doi:10.1016/j.jmb.2003.09.074 J. Mol. Biol. (2003) 334, 489–499

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Domain Motions in GroEL upon Binding of an Oligopeptide

Jimin Wang[1]* and Lingling Chen[2]

[1]*Department of Molecular Biophysics and Biochemistry Yale University, 266 Whitney Avenue, Bass Center, Rm 418 New Haven, CT 06520-8114 USA*

[2]*Department of Biology Indiana University, 915 East Third Street, Bloomington, IN 47405, USA*

**Corresponding author*

GroEL assists protein folding by preventing the interaction of partially folded molecules with other non-native proteins. It binds them, sequesters them, and then releases them so that they can fold in an ATP-driven cycle. Previous studies have also shown that protein substrates, GroES, and oligopeptides bind to partially overlapped sites on the apical domain surfaces of GroEL. In this study, we have determined the crystal structure at 3.0 Å resolution of a symmetric $(GroEL–peptide)_{14}$ complex. The binding of each of these small 12 amino acid residue peptides to GroEL involves interactions between three adjacent apical domains of GroEL. Each peptide interacts primarily with a single GroEL subunit. Residues R231 and R268 from adjacent subunits isolate each substrate-binding pocket, and prevent bound substrates from sliding into adjacent binding pockets. As a consequence of peptide binding, domains rotate and inter-domain interactions are greatly enhanced. The direction of rotation of the apical domain of each GroEL subunit is opposite to that of its intermediate domain. Viewed from outside, the apical domains rotate clockwise within one GroEL ring, while the ATP-induced apical domain rotation is counter-clockwise.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: domain motions; intra-ring negative allostery; opposite allosteric effectors; symmetric GroEL complex

Introduction

The bacterial chaperonin GroEL assists in the folding of a large number of proteins in *Escherichia coli*.[1,2] By binding and sequestering its unfolded substrates, GroEL keeps them away from interacting with other non-native substrates, and it then releases its substrates for refolding in an ATP-dependent manner. In a GroEL-deficient cell line, about 30% of *E. coli* proteins are misfolded, and purified GroEL interacts with about 50% of soluble *E. coli* proteins when they are in non-native states.[3,4] *In vivo* GroEL shows a strong preference for newly synthesized proteins that contain multiple αβ-domains and have molecular masses in the range of 20–60 kDa.[5,6] Many substrates have hydrophobic cores containing β-sheets.[5,6]

The bacterial chaperone GroEL is the structurally and biochemically most extensively studied molecule of its class.[1] It consists of 14 subunits that possess *D*7*2* symmetry,[7] and each subunit has an equatorial, an intermediate, and an apical domain. ATP binds to the equatorial domain.[8] Protein substrates bind to the apical domain surface,[9–14] as does GroES.[15] The bacterial co-chaperonin GroES is a dome-like, heptameric molecule.[16,17] It binds to one end of GroEL to form an asymmetric GroEL–GroES complex that has a large internal cavity.[15,18] It can also bind to both ends of GroEL to form a symmetric complex,[19–22] which appears to be a storage form for excess GroEL and GroES with nucleotides trapped and no accessible binding site for protein substrates.

The affinity of GroEL for protein substrate and ATP varies during the course of a reaction cycle.[2,23,24] The reciprocal interplay of these two affinities is critical to catalysis and is highly regulated. For example, a typical cycle includes a state in which protein substrate is bound with high affinity in the absence of ATP, as well as a state in which the binding of ATP causes a decrease in affinity for protein substrate and results in its release.[24] Concurrently with the latter step and GroES binding, a large internal cavity is generated[15,18] inside which the newly released

Abbreviation used: CM, center of mass.

E-mail address of the corresponding author: wang@mail.csb.yale.edu

0022-2836/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.

protein substrate can refold.[25–27] Affinity regulation thus ensures that binding events occur in an ordered and productive fashion.

Affinity regulation also efficiently couples GroEL-catalyzed ATP hydrolysis to the protein folding reaction.[2,23,28] In the absence of ATP, GroEL can only fold some protein substrates, which appear to be low-affinity substrates with an equilibrium constant for bound and unbound protein close to 1. However, GroEL is rapidly inactivated by the presence of a higher-affinity protein substrate in the absence of ATP; the release of such a substrate requires additional energy. Conversely, the rate of the GroEL-catalyzed ATP hydrolysis is slow in the absence of protein substrate, and is stimulated upon its binding, for example, over 20-fold upon binding of unfolded lactate dehydrogenase.[29]

We previously proposed a structure-based mechanism[30] on the basis of a high-resolution GroEL–KMgATP complex structure and existing biochemical data.[23] The binding of ATP to a GroEL subunit which already has the protein substrate bound is energetically unfavorable. ATP will preferentially bind to the other six free subunits within the same ring. The presence of ATP in the six free subunits through intra-ring allosteric interactions[23,28] permits the binding of ATP by the GroEL subunit with the bound protein substrate. Binding of ATP by the *trans* GroEL ring is prevented through inter-ring allosteric interactions.[23,28] Once ATP is bound by all seven subunits in the substrate-containing ring, a conformational change occurs that switches the protein substrate binding subunit into a form with low affinity for substrate.[23,24] This leads to the release of the bound protein substrate.

Less is known about inter-subunit allosteric communications in GroEL at the atomic level when it binds a protein substrate. Biochemical evidence suggests such communications do exist, because the binding of a protein substrate to its first site on GroEL down-regulates its binding to a second site.[31,32] Here, we examine inter-subunit communications and domain rotations in GroEL upon binding of an oligopeptide. This is an extension of an early study,[11] which showed that the peptide binds to the apical domain surface of GroEL in the same way as it does to that of an isolated apical domain fragment. In this study, we observe that the domain rotations are clockwise viewed from outside and can be likened to the tightening of a screw-top cap on a bottle, which are opposite to the counter-clockwise rotations upon binding of ATP (the loosening of the cap).

Results and Discussion

Overall structure, peptide conformation, and GroEL–peptide interface

The (GroEL–peptide)$_{14}$ crystal structure, one complete complex per asymmetric unit, has been determined by molecular replacement to 3.0 Å resolution with a crystallographic *R*-factor of 23.6% and free *R*-factor of 26.0% (Table 1). The starting model was the structure of GroEL with KMgATP bound,[19] and molecular replacement was done at the domain level. Standard refinement procedures resulted in a model with a *R*-factor of 26.6% and a free *R*-factor of 28.2%. At that stage, both $2F_o - F_c$ and $F_o - F_c$ electron density maps calculated using unbiased model phases showed the well-defined density for bound peptides (Figure 1). Further refinement with the peptide included and the addition of ordered solvent molecules led to a final model with a free *R*-factor of 26.0%. From a Luzzati plot, the estimated coordinate error is 0.4 Å. A least-squares comparison shows that the root-mean-square C^α deviation (rmsd) is about 0.28 Å between the two GroEL rings in any 7-fold orientation (Figure 1B). Thus, the assembly strictly obeys the *D*7*2* symmetry, and the 14-fold redundancy should substantially lower the estimated coordinate errors and allow us to examine subtle global conformational changes.

