] While the proteins were normally imported through the mitochondrial membranes (outer and inner, designated OM and IM), they failed to reach active form. The Rieske Fe/S protein, particularly, is a monomer during its lifetime in the mitochondrial matrix, supporting the idea that folding, not oligomeric assembly, is the step affected. Studies with a fusion protein joining a mitochondrial targeting sequence with monomeric normally cytosolic DHFR, showing the newly imported protein to physically associate with Hsp60 in non-native form and to be released by ATP in a native form, supported a model of binding and ATP-mediated folding by Hsp60.[12] Note also that preexistent assembled Hsp60 mediates the folding of newly imported Hsp60 subunits − wild-type subunits imported into mitochondria bearing the defective version of hsp60 were unable to be folded and assembled to make new Hsp60 complex.[90] DHFR, dihydrofolate reductase; CS, citrate synthase; OTC, ornithine transcarbamylase; LPDH, lipoamide dehydrogenase;[91] KGDH, ketoglutarate dehydrogenase;[91] F1β, subunit of F1ATPase.

they then failed to reach native form, apparently misfolded (Figure 1).

While it seemed likely that it was newly imported monomeric polypeptides that were being misfolded, it remained that many of the proteins studied ultimately formed oligomeric assemblies. Could a defect at the level of oligomeric assembly be excluded? It seemed so, because one of the proteins to be studied in the mutant was the Rieske Fe/S protein, an inner membrane protein that is first translocated to the matrix compartment, has its signal proteolytically removed in two steps, and is then inserted into the inner mitochondrial membrane. During its lifetime in the mitochondrial matrix, where the cleavages occur, it is a monomer. In the mutant cells, the Fe/S protein failed to be properly cleaved, being found in either a noncleaved or once-cleaved state.[11] It apparently had misfolded as a monomer, preventing the second step of cleavage.

The gene that rescued the mitochondrial folding mutant turned out to encode a newly recognized heat inducible protein of mitochondria, first identified by Hallberg,[14] subsequently called heat shock protein 60. With sequencing of the Hsp60 gene,[11,15] it at once became clear that it encoded a chaperonin, with a predicted primary structure greater than 60% identical to GroEL and Rubisco binding protein. Moreover, EM studies of Hallberg revealed that Hsp60 subunits formed the same double ring structure as the other two components.[16] A further study directed by Hartl directly examined the action of Hsp60 in mediating folding of a monomeric protein to its native state, importing monomeric dihydrofolate reductase (DHFR) via a fused signal peptide into isolated *Neurospora* mitochondria.[12] The newly imported protein became associated with Hsp60 in a non-native form and was productively released from it into a form with the properties of the native state by addition of ATP to the mitochondria.

Late in 1989, shortly after these studies in mitochondria, the in vitro reconstitution of folding of denaturant-unfolded subunits of a dimeric Rubisco was reported by Lorimer and his colleagues using purified *Escherichia coli* GroEL and the cooperating component with which it shares an operon, GroES, in the presence of ATP.[17] The reconstitution was carried out in two steps. In the first, the unfolded Rubisco, diluted from denaturant into buffer containing GroEL, became physically associated with the chaperonin in an enzymatically inactive, apparently non-native, form. Such binding forestalled irreversible and quantitative aggregation that otherwise occurred in the absence of chaperonin, associated with failure to recover any enzymatically active protein. In the second step, the addition of GroES and ATP to the GroEL−Rubisco binary complex produced nearly quantitative recovery of native Rubisco on a time scale of several minutes.

The collective of these studies, in organelles and in vitro, shifted incontrovertibly our recognition of where chaperonins act to the level of polypeptide chain folding. They raised the significant new notion that, for many proteins in the living cell, a further step had to be added to the pathway of information transfer from DNA to effector protein, namely, a step of assisted folding. Consistently, cellular studies made clear that the role of chaperonins was an essential one under all conditions − deletion of either GroEL or Hsp60 was observed to be lethal to bacteria and yeast, respectively, at all temperatures.[11,15,18] A second chaperonin family was soon discovered in archaebacteria and the eukaryotic cytosol,[19] and, similarly, where genetic deletion was feasible in the latter case, it was lethal.[20] Thus, while many of these components are induced under heat shock conditions, presumably to provide action in the repair of misfolded proteins, the chaperonins have an essential basal role in mediating polypeptide folding.

How does such a role agree with the principles articulated by Anfinsen? From the time of the early mitochondrial studies, where Hsp60 was found to assist a large number of different proteins, it seemed clear that no steric information could be supplied by the chaperonin. Rather, it seemed that the role was to provide kinetic assistance to the in vivo folding process, binding conformational states that would otherwise misfold and aggregate, as in the Rubisco reconstitution experiment. Whereas folding could often proceed efficiently in a test tube under conditions of high dilution and relatively low temperature without such off-pathway behavior, it was recognized that under cellular conditions, where the temperature is higher and where total solute concentration is very high, misfolding and aggregation become an alternative fate, particularly for larger polypeptides that fold with slow or complex kinetics.

Several examples of misfolding in vivo had already been observed. For example, expression studies in bacteria had noted that, when various eukaryotic proteins were overproduced, they could accrete in inactive forms in morphologically visible refractile inclusion bodies, reflecting their inability to properly fold.[21] Particularly incisive were the studies of Pelham in the mid-1980s, observing first that overexpression of Hsp70 could accelerate the recovery of

normal nucleolar morphology in intact cells following heat shock.[22] Subsequent studies of binding and ATP-driven release of Hsp70 from nuclei of heat shocked cells then led Pelham to suggest that the action of the Hsp70 class of heat shock proteins was to protect proteins from multi-molecular aggregation by binding them in a salt-insensitive manner, via what he proposed were likely to be hydrophobic surfaces that were selectively exposed during heat shock but normally buried to the interior of a protein in its native state.[23] Thus, kinetic assistance could be provided to a large group of proteins by simply binding such surfaces before they could associate with each other to produce multi-molecular aggregation.

The behavior of GroEL in polypeptide binding, as observed in the original Rubisco reconstitution experiment, was consistent with this. For example, when GroEL was supplied to a refolding mixture of Rubisco at later times after dilution from denaturant, it could no longer efficiently bind or refold the protein − aggregation supervened.[17] Chaperonin binding thus competed with the process of misfolding and multi-molecular aggregation. In sum, then, a role of the chaperonins, and of chaperones more generally, is to prevent proteins from lodging in a variety of misfolded kinetically trapped states that are not readily accessible to the native state that can lead to irreversible aggregation. Notably, however, there may be "on-pathway" intermediates that are also recognized and assisted in folding (e.g., ref 94).

There was an additional surprising activity of the chaperonin system, however, beyond that of binding non-native states and preventing their aggregation − productive folding to the native state could be triggered by addition of ATP and a cooperating cochaperonin to a chaperonin−polypeptide binary complex. This was an action that Hsp70 and other chaperone proteins could not readily accomplish. To characterize the mechanism of this unique folding activity, an effort incorporating biochemical and structural studies, over a period of more than a decade, has been carried out by a sizable community of investigators, studying the bacterial chaperonin, GroEL. Whereas initial studies examined the behavior of heterologous substrate proteins, from mammalian mitochondria or even blue-green algae, recent studies indicate that the system also works homologously, with a large number of *E. coli* proteins acted upon in the same way by the bacterial GroEL system both in vitro and in vivo (ref 95 and Chapman et al., unpublished). Yet even with such knowledge, the fate of polypeptide through the increasingly well-defined set of states of the GroEL−GroES machine remains only weakly resolved.

3. Structural States of GroEL and the GroEL–GroES Reaction Cycle

3.1. Architecture of GroEL and GroES

To describe the chaperonin reaction cycle, it is useful to first define the structural features and conformational changes of the chaperonin and its cochaperonin (Figure 2). GroEL in its unliganded state is a cylindrical structure, 135 Å in diameter and 145 Å in height, a homotetradecamer composed of two back-to-back rings each containing seven radially arranged subunits.[24] Each GroEL subunit folds into three domains, an equatorial domain, at the waistline of the cylinder, that houses the nucleotide pocket, the collective of which forms the stable base of the assembly; an apical domain, at the end of the cylinder, which presents a hydrophobic surface at its cavity-facing aspect, the collective surfaces of which can multivalently bind a non-native polypeptide but also, following rigid body elevation and twist, can bind 1:1 with the subunits of the cochaperonin GroES; and an intermediate domain, a small structure with hinge connections at its top and bottom to the other two domains. At the equatorial level of the central cavity of both rings lies the collective of the flexible C-terminal tails of the GroEL subunits, 22 amino acids each, containing a repeating GGM sequence that accounts for their disordered behavior. These tails cannot be resolved crystallographically, but the collective can be seen at low resolution in EM, positioned as an axial mass in the central cavity at the equatorial level of each ring.[25] Although they obstruct the central cavity, they do not appear to play a critical functional role because they can be deleted without effect on GroEL's essential action in maintaining cell viability.[26] Thus, the cylinder contains a central cavity at either apical end that is 45 Å in diameter and ~40 Å in height, of sufficient volume to house a globular non-native protein of ~30 kDa size. Of course, because this cavity is open at its terminal aspect, a larger polypeptide can be accommodated if a portion of it resides outside the cavity in the bulk solution. Such a topology has been observed in small angle neutron scattering of a binary complex of the substrate protein rhodanese in complex with GroEL, a topology resembling a champagne cork.[27]

GroES is a single seven-membered ring whose identical subunits are each composed of a β-barrel body from which extends a hydrophobic loop structure that is mobile and disordered in the standalone state but that becomes stably associated via hydrophobic interaction with a GroEL apical domain upon complex formation (Figure 2).[28−31] GroES has a small central cavity of its own that becomes smoothly contiguous with that of GroEL upon association with it. It thus effectively forms a molecular "lid" for the central cavity. Interestingly, consistent with the GroES cavity playing a participating role in folding of GroEL−GroES-encapsulated polypeptides, an in vivo-selected mutation aimed at improving the efficiency of GroE-mediated folding of GFP in *E. coli* altered a tyrosine (Y71) that projects into the small GroES cavity to a residue with a charged side chain.[32] This mutation, while beneficial to GFP folding, had a strongly inhibitory role on folding of other chaperonin substrates.

3.2. Polypeptide Binding to a GroEL Ring

As mentioned, early studies revealed that polypeptide binding by an open ring of GroEL captures non-native forms of a protein, typically only one polypeptide per ring, and serves to prevent irreversible misfolding and aggregation.[17,33,34] Structural and mutational studies indicated that binding is mediated through hydrophobic residues lining the cavity-facing aspect of the apical domains, positioned on a tier of three structures, two α-helices and an underlying extended segment (Figure 3).[24,35] Alteration of any of these hydrophobic apical residues to hydrophilic character abolishes polypeptide binding. Additional experiments producing GroEL rings as single 7-fold tandemized molecules have allowed evaluation of the arrangement of hydrophobic apical domains required for efficient polypeptide binding.[36] A minimum of three contiguous intact apical surfaces are required. Physical cross-linking experiments corroborate this requirement for multivalent binding, observing substrate simultaneously contacting multiple apical domains.[36]

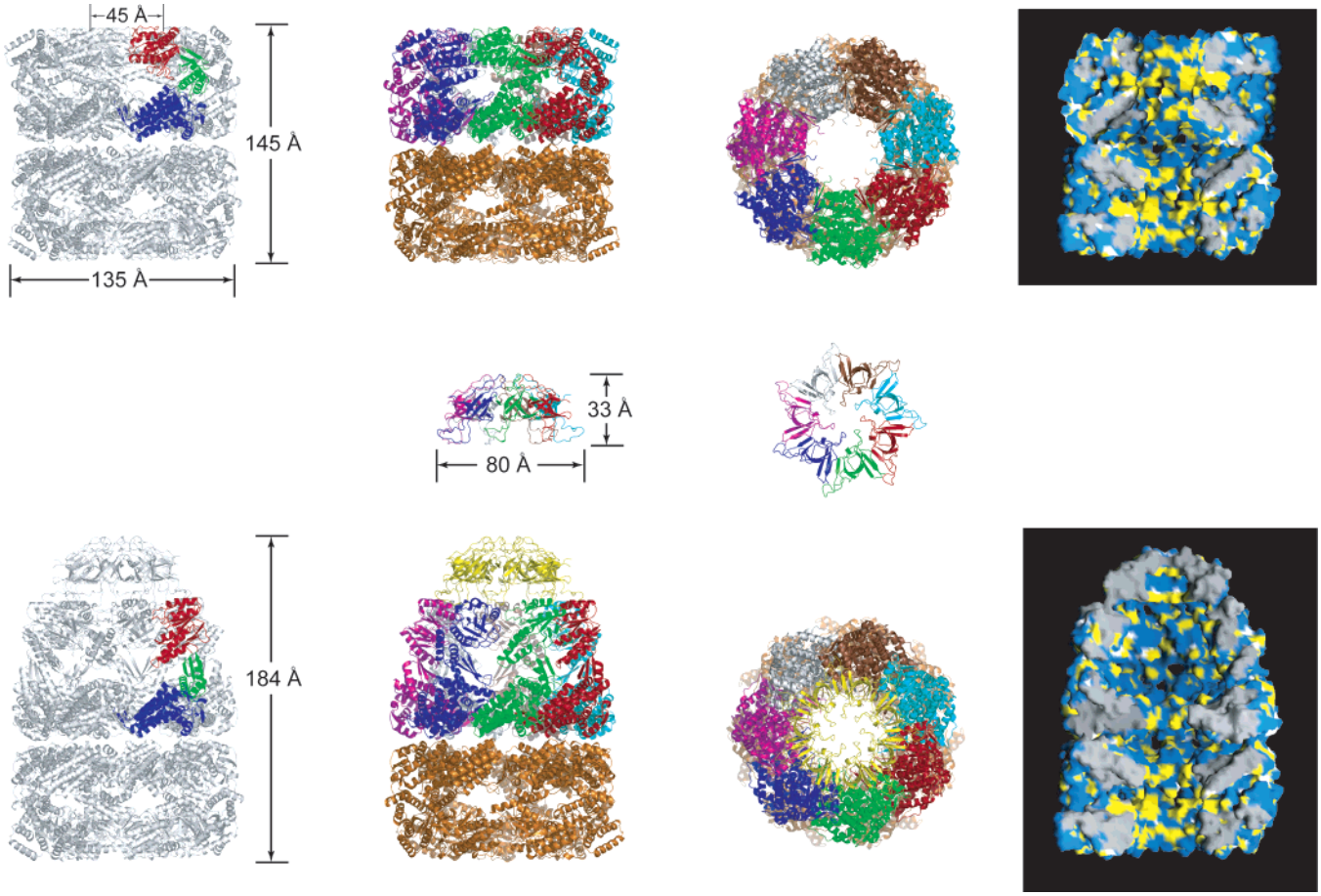


Figure 2. Architecture of the bacterial chaperonin system, GroEL and GroES. Crystallographic models of GroEL, GroES, and GroEL−GroES complexes are shown in the top, middle, and bottom rows, respectively (pdb1oel and 1aon). Left column, a single subunit of GroEL is colored in GroEL and GroEL−GroES: apical domain, red; equatorial domain, blue; intermediate domain, green. Second column, assemblies with GroEL and GroES subunits colored individually in selected rings. Third column, subunits colored individually in end views looking down the central cavity, which measures ~45 Å in diameter. Note, however, that this cavity is obstructed at the equatorial level in both rings by flexible C-terminal tails of each GroEL subunit that are not resolvable crystallographically (see text). Fourth column, space-filling model colored to illustrate hydrophobicity (yellow) and hydrophilicity (blue).

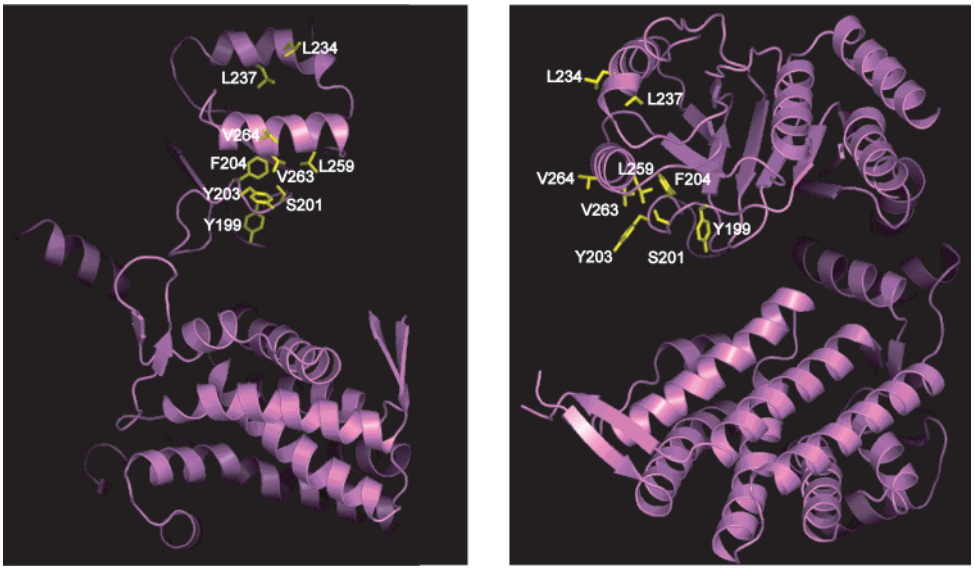


Figure 3. Apical peptide binding surface of a GroEL subunit. Ribbons model shows hydrophobic side chains exposed to solvent, mapping onto three secondary structures, from top to bottom, Helix H, Helix I, and underlying segment, viewed from central cavity, left, and from the side, right. Mutational change of any one of these residues (seven within a ring) abolishes polypeptide binding by GroEL.

The stereochemistry of hydrophobic interactions between substrate and apical domain has been probed using bound peptides, observing in two different X-ray studies the positioning of a peptide in a groove between the two α-helices H and I, with formation of hydrophobic contacts between hydrophobic peptide side chains and those of helices

H and I.[37,38] However, this topology resembles quite closely the positioning of the GroES mobile loop in contact with the GroEL apical domains as it occurs in the GroEL−GroES complex,[30,31,39] leaving as uncertain whether the topologies being observed with peptides are reflecting ones that would be adopted by a non-native polypeptide. Further studies will clearly be needed to address how polypeptide binds across multiple apical faces. Such studies are challenging because bound polypeptide substrates are poorly structured, as described below, obviously not obeying the same 7-fold symmetry that the GroEL−GroES machine itself does (see ref 2).

Bound substrate proteins appear to occupy weakly structured conformations, as indicated particularly by high susceptibility to exogenously added proteases[40,41] and as demonstrated by a number of hydrogen−deuterium exchange experiments that reveal little or no protection from exchange of amide protons of GroEL-bound proteins.[42−47] Moreover, recent NMR studies directly examining an isotopically labeled substrate, human DHFR, while bound to GroEL, indicate that the bound protein lacks any stable structure and that it undergoes significant conformational exchange.[48] Thus, binding by GroEL retains a substrate in a relatively unfolded state.