One global effect of peptide binding to GroEL is a dramatic reduction in thermal motions in the apical domain, resulting in much better definition of the peptide–apical domain complex (Figure 1C and D). The crystallographic temperature factors for the apical domain in GroEL in this structure are reduced by over two fold with reference to the apo structure,[7,33] from above 125 Å^2 to below 50 Å^2, while those for the equatorial and intermediate domains are virtually unchanged. Thus, the apical domain electron density in this structure is better than in any other crystal structures of GroEL.[7,8,30,33] Therefore, we could visualize the details of interactions between the peptide substrate and GroEL quite well, even though the

Table 1. Summary of crystallographic data

Space group	$P2_1$
Unit cell parameters	
a (Å)	135.42
b (Å)	260.69
c (Å)	148.69
β (deg.)	100.94
X-ray data collection	
Resolution (Å)	20–3.0 (3.2–3.0)
Merging *R*-factor (%)	9.9 (14.7)
No. observations	436,111
No. reflections	167,654
Completeness (%)	84.4 (73.9)
Structure refinement	PDB accession 1MNF
Resolution (Å)	20–3.0 (3.2–3.0)
No. reflections	167,654 (22,811)
No. (10%) test reflections	16,677 (2484)
R-factor (%)	23.6 (30.4)
Test free *R*-factor (%)	25.9 (32.2)
No. protein atoms	54,096
No. substrate atoms	1456
No. solvent molecules	212
rmsd bond length (Å)	0.007
rmsd bond angle (deg.)	1.2

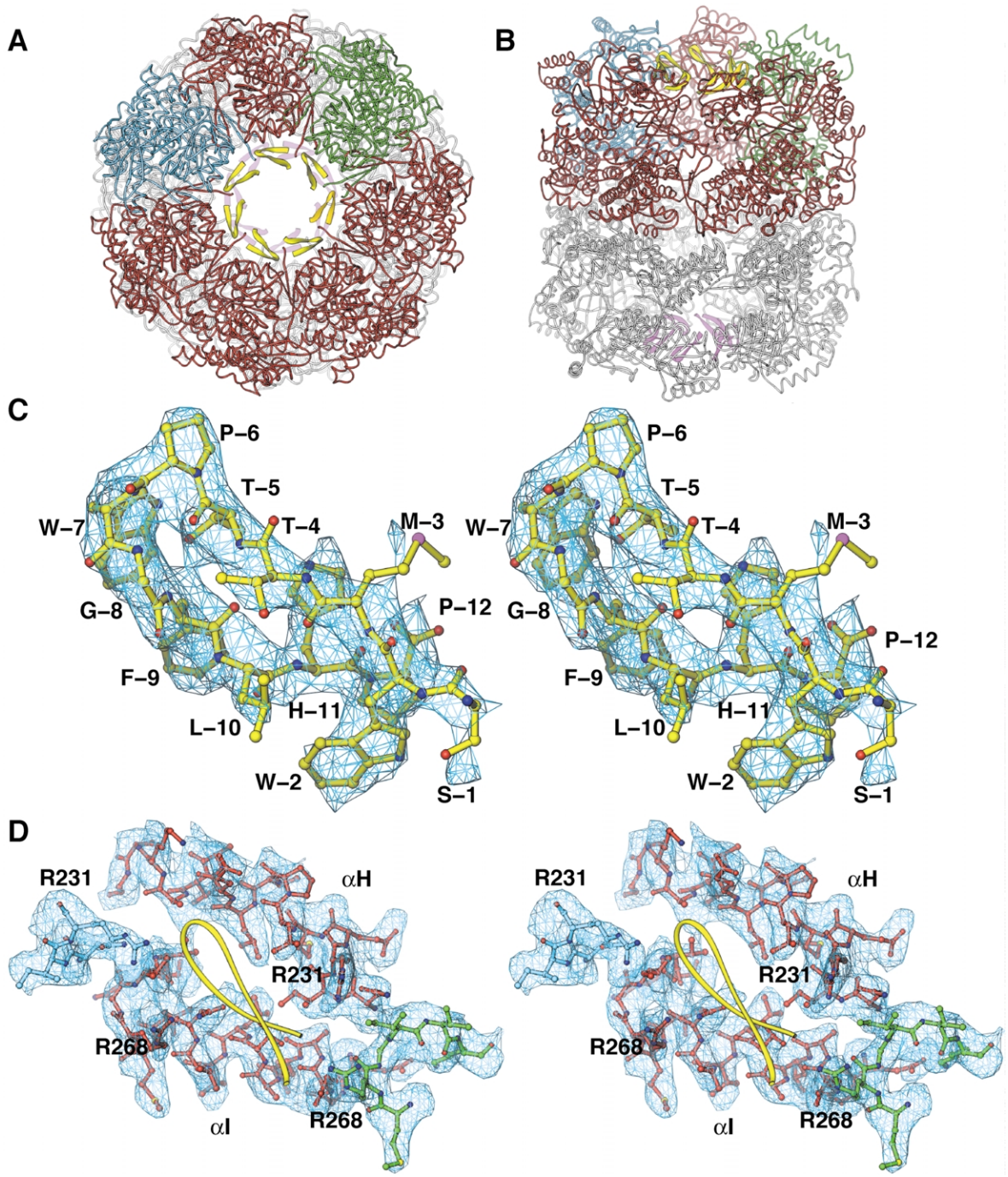

Figure 1. The crystal structure of the (GroEL–peptide)$_{14}$ complex. A, An overall view down the 7-fold axis. The bound peptide is in magenta in the lower GroEL ring (silver) and yellow in the upper ring (red). One GroEL subunit in the upper ring is in cyan, and another in green. B, An approximately orthogonal view (along *x*-axis/+80°). C, Stereo diagram of an unbiased $2F_o - F_c$ map, calculated using phases from a model before the peptide was built into the model, contoured at 1σ shows well-defined electron densities for the peptide which was imported[11] upon least-squares superposition using apical domain C^{α} atoms. Residue Ser1, which was absent from the previous model, was taken from the final refined model. D, An omit $2F_o - F_c$ map, using phases from a final refined model, contoured at 1.2σ shows well-defined electron density for helices H and I with the bound peptide in a yellow ribbon. Three subunits (cyan, red, and green) have the same orientation as in B. Electron density was removed beyond 3.0 Å of the displayed models.

resolution was limited to 3.0 Å. As expected, all interactions that are present in the isolated apical domain in complex with the peptide, for which an independent crystal structure exists,[11] are also present in this structure. In fact, the structure of the isolated apical domain complex perfectly fits into the corresponding region of an unbiased $2F_o - F_c$ map of this GroEL complex (Figure 1C).

When that structure and this one are superimposed using 146 C^{α} atoms, the rmsd value is 0.50 Å; it is 0.51 Å when the 12 C^{α} coordinates of the bound peptide are additionally included.

Overall, the bound peptide adopts a hairpin conformation at Pro6/Trp7/Gly8 with the remaining residues (Ser1 to Thr5 and Phe9 to Pro12) forming a two-stranded antiparallel β-sheet with six internal hydrogen bonds (Figure 2A), consistent with previous observations.[11] However, two new structural features are evident. First, in contrast to the complex with the isolated apical domain where two bound peptides interact with each other and bury 220 Å^2 of non-polar surface between them,[11] the peptides in this complex do not interact; each is segregated from its neighbors. Second, in this complex, Ser1, which was disordered in the isolated complex,[11] is now ordered (Figure 1C), and makes a hydrogen bond with R268 of a neighbor subunit through its backbone carbonyl oxygen atom (Figure 2B).

The extensive hydrophobic nature of the complementary interface surfaces of the apical domains and peptides (Figure 2) explains the observed high affinity of this peptide to GroEL.[11] In this structure, the peptide has a surface of total 1314 Å^2 (83 Å^2 per residue), 76% of which is contributed by non-polar atoms. Both the amount of exposed surface per residue and the fraction of atoms that are non-polar are larger than these in the case for the *Salmonella* virulence effector Sptp (59 Å^2 and 61%), whose conformation is suggested to be unfolded.[34] These values are also much larger than those characteristic of small proteins in the folded state, for example, 30 Å^2 and 55% for barnase[35] and 43 Å^2 and 62% for chymotrypsin inhibitor,[36] respectively. The portion of the peptide surface that interacts with GroEL is even more non-polar in nature (81%) than the rest of the peptide (76%). Moreover, the surface of GroEL and the surface of peptide fit each other very well with a high shape similarity index of 0.71.[37]

Multi-valency in binding of peptide and its implication for binding of protein substrates

One structural feature this complex reveals, but the isolated apical domain complex structure does not,[11] is that bound peptides interact with three adjacent apical domains within each 7-fold symmetric ring (Figures 1 and 2). Of the three apical domains involved, the central one provides most of the contact surface and displays all interactions seen previously in the isolated apical domain complex structure.[11] The apical domains on either side of the central domain provide additional contacts. These contacts are dominated by R231 on one side and R268 on the other (Figure 2B). R268 and R231

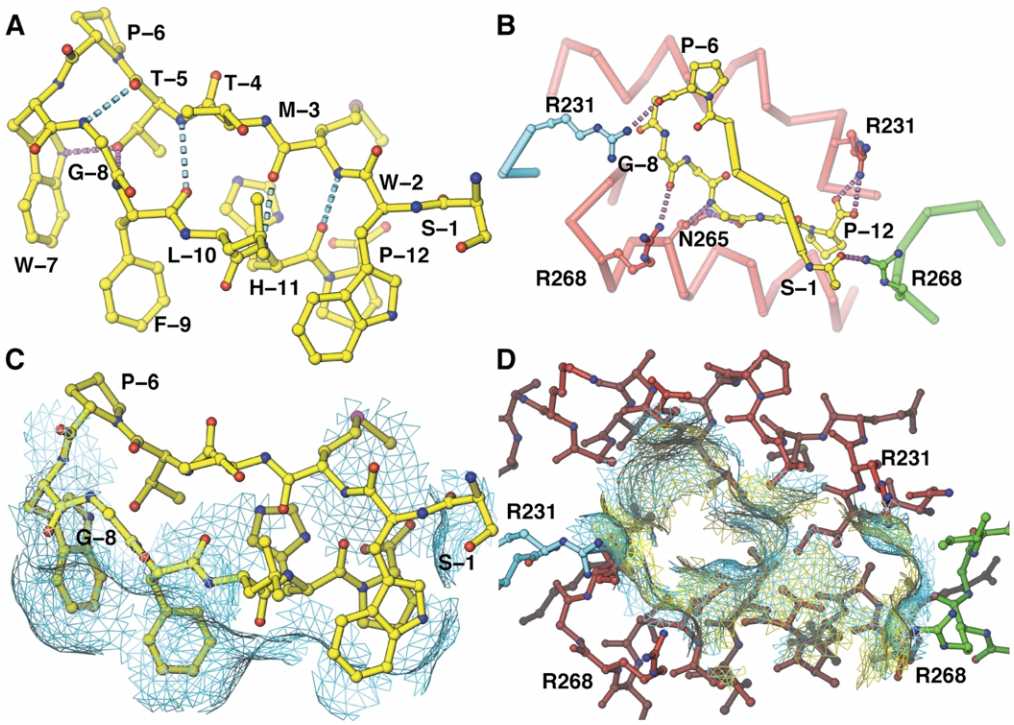

Figure 2. Nature of GroEL–peptide interactions. A, The structure of the bound peptide hairpin includes four backbone hydrogen bonds (cyan dashes) and two backbone/side-chain hydrogen bonds (magenta dashes). B, Four arginine and one asparagine residue from three adjacent GroEL subunits (cyan, red, and green) form seven hydrogen bonds with backbone carbonyl and C-terminal carboxylate oxygen atoms. C, The same strand (residues 7–12) that makes most backbone hydrogen bonds to GroEL also makes most hydrophobic interactions with GroEL. Solvent-accessible surfaces of the bound peptide within 2.5 Å of any GroEL atoms are shown in cyan mesh. D, Two surfaces (GroEL in cyan, and peptide in yellow) are largely complementary to each other.

from three adjacent apical domains make five hydrogen bonds to each peptide: two at the Pro12 carboxyl terminus, one at the Ser1 amino terminus, one at each carbonyl group of Pro6 and Gly8 at the hairpin turn (Figure 2B). Among them, two hydrogen bonds are from two neighboring subunits (R268/Ser1 and R231/Pro6; Figure 2B). Additionally, Asn265 makes two hydrogen bonds with the peptide. Because of the presence of ordered arginine residues at its edge, the binding pocket appears to be deeper and more segregated than it is in the isolated apical domain.[11] Each binding pocket now has a surface area of 300–400 Å^2, a diameter of about 20 Å, and an average depth of 8–10 Å. The distance between the centers of two adjacent pockets is about 18–21 Å.