The act of binding substrate in an open GroEL ring appears to exert an unfolding action on non-native species, effectively pulling a protein out of a kinetic trap back up the energy landscape. An unfolding action by such binding was indicated by early kinetic studies that observed that GroEL could reverse incipient misfolding of subunits of malate dehydrogenase (MDH), rescuing the protein from low order aggregates that could otherwise proceed to irreversible aggregation.[49,50] Such unfolding could be accomplished by either of two mechanisms. One involves thermodynamic partitioning, in which GroEL has greater affinity for less-folded states of any given substrate and shifts an ensemble of states in rapid equilibrium toward the less folded states by binding them.[51] Additional evidence for such an action was presented with a mutant Rnase T1, where GroEL favored binding of a less-structured conformation without altering the microscopic equilibrium between it and a more folded conformation.[52] A second possible mechanism of unfolding is a kinetic one that involves catalyzed unfolding, as observed in a hydrogen−deuterium exchange experiment where incubation of the small RNAse, barnase, with a substoichiometric amount of GroEL in D_2O was associated with its global exchange.[53] This behavior may be the result of multivalent binding.[36] The two mechanisms of unfolding are not mutually exclusive, although the thermodynamic partitioning mechanism may be the major operative one. It remains that our understanding of polypeptide binding is hampered by weakly structured states, the presence of an ensemble of them, and tools that have limits to what they can reveal of these states.

3.3. Rigid Body Movements of GroEL during the Reaction Cycle

3.3.1. ATP Binding

The binding of GroES to GroEL is nucleotide-dependent,[54] with ATP, the physiologic nucleotide, rapidly and cooperatively binding to the seven sites of a GroEL ring.[55−57] As observed by cryoEM, such binding to a ring produces a 20° downward rotation of its intermediate domains and an accompanying movement of its apical domains amounting to a small degree of elevation and 25° counterclockwise twist (Figure 4).[58]

The structural basis to the downward intermediate domain movement as related to ATP binding is not entirely clear. One model would hold that there is a normal range of rotational thermal motion explored by the intermediate domain about its lower hingepoint and that nucleotide binding in the equatorial pocket stabilizes a downward extent of such movement. For example, in an ADP−AlF_x structure of GroEL−GroES, the aluminum fluoride ligand, simulating the terminal phosphate of ATP, forms a bond involving a water-mediated interaction with the intermediate domain, via the carboxylate side chain of residue Asp398, a residue implicated as the base catalyzing ATP hydrolysis (Figure 5).[31] Yet, other bonds may also be able to accomplish stabilization of nucleotide-directed downward movement. Notably, ADP can also promote an apparently similar movement that enables the binding of GroES.[59] In this case, as revealed by a crystal structure of GroEL−GroES−ADP, it appears that the Asp398 side chain is again involved, contacting an oxygen of the β-phosphate via a bound Mg^{2+} ion.[30] In contrast with the cooperative binding of ATP, however, GroEL binds ADP in a noncooperative manner and with 10-fold lower affinity.[60−62]

The ATP-directed downward movement of the intermediate domain is associated with breaking a salt bridge between each apical domain and the neighboring intermediate domain (apical R197-intermediate E386) (Figure 4), freeing the apical domains for their elevation and twisting movement.[58] This also enables a further stabilization of the downward movement of the intermediate domain by the formation of a new electrostatic contact between the freed intermediate domain glutamate (E386) and a side chain positioned on top of the neighboring equatorial domain (K80) (Figure 4b). Thus, the "latched" intermediate domain contacts both the nucleotide pocket of its resident subunit and the equatorial top surface of its neighbor as the result of electrostatic contact switching, which plays a key role in the allosteric adjustments of GroEL.

3.3.2. GroES Binding and the Folding Active State

A GroEL ring with the foregoing rigid body shifts driven by binding of ATP becomes enabled to bind GroES. Such association promotes additional and much larger rigid body movements, producing an end state with 60 degree elevation and 90 degree clockwise twist of the apical domains (Figure 2).[31] The nature of the initial rapid collisional association of GroES has been studied by several avenues, and will be discussed below. But following such initial association, the stable end state with GroES bound with high affinity to ATP-bound GroEL is reached within ~1 s, and in this end state the GroEL apical domains are elevated and twisted, as described, completely removing the hydrophobic polypeptide binding surface from facing the central cavity (Figure 2, right-hand panel). One portion of the hydrophobic surface becomes associated with an edge of the GroES mobile loop, which upon associating with GroEL becomes conformationally ordered as a β-hairpin. Another part of the surface forms a new hydrophobic interface with neighboring apical domains.

A large encapsulated chamber is enclosed underneath GroES in GroES-bound GroEL, comprising a cavity volume (~120 000 Å^3) that is more than double that of an unliganded GroEL ring.[30,31] This is a privileged chamber for polypeptide folding, where the polypeptide substrate, released into it from

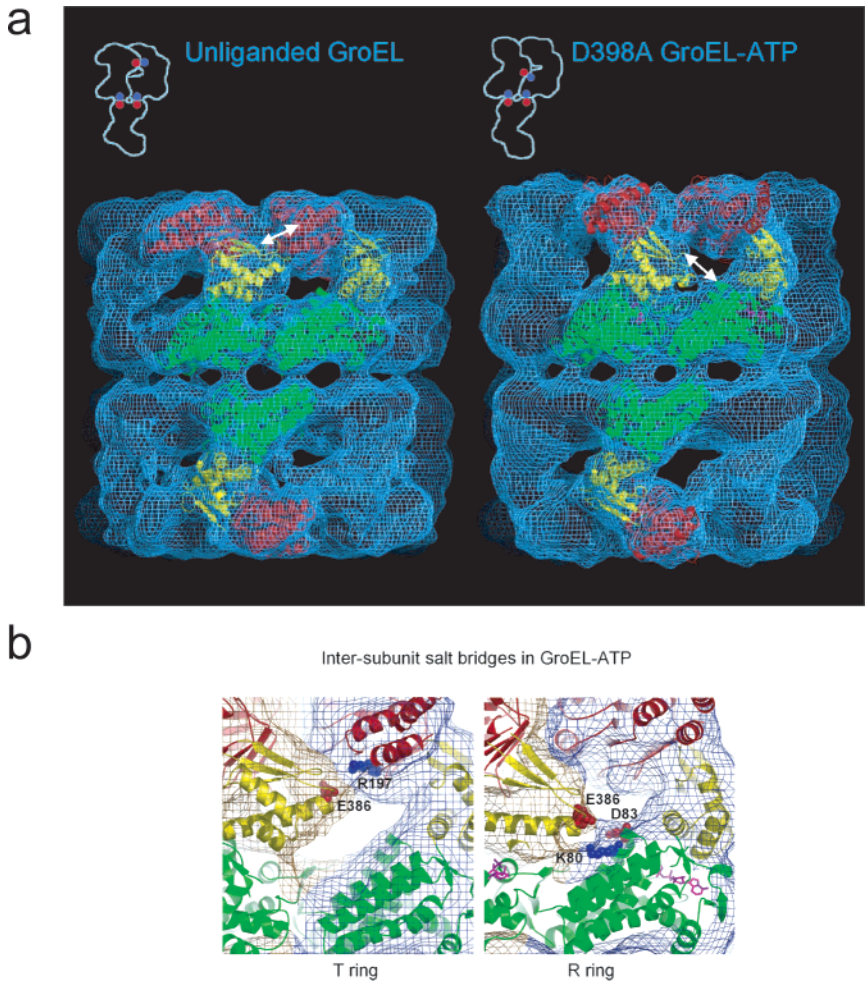


Figure 4. Action of ATP binding in a GroEL ring (in the absence of GroES), comparing unliganded GroEL, left images, and an ATP-bound hydrolysis defective mutant, D398A, right images, by cryoEM image reconstruction.[58] The top images show that ATP produces downward rotational movement of the intermediate domain accompanied by modest elevation and twisting movement of the apical domains of the ATP bound ring, shown as the top ring. The downward movements of the intermediate domain break a salt bridge between E386 and R197 of the neighboring subunit, lower left panel, and E386 forms a new contact with K80 at the top aspect of the neighboring equatorial domain. The broken 386−197 bridge allows the apical domains to move freely, enabling their elevation and twist. Note that the direction of twist is ~15° counterclockwise, opposite to the direction taken upon GroES binding, which totals, from this position, 120° clockwise. Reprinted with permission from ref 58. Copyright 2001 Elsevier.

the hydrophobic binding surface as it elevates and twists, is both isolated from exposure to any other protein, preventing any possibility of aggregation, and exposed to cavity walls that are now hydrophilic in character. This wall character potentially promotes productive folding by energetically disfavoring continued exposure of hydrophobic surface in the folding protein, favoring burial of such side chains to the interior, concomitant with exposure of hydrophilic ones that are accessible to solvent in the native state. In addition, the confinement of the folding substrate protein in the relatively narrow space of this chamber has a likely effect of limiting the extent of off-pathway folding that can occur.