The observed binding of peptide to GroEL has some implications on the binding of protein substrates, even though this peptide is not a typical protein substrate. Using a probe attached to Ser1, we have previously shown that the fluorescently labeled peptide binds with a similar affinity to either intact GroEL or an isolated apical domain fragment.[11] Ser1 was chosen, since it was disordered in the apical domain complex, and it was assumed that it would play no role in binding intact GroEL. However, the current structure shows otherwise: Ser1 plays a key role in enhancing apical–apical domain interactions. Modification of this residue would clearly disrupt the ability of this peptide to induce such interactions. Ser1 was also chosen as the site of modification in another set of experiments[38] in which a mutant (Asn229C) GroEL was disulfide-linked to peptide through an *N*-hydroxysuccinimide maleoyl-β-alaninate moiety. Asn229 in our current structure is buried at an apical–apical domain interface, and Ser1/Asn229 are separated by a C$^\alpha$–C$^\alpha$ distance of 14.8 Å. On the basis of cross-linking experiments it was suggested[38] that the mode of binding by this modified peptide mimicked GroEL–GroES interactions more than GroEL–protein substrate interactions. However, in the context of this structure, the unmodified peptide seems to provide a valid model for substrate binding for a number of reasons, and this complex provide some clues as to how GroEL binds the protein substrates. First, no sequence-specific interactions are observed between the peptide and GroEL. None of the peptide's side-chains, for example, is involved in hydrogen bonds, and all observed hydrogen bonds involve only backbone carbonyl and amide groups of the peptide. This is a basis for the promiscuity of GroEL in binding to a large number of protein substrates. Second, many other peptides found in the selection experiments can bind to GroEL with a variety of affinities.[11] Among them, non-polar residues such as Ile, Val, and Tyr were often found in the positions corresponding to Trp7 and Phe9 of this peptide. These peptides will likely bind to GroEL in the same manner as observed for this peptide in this complex. To explain why peptides with different hydrophobic residues at Trp7 and Phe9 can bind to the apical domain of GroEL, we compared this structure with the isolated apical domain complexes. We show that the ends of helices H and I where the two residues bind have a large conformational variation with the largest C$^\alpha$ displacement of 2.75 Å at Gly269. This variation has been used to accommodate non-polar residues that vary in size.[11,38] Third, many α/β-domain protein substrates contain β-sheets in their hydrophobic cores,[5,6] including a similar distribution of hydrophobic residues along the chain as in this peptide. Therefore, the binding of this peptide in this complex represents at least one such class of protein substrates. Lastly, subunit-specific mutations show that proper binding of two of the three test protein substrates to GroEL requires three adjacent functionally active apical domain surfaces.[39] The observed binding of each peptide to three apical domain surfaces suggests why it is so.

Peptide-induced domain rotations and displacements

Comparison of this substrate complex structure with that of the apo-GroEL[7,33] shows that the substrate binding leads to relative rotations of the three domains (Figure 3). For example, when the equatorial domains are superimposed between the two structures using the domain-averaged coordinates, it becomes obvious that the intermediate domains rotate about 5.2°. However, the rotation of the apical domain between the two structures is less, only about 3.4°, rather than 5.2°, indicating the existence of a relative motion between intermediate and apical domains. Indeed, when the intermediate domains are superimposed between the two structures, the apical domain rotates by 1.9°. The direction of the apical domain rotation with respect to the intermediate domain is opposite to the direction of the intermediate domain with respect to the equatorial domain, leading to some cancellation of rotations in the apical domains (Figure 3(B)). An analysis of domain rotation using individual subunits between the two structures consistently shows the same results, but with slightly larger variations (Figure 3).

The direction of peptide-induced apical domain rotations is clockwise (Figure 4), viewed from outside of the ring assembly, which is opposite to the ATP-induced apical domain rotations. The direction of intermediate domain rotations is counterclockwise in both cases. Thus, relative domain rotations in individual subunits propagate to inter-subunit interactions in the ring assembly.

There are also subtle and important linear domain movements as measured using domain centers of mass (CM) upon binding of peptide. Compared with the apo-GroEL structure,[7,33] the distances of CM to the 7-fold axis in this structure is decreased by 0.68 Å, 0.66 Å, and 0.27 Å for the apical, intermediate and equatorial domains,