Support for a role of the cavity in limiting the range of explored conformations comes from a recent experiment comparing the rates of folding to native form of the substrate Rubisco inside the GroES-encapsulated chamber and alone in solution under "permissive" conditions of lower temperature, where GroEL and GroES are not absolutely required for productive folding.[63] Under such conditions, Rubisco folded more rapidly to native form inside the chamber, presumably a function of a more favorable free energy landscape, one likely limiting particular off-pathway steps. Considering that Rubisco has a subunit size of 52 kDa, just barely able to fit in native form into the GroES-encapsulated chamber, this seems to indicate that more extended conformational states simply could not be populated inside the chamber. In sum, this advantage, plus those of solitary confinement and a hydrophilic cavity lining, translate to the experimental observation that protein substrates maintained in this encapsulated space recover their native state quantitatively − they do not undergo any irreversible misfolding at this site.[64,65]

There may be other significant facets to folding in the encapsulated cavity that as yet have not been probed. Is folding in this space occurring essentially at "infinite dilution," as if in a sea of infinite solvent? The degree of confinement based on relative volumes seems to argue against this, as does the observation that the rotational correlation time of a refolded native protein, GFP, is 4-fold longer when it is inside the cavity than when it is free in solution.[64] Thus, even this relatively small 28 kDa native protein is "seeing" the cavity wall, behaving as if it were 100 kDa. A further aspect of behavior during folding about which mystery remains concerns the fate of water, both at the cavity lining and around the non-native substrate protein. Are there major shifts during the folding reaction, and what

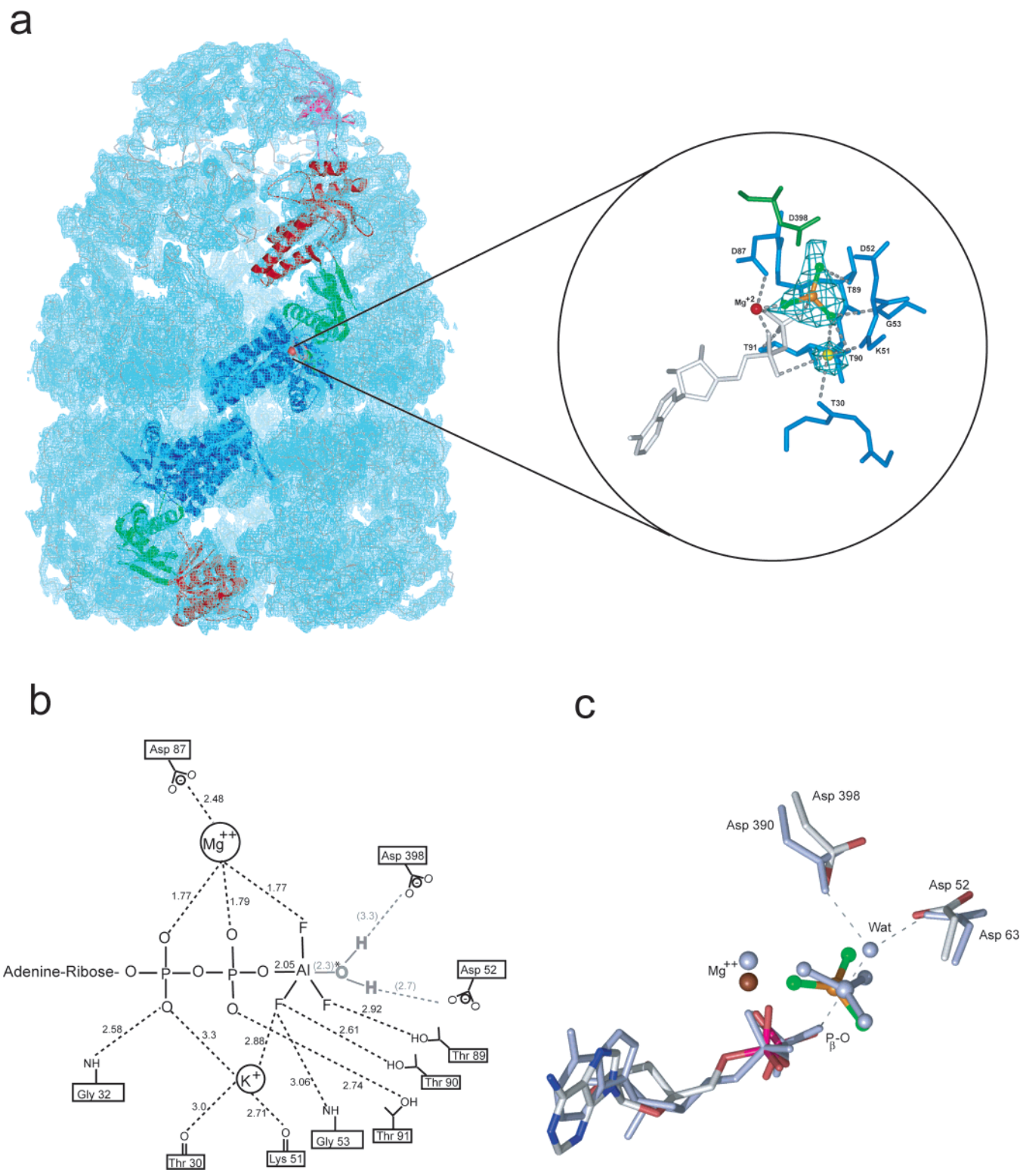


Figure 5. Crystallographic model and electron density map of GroEL−GroES−ADP−AlF_3, taken from ref 31 (panel a), and schematic illustrations of nucleotide pocket (panels b and c). In panel a, the map is calculated with coefficients F_{obs} (50−2.8 Å) and averaged phases resulting from 7-fold NCS averaging, density modification, and phase extension starting from random phases, and is contoured at 1σ. A C_α trace is shown for all of the subunits. One subunit in the cis, GroES-bound, ring, one subunit in GroES, and one in the trans ring are colored to indicate individual domains. The site of the magnesium in the nucleotide pocket of the colored cis GroEL subunit is designated with a red ball. The blowup view shows the contacts between the terminal aluminum fluoride (orange and green) and the surrounding nucleotide pocket of the subunit, with equatorial GroEL residues colored blue and the intermediate domain residue Asp398, involved in activating a water (shown in panels b and c), in green. A potassium ion is designated in yellow. Panel b schematically illustrates contacts between ADP−AlF_3 and the surrounding GroEL nucleotide pocket in the cis ring based on the X-ray structure. The activating water is shown in gray to denote that its electron density is weak in the present structure. However, in the isomorphous nucleotide pocket of the archaeal thermosome, the water is clearly visible.[92] Panel c: Stick model of the nucleotide pocket as in the blowup but showing the water at the in-line position. Reprinted with permission from ref 31. Copyright 2003 Elsevier.

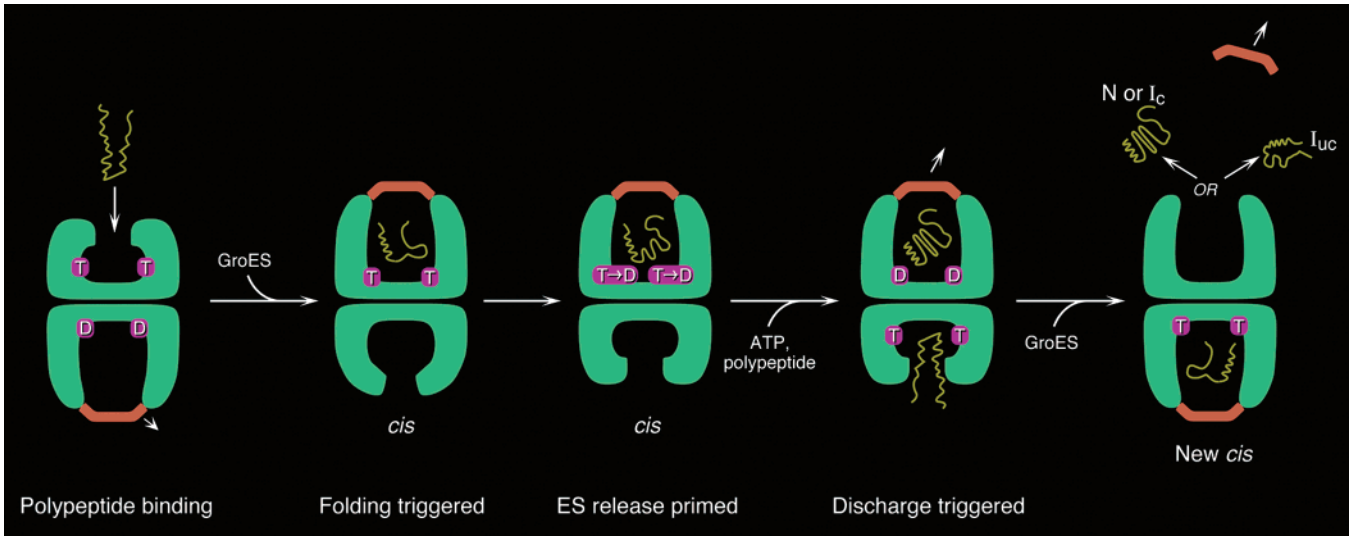


Figure 6. GroEL−GroES reaction cycle. Asymmetric GroEL−GroES−nucleotide complexes are likely the predominant states in the cell. A GroEL−GroES−ADP asymmetric complex binds a non-native substrate polypeptide in a collapsed, largely unstructured state via hydrophobic interactions with its apical domains (panel a), an action that may result in further unfolding of the substrate. Binding of ATP (either before polypeptide or thereafter; the relative rates of arrival have not been measured) followed by GroES to the same ring as polypeptide (panel b) results in large-scale conformational changes of the GroEL apical domains, which move the hydrophobic sites away from the cavity surface and permit GroES to cap the GroEL ring, forming a cis folding-active complex. Concomitant with the ATP/GroES-driven apical movements, polypeptide is ejected into the central cavity (<1 s), where it begins to fold in this sequestered chamber whose walls are now hydrophilic in character. This wall character favors burial of hydrophobic side chains of the substrate protein and exposure of its hydrophilic ones, characteristic of the native state. Folding continues in this longest-lived state of the cycle until ATP hydrolysis in the cis ring (panel c), occurring with a half-time of about 8−10 s, weakens the cis complex and permits ATP (and non-native polypeptide) binding to the trans ring (panel d). This rapidly allosterically discharges the ligands from the cis ring (panel e) (<1 s), regardless of the folding state of the substrate polypeptide. The released polypeptide may have folded to its native state (N) or one committed to it (I_c), or it may have failed to reach the native state (I_{uc}), in which case it can be bound to GroEL again for another attempt at folding. GroES binding to the same ring as ATP and polypeptide forms a new cis complex on the previous trans ring to begin the cycle again. Thus, GroEL alternates its rings between binding-active and folding-active states, using one round of seven ATPs to simultaneously discharge one folding-active complex while forming another. Note that an asymmetric ATP complex has no significant affinity of its trans ring for any of the ligands, whereas once cis ATP hydrolysis occurs, it can accept non-native polypeptide and ATP, and, in an ATP-bound state, can now bind GroES.

are the thermodynamic and structural correlates of such changes?

3.4. Reaction Cycle

The reaction cycle of GroEL and its cooperating component, GroES, is illustrated in Figure 6. The cycle is governed by the inherently asymmetric behavior of the two GroEL rings with respect to binding ATP.[57,96] Within a ring there is cooperative binding of ATP by the seven equatorial sites, with an apparent K_M of ~10 μM and a Hill coefficient of ~3, while there is strong negative cooperativity between the rings (see ref 93 for review). As just described, ATP binding promotes small rigid body movements of the intermediate and apical domains in the ring to which it binds. This allows GroES association, which is followed within a second by much larger rigid body movements of the GroEL intermediate and apical domains, resulting in the end-state GroEL−GroES complex described above. As mentioned, the GroEL apical domains are elevated by 60 degrees and twisted 90 degrees clockwise, removing the hydrophobic binding surface from facing the cavity. Associated with these large movements, a polypeptide substrate protein, initially captured on the hydrophobic cavity wall of an open GroEL ring, is rapidly released into the central cavity of the now hydrophilic GroEL−GroES cis cavity, where it commences to fold.[64,65] Folding proceeds in this cavity for the longest part of the chaperonin cycle, amounting to 8−10 s.[66] At the end of this period, ATP hydrolysis in the cis ring serves to weaken the otherwise very stable association of GroES with GroEL, priming the ring for release of its ligands, GroES, polypeptide substrate, and ADP.[65] The physiological trigger to such release is the cooperative binding of ATP in the opposite ring.[65,67] Moreover, the additional binding of non-native polypeptide further accelerates release.[66]

The nature of the affinity of the trans ring for both ATP and polypeptide following cis hydrolysis has been revealed by recent EM studies comparing the cis ATP with cis ADP asymmetric GroEL−GroES complexes.[77] These show a set of structural shifts occurring in the interface between the rings and within the trans ring, associated with domain rotations in the trans ring, disruption of intra-ring contacts, and expansion of the trans ring, opening both its nucleotide pockets and the apical domains. While this offers an explanation for the acceptance of ligands in trans following cis ATP hydrolysis, how arrival of these ligands allosterically drives dissociation of the cis ligands remains unclear. The acceleration by polypeptide is a particularly interesting allosteric action, because it likely involves an asymmetric set of contacts of polypeptide with the open ring, which is nonetheless transduced across the entire cylinder as an ejection signal that cooperates with the signal produced by ATP.

The binding of ATP and polypeptide to an open ring of an asymmetric complex not only discharges the ligands from the opposite ring, but allows in turn the binding of GroES and the nucleation of a new cis folding-active ring opposite the previous one (Figure 6).[66] Each ring of GroEL thus oscillates back and forth between binding-active and folding-

active modes, out of phase with each other, employing at each cycle a set of seven ATPs to simultaneously discharge an old folding-active ring and nucleate a new one. Notably, it is the energy of ATP binding that carries out the work at each step, triggering folding when GroES binds subsequently in the cis ring and ejecting the ligands when bound in trans. Hydrolysis of seven ATPs per ring, on the other hand, provides directionality to the cycle.[66]

For any given round of folding at GroEL, only a small percentage of initially bound polypeptide substrate molecules reach native form. For example, only ~5% of rhodanese molecules bound to GroEL reach native form in a single round of ATP/GroES-mediated folding,[68] and non-native Rubisco bound to GroEL exchanges with non-native protein in solution much faster than it folds to its native state.[69] Thus, most of the molecules are released in a non-native form that has to be rebound by GroEL for a further trial at folding.[68−71] Furthermore, the non-native states rebound to GroEL appear to be the same or very similar to those initially bound.[45,68] Folding at GroEL appears thus to be an all-or-none process as opposed to an iterative one, in which there is acquisition of progressively more structure with each round at the machine. Finally, in a cellular context, a released non-native polypeptide substrate not only could be rebound by GroEL but also could be recognized by other chaperones or protease components, a kinetic partitioning behavior that governs the fate of non-native polypeptides in the cell.

4. Triggering Productive Folding

The nature of the transition between a binding-active open GroEL ring and the folding-active GroEL-GroES state (see Figure 6) has been the object of considerable experimentation and mechanistic consideration. There are a host of questions about this transition. Why does ATP/GroES but not ADP/GroES trigger productive folding? What are the correlate structural changes produced by these additions? Why does non-native polypeptide not escape during this process? Work of the past few years, described below, begins to address these questions, and further interesting structural and mechanistic questions arise from these observations.

4.1. cis Ternary Complexes Are Formed by Addition of Either ADP/GroES or ATP/GroES to GroEL−Polypeptide Binary Complexes

Early experiments with such GroEL−GroES-dependent substrate proteins as rhodanese, MDH, or Rubisco made clear that their productive folding occurs only in the presence of GroEL, GroES, and ATP, unable to be supported by ADP.[17,40,49,50,72] Subsequent experiments concerning substrate topology indicated, however, that such substrates as rhodanese or MDH could be encapsulated by GroES in the presence of either ATP or ADP, and these substrates were no longer susceptible to exogenously added protease as they were when bound in an open GroEL ring.[73,74] Yet when fluorescent reporters in these substrate proteins were analyzed, it became clear that these substrates do not undergo any conformational change upon GroES binding in the presence of ADP. For example, pyrene-labeled rhodanese exhibited a rapid drop of fluorescence anisotropy beginning in the dead time of stopped-flow mixing when ATP/GroES was added, but no change at all occurred upon addition of ADP/GroES.[64] In the case of Rubisco, a similar rapid change in anisotropy of endogenous tryptophans was observed immediately upon addition of GroES/ATP, followed by a later increasing phase, correlating directly with production of the native enzymatically active state,[65] but neither change was seen with ADP/GroES. In the case of tryptophan-substituted MDH, fluorescence intensity change commenced in the dead time of mixing with ATP/GroES, but no change was observed in ADP/GroES.[31] Thus, the fluorescence experiments in ATP, while reporting on local motions and environment of the respective fluorophores, are almost certainly reflecting more globally the rapid and complete release of substrate polypeptide from the GroEL cavity wall, attendant to the rigid body movements occurring in the GroES-bound GroEL ring on the same time scale (see below). Moreover, with respect to nucleotide, these experiments indicate that ADP/GroES is unable to trigger such release, even though, in the case of rhodanese and MDH, GroES can nevertheless be bound to the same GroEL ring as polypeptide in ADP and can encapsulate the substrate protein.

Further evidence for this latter conclusion comes from gel filtration analyses of rhodanese−SR1−GroES−ADP complexes. SR1 is the single ring version of GroEL that is fully active in folding substrates but cannot release them or GroES except at low temperature.[31] Upon treatment of rhodanese−SR1−GroES−ADP complexes at 4 °C, rhodanese remains quantitatively physically associated with SR1, evidence that it was never released during the initial complex formation. By contrast, none of the rhodanese molecules remain physically associated with SR1 when ATP is used, with a significant percentage recovered at the migration position of the native enzyme (33 kDa), indicating that they had been released from the GroEL cavity wall and folded.

Thus, the γ-phosphate moiety of ATP is critical to triggering release of polypeptide substrates from the cavity wall, enabling folding to proceed in ternary substrate−GroEL−GroES complexes. Could this moiety be added independently to a folding-inactive ADP complex and drive its transition to a folding-active state? Aluminum fluoride has been shown to act, with GDP or ADP, as an analogue of the transition state for ATP or GTP hydrolysis, triggering functional activation of a variety of nucleotide-dependent machines. For example, GDP-aluminum fluoride triggers heterotrimeric G protein-mediated signal transduction.[75] Here, such a complex would mimic the transition state for ATP hydrolysis, potentially able to trigger productive folding. This is precisely what was observed when aluminum fluoride was added to a folding-inactive rhodanese−GroEL−GroES−ADP complex.[31] As shown by fluorescence anisotropy studies, rhodanese was rapidly ejected from the cavity wall, and its folding to the native state proceeded with kinetics resembling a wild-type or SR1-mediated reaction. That a γ-phosphate mimic could be effective in triggering folding not just at the transition state for ATP hydrolysis but at an earlier point along the reaction coordinate was revealed by identical triggering of productive folding by addition of beryllium fluoride to an ADP complex, forming an analogue of the ground state of ATP. Thus, the γ-phosphate could be supplied as an independent entity to mimic the γ-phosphate of ATP in a state somewhere between the ground state and transition state of hydrolysis to trigger productive folding by GroEL−GroES. By contrast, a posthydrolysis analogue, ADP-vanadate, could not support folding, although like ADP, it could support production of GroEL−GroES complexes.[31]

Because aluminum fluoride could be added independently to ADP−GroEL−GroES complexes, the energetics of bind-

ing could be probed in equilibrium binding measurements. A competition assay with ^{7}BeF allowed an estimate of the affinity for aluminum fluoride and in turn an estimate of $\Delta G°$ for aluminum fluoride binding of ~45 kcal/mol of GroEL rings.[31] This is a substantial amount of free energy, roughly equal to the sum of the energy for the steps of ADP plus GroES binding measured from calorimetry and Hummell-Dreyer analyses,[31] respectively. Thus, the free energy of binding of the γ-phosphate of ATP functions as a critical element of the folding trigger. What does this energy do? For some time, we and others presumed that it was used to produce a different structure of the entire complex that was uniquely configured to release polypeptide into the encapsulated chamber and drive its folding. Recent experiments suggest another explanation.