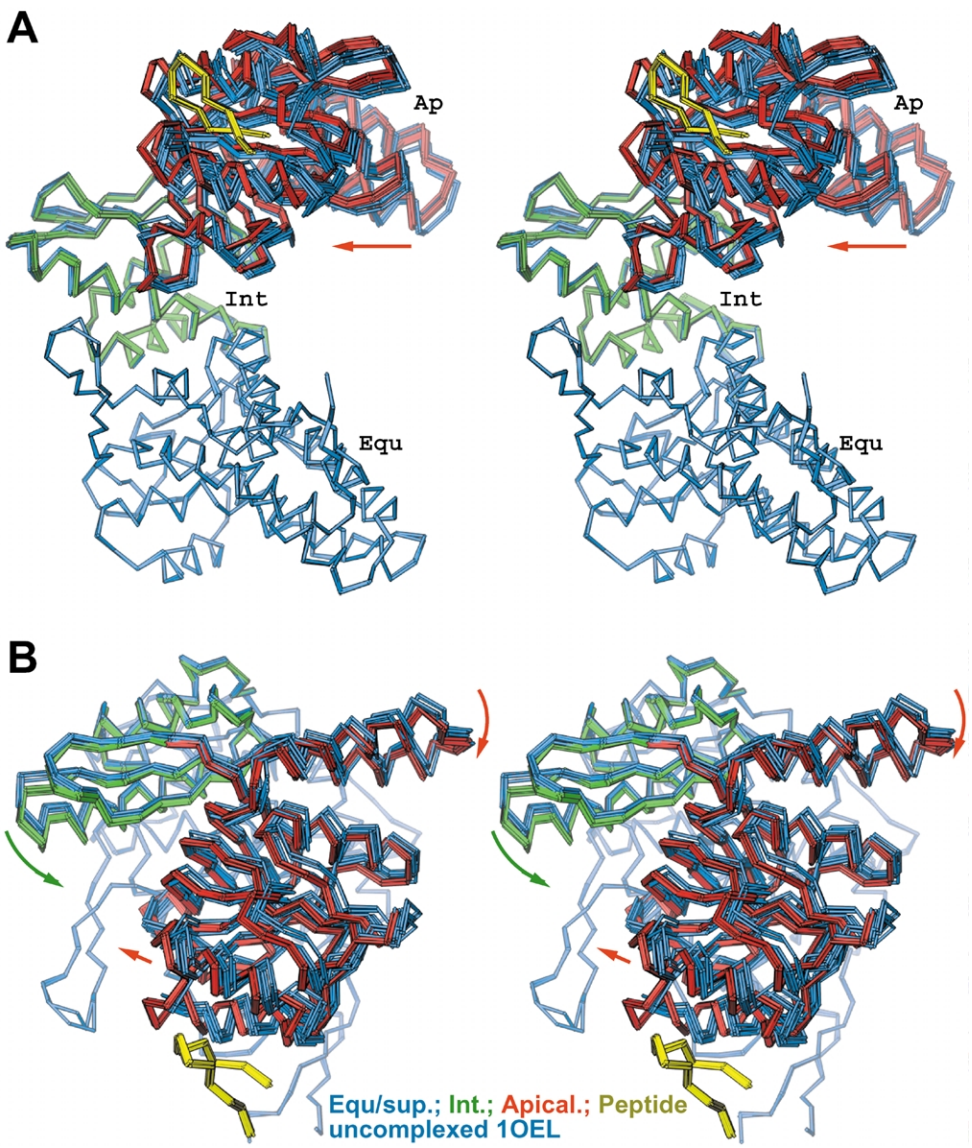

Figure 3. Peptide-induced domain rotations. A, Subunit comparison in a view approximately perpendicular to the 7-fold axis. B, View along the 7-fold axis. All subunits are superimposed using equatorial domains (cyan). Subunits in the apo structure[33] are in cyan. Domains of subunits in this structure are green and red for the intermediate and apical domains, respectively. Red and green arrows indicate the rotations of the apical and intermediate domains in this structure.

respectively. Additionally, the apical domains are displaced away from the opposite ring by 0.44 Å, while the intermediate and equatorial domains are displaced towards the opposite ring by 0.28 Å and 0.25 Å, respectively.

Both rotational and linear domain motions observed in this structure in comparison with the apo-GroEL structure[7,33] are consistent with the fact that the bound peptide enhances the inter-subunit domain–domain interactions. As a result, the apical domains become highly ordered as shown in the well-defined electron density in this structure. On the other hand, the apical domains are partially free upon binding of ATP.[30] When rotational and linear domain motions are combined, the observed clockwise apical domain motions in this structure with respect to the apo-GroEL structure[7,33] can be likened to the tightening of a screw-top cap on a bottle, while the previously observed counter-clockwise apical domain rotations upon binding of ATP[30] can then be likened to loosening it (Figure 5). The two precisely opposite domain motions imply that the ATPase and protein refolding activities of GroEL can be regulated by each other through binding, consistent with known biochemical data.[23,24,28,29]

Because of limited space available, the extent of the movements of subunits towards the 7-fold axis upon binding of peptide is limited in this structure. We predict that in the sub-stoichiometric situations such as 1:14 peptide/GroEL ratio, the movement of three apical domains that interact with the bound single peptide would be much larger than what has been observed here, leading to a highly

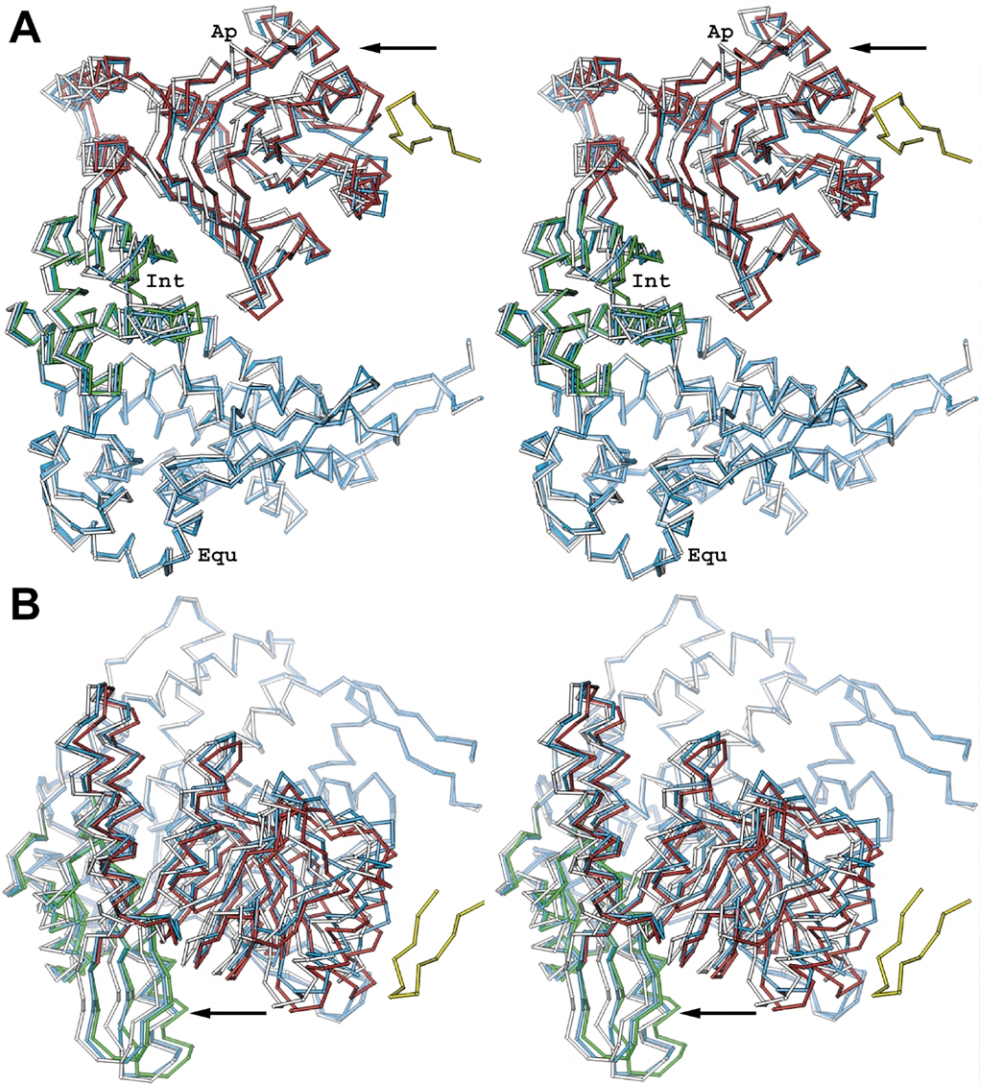

Figure 4. Peptide-induced domain rotations are distinct from ATP-induced ones. A, View nearly perpendicular to the 7-fold axis. B, View along the 7-fold axis. Averaged coordinates are used for the apo structure[33] (cyan), the KMgATP bound structure[30] (silver), and this complex structure (apical domain, red; intermediate domain, green; and equatorial domain, cyan) with the bound peptide in yellow. Arrows indicate the rotations of both intermediate and apical domains upon binding of ATP.