4.2. A Structural Conundrum – GroEL–GroES Complexes Formed in Either ATP or ADP Reach the Same Structural End State

With ATP/GroES able to release substrates into the central cavity of GroEL but ADP/GroES unable to do so, it seemed likely that the respective end-state ternary complexes would differ structurally from each other. An earlier crystallographic study had examined a GroEL−GroES−ADP crystal, formed by adding ADP/GroES to GroEL.[30] This asymmetric complex exhibited GroES bound to one ring, containing ADP in each of its seven equatorial nucleotide sites. The difficulty in obtaining an ATP/GroES-formed complex to analyze is that ongoing ATP hydrolysis, occurring normally after ~10 s in the GroES-bound (cis) ring will produce a GroEL−GroES−ADP one. Attempts to prevent this from occurring, using either the hydrolysis-deficient mutant D398A[65] or a solution depleted of the monovalent cation potassium, critical to hydrolysis,[69,76] could not prevent such conversion from occurring on a time scale of some hours. But, by contrast, the ability to form stable GroEL−GroES−ADP−AlFx complexes, which were folding-active, suggested that such complexes could be used to represent the structure of an ATP-bound complex.[31] These complexes were stable on the several day time scale, with GroES unable to be dissociated by 0.4 M guanidine HCl, a stability commensurate with GroEL−GroES−ADP complexes. A crystal structure of the aluminum fluoride complex was thus obtained at ~3.0 Å resolution (Figure 5). Aluminum fluoride could be readily resolved with full occupancy in the cis ring, forming bonds with the nucleotide pocket and a water-mediated contact with intermediate domain Asp398. Surprisingly, however, when the positions and conformations of the intermediate and apical domains of the GroES-bound GroEL ring were examined, no difference with GroEL−GroES−ADP was observed.[31] Was this the result of crystal lattice restraints? To address this, both further crystallographic work and cryoEM studies were carried out. The former study involved forming a complex of SR1 with GroES and ADP-aluminum fluoride, and produced a crystal in a different space group, implying a different set of lattice contacts. Although this structure was obtained at low resolution, ~7 Å, it nevertheless exhibited identical apical and intermediate domain positioning.[31]

The cryoEM study involved analysis of hydrolysis-defective D398A GroEL in complex with ATP and GroES.[77] Here the complex was free in solution, devoid of any lattice restraints. A large number of particles were analyzed and were directly compared with GroEL−GroES−ADP complexes frozen and analyzed identically. These studies, achieving 8−10 Å resolution, again showed no difference between ADP and ATP complexes at the level of the cis ring, with the EM-derived models agreeing well with the crystallographic ones. Thus, regardless of nucleotide, it seemed that the same GroEL−GroES end state was being reached. How could one such state be folding-active and the other not?

4.3. Resolution of the Conundrum – Polypeptide Substrate Acts as a Load against which only ATP Can Drive Apical Movement

After much consideration, it occurred to us that all of the foregoing structural analyses had been carried out in the absence of substrate polypeptide. If substrate was acting as a "load" on the GroEL apical domains, then perhaps only ATP/GroES could overcome the load, whereas ADP/GroES could not. The most direct way to test this hypothesis would be to monitor apical movement in real time, comparing the effects of the respective additions to complexes without and with bound polypeptide on such movement. To carry out such monitoring, fluorescence resonance energy transfer (FRET) was employed.[78] Two cysteine substitutions were produced in a cysteine-deficient version of GroEL, one on the immobile equatorial "base" of GroEL at the cavity-facing aspect and one at a lateral position on the mobile apical domain. These positions were chosen so that, when appropriate fluorophores were covalently linked to the cysteines, FRET would be produced in the unliganded state and would decrease when opening of the apical domain produced a major increase in the distance between the two fluorophores. With such a system established, time-dependent measurements of FRET were taken in the absence and presence of substrate polypeptides, examining for effect on nucleotide/GroES-driven apical movement.

In the absence of substrate polypeptide, there was rapid loss of FRET upon either ATP/GroES binding or ADP/GroES binding to GroEL, reaching the full extent of donor fluorescence dequenching within 1 s, reflecting rapid opening of the apical domains in the absence of the substrate polypeptide. When either MDH or rhodanese occupied a GroEL ring, however, there was now an enormous effect of nucleotide with respect to how fast the apical domains could open. ATP/GroES was still able to open the apical domains on the time scale of 1−2 s, a rate approximately one-fourth that in the absence of substrate. By contrast, in ADP/GroES, the apical domains did not open on a physiologic time scale, requiring nearly a minute (compare with 10−12 s of the entire ATPase cycle), consistent with the inability of ADP/GroES to support productive folding.

Similar results were obtained using either GroEL or SR1, the latter an obligately cis-forming assembly. These results indicated that substrate polypeptides, rhodanese and MDH, which bind the apical domains multivalently,[36] act as a load on them that can only be countered by ATP/GroES. The 45 kcal/mol of free energy provided by the γ-phosphate apparently produces the "power stroke" that moves these domains against the load.

4.4. Nature of the Polypeptide "Load"

What is the nature of the load imposed by substrate polypeptide on the apical domains? Is it highly structured regions of polypeptide that have to be pulled apart? This

had been suggested by studies of the substrate Rubisco, observing in tritium exchange experiments that when the tritiated protein was either diluted into buffer, where it forms a metastable intermediate state, or diluted into a mixture containing GroEL, with which it forms a binary complex, it retained about a dozen tritiums in a highly protected state (protection factors $>10^5-10^6$).[79] This protection was observed to be lost upon addition of ATP/GroES, suggesting that the elevation and twisting movements of GroEL during GroES binding produced an effect of mechanical stretching.[79] However, other substrate proteins − cyclophilin, α-lactalbumin (3 disulfide), β-lactamase, human DHFR, mitochondrial MDH − have exhibited no such protection while bound to GroEL.[42−47] For example, in the case of MDH, experiments carried out using pulsed deuterium exchange of MDH while bound to GroEL have indicated a maximum protection of no greater than 100 of a small region of the protein.[47] Moreover, recent experiments with Rubisco have not reproduced the extent of protection reported originally, under identical conditions.[80] Thus, it appears likely that polypeptide is presenting a load to the apical domains through some property other than being highly structured. It would seem more likely that it is the direct contact with the apical domains, mediated through both hydrophobic interactions and hydrogen bonds,[35,37,38] that is conferring the properties of a load, but additional experiments will be necessary to establish this.

4.5. Is the Polypeptide−GroEL−GroES−ADP Complex Representative of a Collision State?

Substrate polypeptide-loaded GroEL−GroES−ADP complexes appear stalled in an early or incompletely formed GroES−GroEL state, but addition of γ-phosphate analogues, beryllium fluoride or aluminum fluoride, produces rapid onset of apical movement, which proceeds to the full extent, triggering release of polypeptide into the central cavity and initiation of productive folding.[31,78] The substrate−ADP complex thus appears to lie on a productive pathway, potentially reflecting a physiological collision state between GroES and GroEL. Could such a state also occur in the normally productive nucleotide, ATP? As indicated from the FRET studies monitoring apical domain movement, the presence of bound substrate polypeptide does have effects on the *rate of apical movement in ATP*, with at least 4−6-fold slowing observed. The rate of this movement could be compared with the *rate of association of GroES* with substrate-bound GroEL in ATP, measured by FRET between fluorophore-labeled GroES and SR1 (obligately cis) following stopped-flow mixing.[78] In this setting, GroES associates with SR1 at a rate only 2−3-fold slower than in the absence of substrate. By contrast, apical movement is slowed from 4- to 20-fold, with the rate dependent on the particular bound substrate and the concentration of GroES. The rate of GroES binding increased linearly with concentration, consistent with a bimolecular reaction; by contrast, the rate of apical movement increased only at low concentrations of GroES, reaching a plateau at 0.4 μM (with SR1 at 1 μM).[78] Thus, it seems that GroES can assist apical domain movement, but only to a certain extent.

Directly comparing the rates of GroES association and apical movement in the context of substrate-bound SR1 at the physiologic concentration of GroES (1 μM), there is a 10−20-fold slower rate of apical movement, implying the presence of a GroES−GroEL collision state when substrate polypeptide is present. This state, in which GroES has contacted GroEL but the apical domains have yet to move and polypeptide remains bound to them, may correspond to the polypeptide−GroEL−GroES−ADP complexes observed in the FRET experiments.

What, at a higher level of resolution, do such putative collision complexes look like? How does GroES form an initial complex with GroEL? Do the mobile loops of GroES, disordered in the standalone state as determined from NMR studies,[81] already make contact with the hydrophobic peptide-binding apical surface of GroEL at the initial interaction? Notably, if the loops are positioned in a downward position, as is the case when GroES has formed a stable complex with GroEL, then they align reasonably well for such interaction. But is it at all feasible for GroES to make contact with the apical surfaces of GroEL if substrate polypeptide is already occupying them? Or does GroES bind initially only to those apical domains that are not substrate-occupied? An earlier study using covalent ring assemblies with different numbers of apical domains defective for binding polypeptide and GroES seems to support that GroES can contact the same apical domains as polypeptide.[36] A covalent ring assembly with three consecutive nonmutant apical domains was able to productively bind MDH and fold it upon addition of GroES and ATP, suggesting that the wild-type domains bound both substrate and co-chaperonin. Of course, whether GroES was in part recruited through the mutant domains could not be resolved by those experiments, although when all seven apical domains were mutant, no GroES binding occurred. This observation would indicate that in a putative collision complex with such a mutant assembly, three GroES mobile loops could be binding to one aspect of the surface of the three wild-type subunits, while polypeptide is bound to the same apical domains at another aspect of their surface. Further physical studies may be able to resolve whether this is the case. Notably, however, other studies with covalent rings indicate that only a single nonmutant apical domain is sufficient to bind GroES to the level of 68% of a wild-type complex.[36] Thus, considering an alternate model in which GroES could only bind to a wild-type apical domain vacant of polypeptide, if polypeptide was even transiently releasing from a single apical domain of a ring, this could perhaps be sufficient to allow GroES recruitment to occur. Clearly, trapping collision states and analyzing their structures by EM and X-ray crystallography will be essential to addressing these questions.