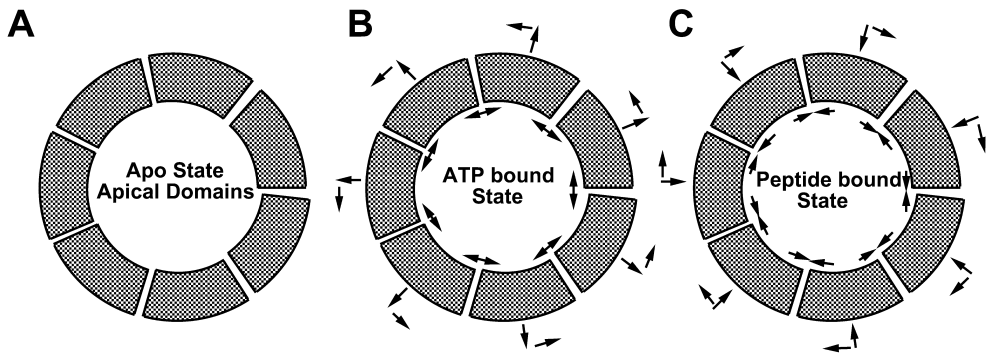

Figure 5. Schematic drawing of the observed motions in the apical domains. A, The apical domains in the apo state. B, The motions of the apical domains upon binding of ATP, leading to the liberation of individual apical domains. C, The motions of the apical domains upon binding of peptide, leading to the enhanced inter-domain interactions. Arrows indicate the directions of domain motions.

asymmetric ring structure. We also predict that the same type of domain motions occur when protein substrates bind.

Allostery in peptide binding and implication for GroEL-assisted protein folding

The observed peptide-induced domain rotations suggest a new allosteric property of GroEL for binding of protein substrates within one ring. Within one GroEL ring, the mechanics of subunit allosteric communications and domain rotations have been addressed elsewhere.[30] With respect to the reference equatorial domain, the rotation of the remote apical domain may occur in the same or opposite direction as the rotation of the immediately adjacent intermediate domain. When in the same direction as in the case of ATP binding, the communication is positive allosteric,[30] when in the opposite direction as in this complex upon binding of peptide, it is negative (Figure 5).

The observed peptide-induced equatorial domain displacements suggest a new allosteric property of GroEL for binding of protein substrates in two rings. Sigler and colleagues suggest that inter-ring communications are carried out through ring puckering on the basis of the asymmetric GroEL–GroES complex structure.[15] When the nucleotide and GroES bind to the *cis* GroEL ring, this ring is puckered relative to the *trans* ring, thus preventing ATP and GroES from binding to the *trans* ring. We observed in this complex that the center of the equatorial domain was displaced towards the opposite ring. We predict that this displacement would cause similar ring puckering when peptides are not bound in the second ring in the asymmetric $(GroEL)_{14}(peptide)_7$ complex. This ring puckering would discourage the binding of protein substrates in the opposite ring.

The two new GroEL allosteric properties for the binding of protein substrates are also consistent with the existing biochemical data.[37] The simultaneous occupation of all 14 apical domain surfaces by 14 peptide substrates does not represent the binding mode of protein substrates to GroEL. Typically, a protein substrate may cooperatively bind to three adjacent apical domain surfaces,[39] excluding the possibility of more than two binding sites per GroEL ring.[31,32,40] In general, the binding affinity for a second protein substrate, provided by a second site, appears to be much lower than the one provided by the first site.[31] Depending on protein substrates, the difference in affinity between the two sites could be as large as several orders of magnitude. These observations have already implied the existence of negative allosteric properties of GroEL for binding of protein substrates, even though this has not been explicitly stated.

The observed domain motions upon binding of the substrate to GroEL can also explain how GroEL takes two different pathways, one an encapsulation and the other a non-encapsulation pathways, for the protein substrates of different size for maximal efficiency.[30,41,42] When a protein substrate of small or intermediate size binds to one or a few GroEL subunits (Figure 6), the apical domains of these subunits will move towards the central axis. The apical domains of the other unoccupied subunits will move away from it (Figure 6(B)). In doing so, the apical domains of unoccupied subunits begin to undergo rotations, before ATP binding, in the same directions that are induced upon binding of ATP. These GroEL subunits thus exhibit higher affinity for ATP than others in the opposite ring,[23,30] and they preferentially bind ATP. Therefore, small proteins take an encapsulation pathway inside the GroEL–GroES internal cavity for folding. This is an efficient pathway in which folding intermediates are completely

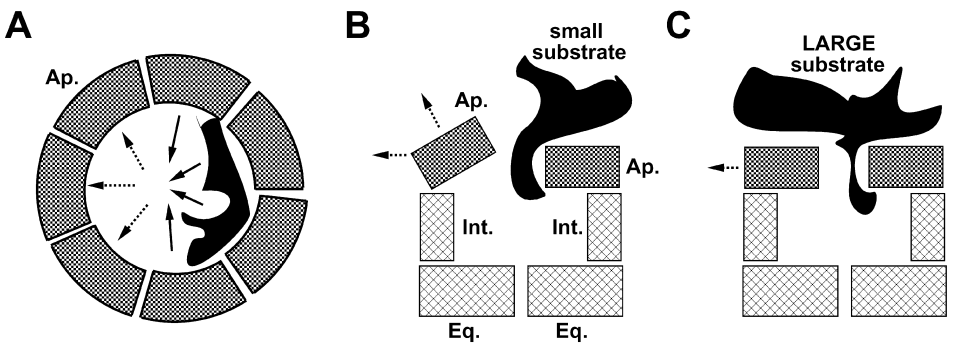

Figure 6. Implications of protein–substrate induced domain motions in protein folding mechanism. A, The binding of the protein substrate (the shaded object) may occur at three to four consecutive apical domains (right side). These domains would move into the central axis according to what has been observed here. They would push the remaining apical domains (left side) away from the center. B, When enough space is available for the movements of the apical domains such as the binding of a small protein substrate, the apical domains (left side) of subunits with no substrate bound will be free, and these subunits thus become a higher-affinity state than subunits in the opposite ring. C, When no space is available for the movements of the apical domains such as the binding of a large protein substrate, the movements of these apical domains are suppressed, and these subunits thus become a lower-affinity state than subunits in the opposite ring. Filled arrows indicate the directions of the observed domain motions; broken arrows indicate the directions of predicted domain motions.

sequestered from other non-native proteins. Small protein substrates may be forced to use a second folding pathway without encapsulation as shown in a *trans*-only GroEL–GroES complex after GroES is closely tethered to the *cis*-GroEL ring.[42] However, this is an inefficient folding pathway, because the folding intermediates are not sequestered and they may interact with other non-native proteins. When a protein substrate of large size binds to one or a few GroEL subunits in one ring, it may block possible movements of apical domains of the entire ring (Figure 6). In this case, none of GroEL subunits in the ring can complete domain rotations in the same directions that are induced upon binding of ATP. These GroEL subunits will now exhibit lower affinity for ATP than others are in the opposite ring.[30] ATP preferentially binds the opposite ring as the protein substrate, and the folding reaction takes place in the non-encapsulation pathway that are only available to large protein substrates. In summary, the size of the protein substrate becomes a determinant factor in the GroEL-assisted protein folding mechanism,[30] by controlling the mechanisms of domain motions.

Materials and Methods

Peptide binding assay, crystallization, and X-ray data processing

Wild-type GroEL was over-expressed in, and purified from *E. coli*.[7,15] A peptide substrate was selected using bio-panning experiments, and binding assays were carried out, as described.[11] The GroEL–peptide complex was crystallized under identical conditions as GroEL–(ATPγS),[8] except for the replacement of ATPγS with the peptide. The resulting crystals belonged to the same $P2_1$ space group as the GroEL–(ATPγS) crystals[8] and crystals had a similar unit cell $a = 135.42$ Å, $b = 260.69$ Å, $c = 148.69$ Å, $\beta = 100.94°$. X-ray diffraction data were collected at the Structural Biology Center ID19 beam line at Advanced Photon Source and the statistics are summarized in Table 1.

Structure determination and refinement

The peptide-bound GroEL structure was determined by the domain-wise molecular replacement using a GroEL/ATP crystal structure as the starting model.[30] Structure refinement included three stages of rigid-body refinement (one body per tetradecamer, one body per subunit, and then three bodies per subunit), followed by alternating Powell minimization and individual temperature *B*-factor refinement. All steps were executed using the program CNS.[43] During the refinement, each of the three domains for 14 GroEL subunits was restrained to be in a similar conformation. Any deviation from their averaged coordinates was penalized using the values of 300 kcal/mol per Å for positions and 2 kcal/mol per Å^2 for the *B*-factors. Attempts were made to use lower restraint values, but they did not result in improved refinement statistics and thus were aborted. During the refinement, it was concerned that the refined peptide-bound GroEL structure resembled the apo-GroEL structure[33] more closely than the KMgATP-bound GroEL structure used for the molecular replacement[30] in both overall subunit conformation and ring symmetry. We therefore ran a test refinement starting with the apo-GroEL structure[33] using nearly identical refinement protocols as just described,[11] resulting in a model with a free *R*-factor of 31%. This implies an accurate starting model is important for the success of structure refinement at that resolution.

Averaged coordinates and domain motion analysis

GroEL subunits are divided into the equatorial domains (residues 2–136/410–525), intermediate domains (residues 137–188/378–409), and apical domains (residues 189–377). In order to obtain an averaged subunit conformation in each of its three states (this structure, the apo-structure,[7,33] and the ATP-bound structure[8,30]), independent subunits were simultaneously superimposed using C^α coordinates of the equatorial domain residues, and the resulting coordinates averaged. Domain motion angles were then deduced from these averaged subunit conformations as described elsewhere.[44]

In order to understand conformational changes within ring assembly, each of the apo and ATP-bound GroEL heptamers was superimposed onto that of this structure using the C^α coordinates of all residues. The CM for each domain was calculated. The displacement of CM between the two structures represents a global motion of each domain. When CM moves towards the 7-fold axis, the ring contracts; when CM moves away from it, the ring expands. The movement of CM along the 7-fold axis is a relative displacement of the domain with respect to the opposite ring.

Figures and coordinates

Figures were made using the program Ribbons.[45] Coordinates and X-ray data are available from the RCSB database under the accession number 1MNF.

Acknowledgements

We thank Professor P.B. Moore for editing the manuscript and Drs J. Pata, S. Kamtekar, and G. Blaha for discussion. J.W. acknowledges funding from Yale University.

References

1. Sigler, P. B., Xu, Z., Rye, H., Burston, S. G., Fenton, W. A. & Horwich, A. L. (1998). Structure and function in GroEL-mediated protein folding. *Annu. Rev. Biochem.* **67**, 581–608.
2. Thirumalai, D. & Lorimer, G. H. (2001). Chaperonin-mediated protein folding. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 245–269.
3. Viitanen, P. V., Gatenby, A. A. & Lorimer, G. H. (1992). Purified chaperonin 60 (GroEL) interacts with the non-native states of a multitude of *Escherichia coli* proteins. *Protein Sci.* **1**, 363–369.
4. Horwich, A. L., Low, K. B., Fenton, W. A., Hirshfield, I. N. & Furtak, K. (1993). Folding *in vivo* of bacterial