Interestingly, two different chaperonin mutants appear to behave much like the ADP ternary complexes, arresting in what could be collision-like states. One, first described by Kawata and co-workers,[82] affects a residue in the ascending limb of the intermediate domain, substituting tryptophan for Cys138. This mutant arrests in a ternary complex with rhodanese and GroES in ATP at 25 °C, unable to refold rhodanese. The rhodanese in such complexes is encapsulated underneath GroES as indicated by its resistance to exogenously added proteinase K. Subsequent upshift of temperature to 37 °C reverses this arrest and activates rhodanese refolding. These complexes may thus resemble ADP cis ternary complexes, and structural comparison, for example in EM, should indicate whether the apical domains of such complexes are present in a nonelevated state. Notably, these two potentially structurally similar states are populated by two different mechanisms. The ADP-arrested state is populated by an effect of polypeptide load in the absence of a

γ-phosphate (and thus absence of hydrogen bonds that contribute to the power stroke), whereas the arrested state of the C138W mutant is likely populated by failure of the power stroke to be executed against the substrate load by interference at 25 °C with nucleotide-directed intermediate domain movement or the allosteric transmission of movement.

4.6. Other Potential States along the Reaction Pathway

The transition from the binding-active to the folding-active state of the GroEL machine has generally been regarded as a single structural shift involving the coordinated movement of the apical domains coupled to ATP and GroES binding.[31,83,84] The recent FRET experiments directly observing apical movement directly observe such rapid movement of the apical domains, completed within ~1−2 s.[78] As a corollary, release and initiation of folding of bound polypeptide have been observed to begin coincident with GroES encapsulation and movement of the apical domains, commencing on the time scale of under 1 s.[31,64,65,78] Recently, Yoshida and co-workers have suggested a more complex pathway, based on a series of single-molecule kinetic experiments examining primarily the refolding of GFP.[85] The regain of fluorescence of GFP upon spontaneous refolding from denaturant is a single exponential process, but restoration of fluorescence assisted by GroEL−GroES−ATP, both in ensemble and single molecule studies, shows a delay of about 3 s. This result has been interpreted to mean that the commencement of steps of refolding experiences a similar lag and that the cause of the lag in appearance of fluorescence is the failure of GFP to be released from the apical binding sites until a relatively slow transition ($\tau \sim 3$ s) occurs in the apical domains. This transition, proposed to occur before a rapid step of ATP hydrolysis, is proposed to be followed by a second transition with slightly slower kinetics ($\tau \sim 5$ s) in the resulting GroEL−GroES−polypeptide−ADP complex before discharge of the ligands at the end of the cycle. Only the second ATP complex and the ADP ones would be folding-active. This proposed kinetic mechanism with two transitions of similar rates has been called a "two timer" mechanism.[85]

Although similar kinetic constants have been recovered from bulk-phase FRET experiments examining the relative positions of non-native substrates and the GroEL equatorial domain during refolding,[85] a number of other results do not show these two phases and support the simpler mechanism described above. For example, the time courses of release of rhodanese, Rubisco, and MDH from the apical domains have been examined by changes in fluorescence anisotropy (rhodanese and Rubisco)[31,64,65] or intensity (MDH).[31] The anisotropies show an immediate rapid drop ($\tau \sim 1$ s), interpreted as reflecting release from the apical domains, followed in the case of Rubisco by a single-exponential rising phase with a rate constant corresponding to the rate of regain of enzymatic activity.[65] In the case of MDH, fluorescence intensity begins to change immediately upon stopped-flow addition of ATP/GroES to an SR1−MDH binary complex and proceeds as a single-exponential process with a rate corresponding with regain of activity (ref 31, see Supplementary Figure 2).

Other experiments are likewise inconsistent with a lag in the initiation of folding. For example, the recent examination by FRET of apical domain movement upon addition of GroES/ATP to binary polypeptide−GroEL complexes shows that it occurs as a rapid ($\tau \sim 1.5$ s), largely single-exponential process; notably, there is no lag.[78] Likewise, recently reported FRET experiments directed at following the relative motions of the N- and C-termini of Rubisco during GroEL-mediated refolding show an immediate change of their positions when GroES and ATP are added ($\tau \sim 0.7$ s), indicating that release from the apical domains and initiation of folding occur without any discernible lag.[86] Finally, a large number of kinetic experiments examining changes in GroEL structure during ATP binding and hydrolysis have failed to reveal evidence for the proposed lag in ATP hydrolysis required by the "two timer" mechanism.[59,69,87,88] Thus, the bulk of existing evidence appears to support the model of GroEL action in which ATP and GroES binding drive major conformational changes in GroEL that simultaneously and immediately release a non-native substrate protein from the apical binding sites and initiate refolding.

5. References

(1) Horwich, A. L.; Fenton, W. A. *The Enzymes* **2003**, *23*, 400.
(2) Fenton, W. A.; Horwich, A. L. *Quart. Rev. Biophys.* **2003**, *36*, 229.
(3) Anfinsen, C. B. *Science* **1973**, *181*, 223.
(4) Georgopoulos, C. P.; Hendrix, R. W.; Kaiser, A. D.; Wood, W. B. *Nat. New Biol.* **1972**, *239*, 38.
(5) Takano, T.; Kakefuda, T. *Nat. New Biol.* **1972**, *239*, 34.
(6) Barraclough, R.; Ellis, R. J. *Biochim. Biophys. Acta* **1980**, *608*, 19.
(7) Hemmingsen, S. M.; Woolford, C.; van der Vies, S. M.; Tilly, K.; Dennis, D. T.; Georgopoulos, C. P.; Hendrix, R. W.; Ellis, R. J. *Nature* **1988**, *333*, 330.
(8) Hendrix, R. W. *J. Mol. Biol.* **1979**, *129*, 375.
(9) Hohn, T.; Hohn, B.; Engel, A.; Wurtz, M.; Smith, P. R. *J. Mol. Biol.* **1979**, *129*, 359.
(10) Pushkin, A. V.; Tsuprun, V. L.; Solovjeva, N. A.; Shubin, V. V.; Evstigneeva, Z. G.; Kretovich, W. L. *Biochim. Biophys. Acta* **1982**, *704*, 379.
(11) Cheng, M. Y.; Hartl, F. U.; Martin, J.; Pollock, R. A.; Kalousek, F.; Neupert, W.; Hallberg, E. M.; Hallberg, R. L.; Horwich, A. L. *Nature* **1989**, *337*, 620.
(12) Ostermann, J.; Horwich, A. L.; Neupert, W.; Hartl, F.-U. *Nature* **1989**, *341*, 125.
(13) Eilers, M.; Schatz, G. *Nature* **1986**, *322*, 228.
(14) McMullen, T. W.; Hallberg, R. L. *Mol. Cell. Biol.* **1987**, *7*, 4414.
(15) Reading, D. S.; Hallberg, R. L.; Myers, A. M. *Nature* **1989**, *337*, 655.
(16) McMullen, T. W.; Hallberg, R. L. *Mol. Cell. Biol.* **1988**, *8*, 371.
(17) Goloubinoff, P.; Christeller, J. T.; Gatenby, A. A.; Lorimer, G. H. *Nature* **1989**, *342*, 884.
(18) Fayet, O.; Ziegelhoffer, T.; Georgopoulos, C. *J. Bacteriol.* **1989**, *171*, 1379.
(19) Trent, J. D.; Nimmesgern, E.; Wall, J. S.; Hartl, F.-U.; Horwich, A. L. *Nature* **1991**, *354*, 490.
(20) Ursic, D.; Culbertson, M. R. *Mol. Cell. Biol.* **1991**, *11*, 2629.
(21) Williams, D. C.; Van Frank, R. M.; Muth, W. L.; Burnett, J. P. *Science* **1982**, *215*, 687.
(22) Pelham, H. R. B. *EMBO J.* **1984**, *3*, 3095.
(23) Lewis, M. J.; Pelham, H. R. B. *EMBO J.* **1985**, *4*, 3137.
(24) Braig, K.; Otwinowski, Z.; Hegde, R.; Boisvert, D. C.; Joachimiak, A.; Horwich, A. L.; Sigler, P. B. *Nature* **1994**, *371*, 578.
(25) Saibil, H. R.; Zheng, D.; Roseman, A. M.; Hunter, A. S.; Watson, G. M. F.; Chen, S.; auf der Mauer, A.; O'Hara, B. P.; Wood, S. P.; Mann, N. H.; Barnett, L. K.; Ellis, R. J. *Curr. Biol.* **1993**, *3*, 265.
(26) Burnett, B. P.; Horwich, A. L.; Low, K. B. *J. Bacteriol.* **1994**, *176*, 6980.
(27) Thiyagarajan, P.; Henderson, S. J.; Joachimiak, A. *Structure* **1996**, *4*, 79.
(28) Hunt, J. F.; Weaver, A. J.; Landry, S. J.; Gierasch, L.; Deisenhofer, J. *Nature* **1996**, *379*, 37.
(29) Mande, S. C.; Mehra, V.; Bloom, B. R.; Hol, W. B. *Science* **1996**, *271*, 203.
(30) Xu, Z.; Horwich, A. L.; Sigler, P. B. *Nature* **1997**, *388*, 741.
(31) Chaudhry, C.; Farr, G. W.; Todd, M. J.; Rye, H. S.; Brunger, A. T.; Adams, P. D.; Horwich, A. L.; Sigler, P. B. *EMBO J.* **2003**, *22*, 4877.
(32) Wang, J.; Herman, C.; Tipton, K. A.; Gross, C. A.; Weissman, J. S. *Cell* **2002**, *111*, 1027.