cytoplasmic proteins: role of GroEL. *Cell*, **74**, 909–917.
5. Houry, W. A., Frishman, D., Echerskorn, C., Lottspeich, F. & Hartl, F. U. (1999). Identification of *in vivo* substrates of the chaperonin GroEL. *Nature*, **402**, 147–154.
6. Houry, W. A. (2001). Mechanism of substrate recognition by the chaperonin GroEL. *Biochem. Cell Biol.* **79**, 569–577.
7. Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L. & Sigler, P. B. (1994). The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature*, **371**, 578–586.
8. Boisvert, D. C., Wang, J., Otwinoski, Z., Horwich, A. L. & Sigler, P. B. (1996). The 2.4 Å crystal structure of the bacterial chaperonin GroEL complexed with ATPγS. *Nature Struct. Biol.* **3**, 170–177.
9. Fenton, W. A., Kashi, Y., Furtak, K. & Horwich, A. L. (1994). Residues in GroEL required for polypeptide binding and release. *Nature*, **371**, 614–619.
10. Buckle, A. M., Zahn, R. & Fersht, A. R. (1997). A structural model for GroEL-polypeptide recognition. *Proc. Natl Acad. Sci. USA*, **94**, 3571–3575.
11. Chen, L. & Sigler, P. B. (1999). The crystal structure of a GroEL/peptide complex: plasticity as a basis for substrate diversity. *Cell*, **99**, 757–768.
12. Wang, Q., Buckle, A. M. & Fersht, A. R. (2000). Stabilization of GroEL minichaperones by core and surface mutations. *J. Mol. Biol.* **298**, 917–926.
13. Wang, Q., Buckle, A. M. & Fersht, A. R. (2000). From minichaperone to GroEL 1: information on GroEL-polypeptide interactions from crystal packing of minichaperones. *J. Mol. Biol.* **304**, 873–881.
14. Zahn, R., Buckle, A. M., Perrett, S., Johnson, C. M., Corrales, F. J., Golbik, R. & Fersht, A. R. (1996). Catalysis of amide protein exchange by the molecular chaperones GroEL and SecB. *Proc. Natl Acad. Sci. USA*, **93**, 15024–15029.
15. Xu, Z., Horwich, A. L. & Sigler, P. B. (1997). The crystal structure of the asymmetric GroEL–GroES-(ADP)7 chaperonin complex. *Nature*, **388**, 741–750.
16. Hunt, J. F., Weaver, A. J., Landry, S. J., Gierasch, L. & Deisenhofer, J. (1996). The crystal structure of the GroES cochaperonin at 2.8 Å resolution. *Nature*, **379**, 37–45.
17. Mande, S. C., Mehra, V., Bloom, B. R. & Hol, W. G. J. (1996). Structure of the heat shock protein chaperonin-10 of *Mycobacterium laprae*. *Science*, **271**, 203–207.
18. Chen, S., Roseman, A. M., Hunter, A. S., Wood, S. P., Burton, S. G., Ranson, N. A. *et al.* (1994). Location of a folding protein and shape changes in GroEL–GroES complexes imaged by cryo-electron microscopy. *Nature*, **371**, 261–264.
19. Azem, A., Kessel, M. & Goloubinoff, P. (1994). Characterization of a functional $GroEL_{14}(GroES_7)_2$ chaperonin hetero-oligomer. *Science*, **265**, 653–656.
20. Schmidt, M., Rutkat, K., Rachel, R., Pfeifer, G., Jaenicke, R., Viitanen, P. *et al.* (1994). Symmetric complexes of GroE chaperonins as part of the functional cycle. *Science*, **265**, 656–659.
21. Sparrer, H., Rutkat, K. & Buchner, J. (1997). Catalysis of protein folding by symmetric chaperone complexes. *Proc. Natl Acad. Sci. USA*, **94**, 1096–1100.
22. Beibinger, M., Rutkat, K. & Buchner, J. (1999). Catalysis, commitment and encapsulation during GroE-mediated folding. *J. Mol. Biol.* **289**, 1075–1092.
23. Horovitz, A., Friedmann, Y., Kafri, G. & Yifrach, O. (2001). Review: allostery in chaperonins. *J. Struct. Biol.* **135**, 104–114.
24. Yifrach, O. & Horovitz, A. (1996). Allosteric control by ATP of non-folded protein binding to GroEL. *J. Mol. Biol.* **255**, 356–361.
25. Burston, S. G., Weissman, J. S., Farr, G. W., Fenton, W. A. & Horwich, A. L. (1996). Release of both native and non-native proteins from a *cis*-only GroEL ternary complex. *Nature*, **383**, 96–99.
26. Rye, H. S., Burston, S. G., Fenton, W. A., Beechem, J. M., Xu, Z., Sigler, P. B. & Horwich, A. L. (1997). Distinct actions of *cis* and *trans* ATP within the double ring of the chaperonin GroEL. *Nature*, **388**, 792–798.
27. Rye, H. S., Roseman, A. M., Chen, S., Furtak, K., Fenton, W. A., Saibil, H. R. & Horwich, A. L. (1999). GroEL–GroES cycling: ATP and non-native polypeptide direct alternation of folding-active rings. *Cell*, **97**, 325–338.
28. Yifrach, O. & Horovitz, A. (1995). Nested cooperativity in the ATPase activity of the oligomeric chaperonin GroEL. *Biochemistry*, **34**, 5303–5308.
29. Jackson, G. S., Staniforth, R. A., Halsall, D. J., Atkinson, T., Holbrook, J. J., Clarke, A. R. & Burston, S. G. (1993). Binding and hydrolysis of nucleotides in the chaperonie catalytic cycle: implications for the mechanism of assisted protein folding. *Biochemistry*, **32**, 2554–2563.
30. Wang, J. & Boisvert, D. C. (2003). Structural basis for GroEL-assisted protein folding from the crystal structure of $(GroEL–KMgATP)_{14}$ at 2.0 Å resolution. *J. Mol. Biol.* **327**, 843–855.
31. Makio, T., Arai, M. & Kuwajima, K. (1999). Chaperonin-affected refolding of a-lactalbumin: effects of nucleotides and the co-chaperonin GroES. *J. Mol. Biol.* **293**, 125–137.
32. Aoki, K., Taguchi, H., Shindo, Y., Yoshida, M., Ogasahara, K., Yutani, K. & Tanaka, N. (1997). Calorimetric observation of a GroEL–protein interaction with little contribution of hydrophobic interaction. *J. Biol. Chem.* **272**, 32158–32162.
33. Braig, K., Adams, P. & Brunger, A. T. (1995). Conformational variability in the refined structure of the chaperonin GroEL at 2.8 Å resolution. *Nature Struct. Biol.* **2**, 1083–1094.
34. Stebbins, C. E. & Galan, J. E. (2001). Maintenance of an unfolded polypeptide by a cognate chaperone in bacterial type III secretion. *Nature*, **414**, 77–81.
35. McPhalen, C. A. & James, M. N. (1987). Molecular structure of a new family of ribonucleases. *Biochemistry*, **26**, 261–269.
36. Mauguen, Y., Hartley, R. W., Dodson, E. J., Dodson, G. G., Bricogne, G., Chothia, C. & Jack, A. (1982). Crystal and molecular structure of the serine proteinase inhibitor CI-2 from barley seeds. *Nature*, **297**, 162–164.
37. Lawrence, M. C. & Colman, P. M. (1993). Shape complementarity at protein/protein interface. *J. Mol. Biol.* **234**, 946–950.
38. Ashcroft, A. E., Brinker, A., Coyle, J. E., Weber, F., Kaiser, M., Moroder, L. *et al.* (2003). Structural plasticity and non-covalent substrate binding in the GroEL apical domain. *J. Biol. Chem.* **277**, 33115–33126.
39. Farr, G. W., Furtak, K., Rowland, M. B., Ranson, N. A., Saibil, H. R., Kirchhausen, T. & Horwich, A. L. (2000). Multivalent binding of non-native substrate proteins by the chaperonin GroEL. *Cell*, **100**, 561–573.

40. Laminet, A. A., Ziegelhoffer, T., Georgopoulos, C. & Pluckthun, A. (1990). The *Escherichia coli* heat shock proteins GroEL and GroES modulate the folding of the beta-lactamase precursor. *EMBO J.* **9**, 2315–2319.
41. Chaudhuri, T. K., Farr, G. W., Fenton, W. A., Rospert, S. & Horwich, A. L. (2001). GroEL/GroES-mediated folding of a protein too large to be encapsulated. *Cell*, **107**, 235–246.
42. Farr, G. W., Fenton, W. A., Chaudhuri, T. K., Clare, D. K., Saibil, H. R. & Horwich, A. L. (2003). Folding with and without encapsulation by *cis*- and *trans*-only GroEL–GroES complexes. *EMBO J.* **22**, 3220–3230.
43. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grusse-Kunstleve, R. W. *et al.* (1998). Crystallography NMR system: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **50**, 905–921.
44. Wang, J., Song, J. J., Seong, I. S., Franklin, M. C., Kamtekar, S., Eom, S. H. & Chung, C. H. (2001). Nucleotide-dependent conformational changes in a protease-associated ATPase HslU. *Structure*, **9**, 1107–1116.
45. Carson, M. (1991). Ribbons 2.0. *J. Appl. Crystallog.* **24**, 958–961.

Edited by W. Baumeister

(Received 5 June 2003; received in revised form 18 September 2003; accepted 29 September 2003)