(33) Langer, T.; Pfeifer, G.; Martin, J.; Baumeister, W.; Hartl, F.-U. *EMBO J*. **1992**, *11*, 4757.
(34) Braig, K.; Simon, M.; Furuya, F.; Hainfeld, J. F.; Horwich, A. L. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3978.
(35) Fenton, W. A.; Kashi, Y.; Furtak, K.; Horwich, A. L. *Nature* **1994**, *371*, 614.
(36) Farr, G. W.; Furtak, K.; Rowland, M. B.; Ranson, N. A.; Saibil, H. R.; Kirchhausen, T.; Horwich, A. L. *Cell* **2000**, *100*, 561.
(37) Buckle, A. M.; Zahn, R.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 3571.
(38) Chen, L.; Sigler, P. B. *Cell* **1999**, *99*, 757.
(39) Shewmaker, F.; Maskos, K.; Simmerling, C.; Landry, S. J. *J. Biol. Chem*. **2001**, *276*, 31257.
(40) Martin, J.; Langer, T.; Boteva, R.; Schramel, A.; Horwich, A. L.; Hartl, F. U. *Nature* **1991**, *352*, 36.
(41) Hlodan, R.; Tempst, P.; Hartl, F. U. *Nat. Struct. Biol*. **1995**, *2*, 587.
(42) Zahn, R.; Spitzfaden, C.; Ottiger, M.; Wüthrich, K.; Plückthun, A. *Nature* **1994**, *368*, 261.
(43) Robinson, C. V.; Gross, M.; Eyles, S. J.; Ewbank, J. J.; Mayhew, M.; Hartl, F. U.; Dobson, C. M.; Radford, S. E. *Nature* **1994**, *372*, 646.
(44) Gervasoni, P.; Gehrig, P.; Plückthun, A. *J. Mol. Biol*. **1998**, *275*, 663.
(45) Gross, M.; Robinson, C. V.; Mayhew, M.; Hartl, F. U.; Radford, S. E. *Protein Sci*. **1996**, *5*, 2506.
(46) Goldberg, M. S.; Zhang, J.; Sondek, S.; Matthews, C. R.; Fox, R. O.; Horwich, A. L. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 1080.
(47) Chen, J.; Walter, S.; Horwich, A. L.; Smith, D. L. *Nat. Struct. Biol*. **2001**, *8*, 721.
(48) Horst, R.; Bertelsen, E.; Fiaux, J.; Wider, G.; Horwich, A. L.; Wüthrich, K. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13715.
(49) Ranson, N. A.; Dunster, N. J.; Burston, S. G.; Clarke, A. R. *J. Mol. Biol*. **1995**, *250*, 581.
(50) Peralta, F.; Hartman, D. J.; Hoogenraad, J. J.; Høj, P. B. *FEBS Lett*. **1994**, *339*, 45.
(51) Zahn, R.; Plückthun, A. *J. Mol. Biol*. **1994**, *242*, 165.
(52) Walter, S.; Lorimer, G. H.; Schmid, F. X. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 9425.
(53) Zahn, R.; Perrett, S.; Fersht, A. R. *J. Mol. Biol*. **1996**, *261*, 43.
(54) Chandrasekhar, G. N.; Tilly, K.; Woolford, C.; Hendrix, R.; Georgopoulos, C. *J. Biol. Chem*. **1986**, *261*, 12414.
(55) Gray, T. E.; Fersht, A. R. *FEBS Lett* **1991**, *292*, 254.
(56) Bochkareva, E. S.; Lissin, N. M.; Flynn, G. C.; Rothman, J. E.; Girshovich, A. S. *J. Biol. Chem*. **1992**, *267*, 6796.
(57) Yifrach, O.; Horovitz, A. *Biochemistry* **1995**, *34*, 5303.
(58) Ranson, N. A.; Farr, G. W.; Roseman, A. M.; Gowen, B.; Fenton, W. A.; Horwich, A. L.; Saibil, H. R. *Cell* **2001**, *107*, 869.
(59) Jackson, G. S.; Staniforth, R. A.; Halsall, D. J.; Atkinson, T.; Holbrook, J. J.; Clarke, A. R.; Burston, S. G. *Biochemistry* **1993**, *32*, 2554.
(60) Cliff, M. J.; Kad, N. M.; Hay, N.; Lund, P. A.; Webb, M. R.; Burston, S. G.; Clarke, A. R. *J. Mol. Biol*. **1999**, *293*, 667.
(61) Inobe, T.; Makio, T.; Takasu-Ishikawa, E.; Terada, T. P.; Kuwajima, K. *Biochim. Biophys. Acta* **2001**, *1545*, 160.
(62) Horovitz, A.; Fridmann, Y.; Kafri, G.; Yifrach, O. *J. Struct. Biol*. **2001**, *135*, 104.
(63) Brinker, A.; Pfeifer, G.; Kerner, M. J.; Naylor, D. J.; Hartl, F. U.; Hayer-Hartl, M. *Cell* **2001**, *107*, 223.
(64) Weissman, J. S.; Rye, H. S.; Fenton, W. A.; Beechem, J. M.; Horwich, A. L. *Cell* **1996**, *84*, 481.
(65) Rye, H. S.; Burston, S. G.; Fenton, W. A.; Beechem, J. M.; Xu, Z.; Sigler, P. B.; Horwich, A. L. *Nature* **1997**, *388*, 792.
(66) Rye, H. S.; Roseman, A. M.; Chen, S.; Furtak, K.; Fenton, W. A.; Saibil, H. R.; Horwich, A. L. *Cell* **1999**, *97*, 325.
(67) Fridmann, Y.; Kafri, G.; Danziger, O.; Horovitz, A. *Biochemistry* **2002**, *41*, 5938.
(68) Weissman, J. S.; Kashi, Y.; Fenton, W. A.; Horwich, A. L. *Cell* **1994**, *78*, 693.
(69) Todd, M.; Viitanen, P. V.; Lorimer, G. H. *Science* **1994**, *265*, 659.
(70) Smith, K. E.; Fisher, M. T. *J. Biol. Chem*. **1995**, *270*, 21517.
(71) Taguchi, H.; Yoshida, M. *FEBS Lett*. **1995**, *359*, 195.
(72) Mendoza, J. A.; Rogers, E.; Lorimer, G. H.; Horowitz, P. M. *J. Biol. Chem*. **1991**, *266*, 13044.
(73) Weissman, J. S.; Hohl, C. M.; Kovalenko, O.; Kashi, Y.; Chen, S.; Braig, K.; Saibil, H. R.; Fenton, W. A.; Horwich, A. L. *Cell* **1995**, *83*, 577.
(74) Farr, G. W.; Fenton, W. A.; Chaudhuri, T. K.; Clarke, D. K.; Saibil, H. R.; Horwich, A. L. *EMBO J*. **2003**, *22*, 3220.
(75) Bigay, J.; Deterre, P.; Pfister, C.; Chabre, M. *EMBO J*. **1987**, *6*, 2907.
(76) Viitanen, P. V.; Lubben, T. H.; Reed, J.; Goloubinoff, P.; O′;Keefe, D. P.; Lorimer, G. H. *Biochemistry* **1990**, *29*, 5665.
(77) Ranson, N. A.; Clare, D. K.; Farr, G. W.; Houldershaw, D.; Horwich, A. L.; Saibil, H. R. *Nat. Struct. Mol. Biol*. **2006**, *13*, 147.
(78) Motojima, F.; Chaudhry, C.; Fenton, W. A.; Farr, G. W.; Horwich, A. L. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 15005.
(79) Shtilerman, M.; Lorimer, G. H.; Englander, S. W. *Science* **1999**, *284*, 822.
(80) Park, E. S.; Fenton, W. A.; Horwich, A. L. *FEBS Lett*. **2005**, *579*, 1183.
(81) Landry, S. J.; Zeilstra-Ryalls, J.; Fayet, O.; Georgopoulos, C.; Gierasch, L. M. *Nature* **1993**, *364*, 255.
(82) Miyazaki, T.; Yoshimi, T.; Furutsu, Y.; Hongo, K.; Mizobata, T.; Kanemori, M.; Kawata, Y. *J. Biol. Chem*. **2002**, *277*, 50621.
(83) Ma, J.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 8502.
(84) Ma, J.; Sigler, P. B.; Xu, Z.; Karplus, M. *J. Mol. Biol*. **2000**, *302*, 303.
(85) Ueno, T.; Taguchi, H.; Tadakuma, H.; Yoshida, M.; Funatsu, T. *Mol. Cell* **2004**, *14*, 423.
(86) Lin, Z.; Rye, H. S. *Mol. Cell* **2004**, *16*, 23.
(87) Yifrach, O.; Horowitz, A. *Biochemistry* **1998**, *37*, 7083.
(88) Kad, N. M.; Ranson, N. A.; Cliff, M. J.; Clarke, A. R. *J. Mol. Biol*. **1998**, *278*, 267.
(89) Burston, S. G.; Ranson, N. A.; Clarke, A. R. *J. Mol. Biol*. **1995**, *249*, 138.
(90) Cheng, M.-Y.; Hartl, F.-U.; and Horwich, A. L. *Nature* **1990**, *348*, 455.
(91) Glick, B. S.; Brandt, A.; Cunningham, K.; Muller, S.; Hallberg, R. L.; Schatz, G. *Cell* **1992**, *69*, 809.
(92) Ditzel, L.; Lowe, J.; Stock, D.; Stetter, K. O.; Huber, H.; Huber, R.; Steinbacher, S. *Cell* **1998**, *93*, 125.
(93) Horovitz, A.; Willison, K. R. *Curr. Opin. Struct. Biol*. **2005**, *15*, 646.
(94) Bhutani, N.; Udgaonkar, J. B. *J. Mol. Biol*. **2001**, *314*, 1167.
(95) Kerner, M. J.; Naylor, D. M.; Ishihama, Y.; Maier, T.; Chang, H.-C.; Stines, A. P.; Georgopoulos, C.; Frishman, D.; Hayer-Hartl, M.; Mann, M.; Hartl, F. U. *Cell* **2005**, *122*, 209.
(96) Saibil, H. R. and Ranson, N. A. *Trends Biochem. Sci*. **2002**, *27*, 627.

CR040435V