doi:10.1016/j.jmb.2005.02.027 J. Mol. Biol. (2005) 348, 219–230

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

The 13 Å Structure of a Chaperonin GroEL–Protein Substrate Complex by Cryo-electron Microscopy

Scott Falke[1], Florence Tama[2], Charles L. Brooks III[2], Edward P. Gogol[3]* and Mark T. Fisher[1]*

[1]*Department of Biochemistry and Molecular Biology University of Kansas Medical Center, Kansas City, KS 66160 USA*

[2]*Department of Molecular Biology (TPC6), The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla CA 92037, USA*

[3]*School of Biological Sciences University of Missouri-Kansas City, Kansas City, MO 64110 USA*

The 13 Å resolution structures of GroEL bound to a single monomer of the protein substrate glutamine synthetase (GS_m), as well as that of unliganded GroEL have been determined from a heterogeneous image population using cryo-electron microscopy (cryo-EM) coupled with single-particle image classification and reconstruction techniques. We combined structural data from cryo-EM maps and dynamic modeling, taking advantage of the known X-ray crystallographic structure and normal mode flexible fitting (NMFF) analysis, to describe the changes that occur in GroEL structure induced by GS_m binding. The NMFF analysis reveals that the molecular movements induced by GS_m binding propagate throughout the GroEL structure. The modeled molecular motions show that some domains undergo en bloc movements, while others show more complex independent internal movements. Interestingly, the substrate-bound apical domains of both the *cis* (GS_m-bound ring) and *trans* (the opposite substrate-free ring) show counterclockwise rotations, in the same direction (though not as dramatic) as those documented for the ATP-GroEL-induced structure changes. The structural changes from the allosteric substrate protein-induced negative cooperativity between the GroEL rings involves upward concerted movements of both *cis* and *trans* equatorial domains toward the GS_m-bound ring, while the inter-ring distances between the heptamer contact residues are maintained. Furthermore, the NMFF analysis identifies the secondary structural elements that are involved in the observed ~5 Å reduction in the diameter of the cavity opening in the unbound *trans* ring. Understanding the molecular basis of these substrate protein-induced structural changes across the heptamer rings provides insight into the origins of the allosteric negative cooperative effects that are transmitted over long distances (~140 Å).

**Corresponding authors*

Keywords: GroEL; chaperonin; electron microscopy; protein folding

Introduction

The GroEL chaperonin is an absolutely essential allosteric protein machine that binds partially folded protein intermediates, prevents protein misfolding and assists in the folding of proteins in an ATP-dependent manner. The GroEL oligomer contains 14 homologous subunits (57 kDa each) arranged in two heptameric rings, stacked in a back-to-back fashion. Each subunit is divided into three functional domains. The apical domain is responsible for binding protein substrates and the cochaperonin GroES, the equatorial domain binds and hydrolyzes ATP, and the intervening intermediate domain transmits allosteric signals between the apical and equatorial domains. The allosteric regulation of the GroEL oligomer by ATP is evident as nested cooperativity, which involves positive intra-ring cooperativity and negative inter-ring cooperativity.[1] Although most structural studies of GroEL to date have focused on the ATP/ADP-dependent transitions, it is now clear that the binding reaction of the protein substrate

Present address: S. Falke, Department of Biology, William Jewell College, 500 College Hill, Liberty, MO 64068, USA.

Abbreviations used: EM, electron microscopy; GSm, a single large monomer of glutamine synthetase; NMFF, normal mode flexible fitting.

E-mail addresses of the corresponding authors: egogol@umkc.edu; mfisher1@kumc.edu

with GroEL influences the allostery, dynamics and structure of the chaperonin machinery.

At an intra-ring level of allostery, the GroEL heptamer binds and hydrolyzes seven ATP molecules in a positive cooperative manner.[2,3] At this structural level, protein substrate binding influences the intra-ring positive cooperativity by opposing the intra-ring ATP binding.[4] Within the heptamer, the antagonistic protein substrate and ATP binding is thought to control concerted shifts between two opposing global conformations.[1] For example, protein substrate binding favors the conformer (the taut or T state) that has a strong affinity toward binding protein substrate and a weakened affinity for ATP. The ATP binding reaction opposes this conformational shift and favors the conformer (referred to as the relaxed or R state) that has a stronger affinity for nucleotide and a weakened affinity for the protein substrate. The protein substrate affinity is weakened further when the cochaperonin GroES binds preferentially to the nucleotide-bound GroEL, triggering massive conformational changes, to create a nanostructured chamber inside the GroEL–GroES complex that momentarily encapsulates small to medium-sized folding protein substrates (~20–50 kDa), allowing them to fold sequestered from bulk solvent.[5]

The influence of the protein substrate on the GroEL inter-ring allostery is most important during the functional chaperonin cycle. The particular allostery manifests negative cooperative effects on inter-ring ATP, polypeptide and GroES binding. That is, once ATP, polypeptide or GroES is bound on one ring, the affinities for further ATP or polypeptide or GroES binding on the opposing ring decrease. Functionally, the long-range inter-ring transmission of binding energy of the polypeptide (over ~140 Å) has dramatic effects on the chaperonin cycle. Horwich and co-workers observed that polypeptide binding on one ring facilitates the dissociation of bound GroES and entrapped protein on the opposing ring.[6] Furthermore, as dictated by the tenets of thermodynamic reciprocity, GroES binding to one heptamer facilitates the dissociation of the protein substrate from the opposite ring.[7–10]

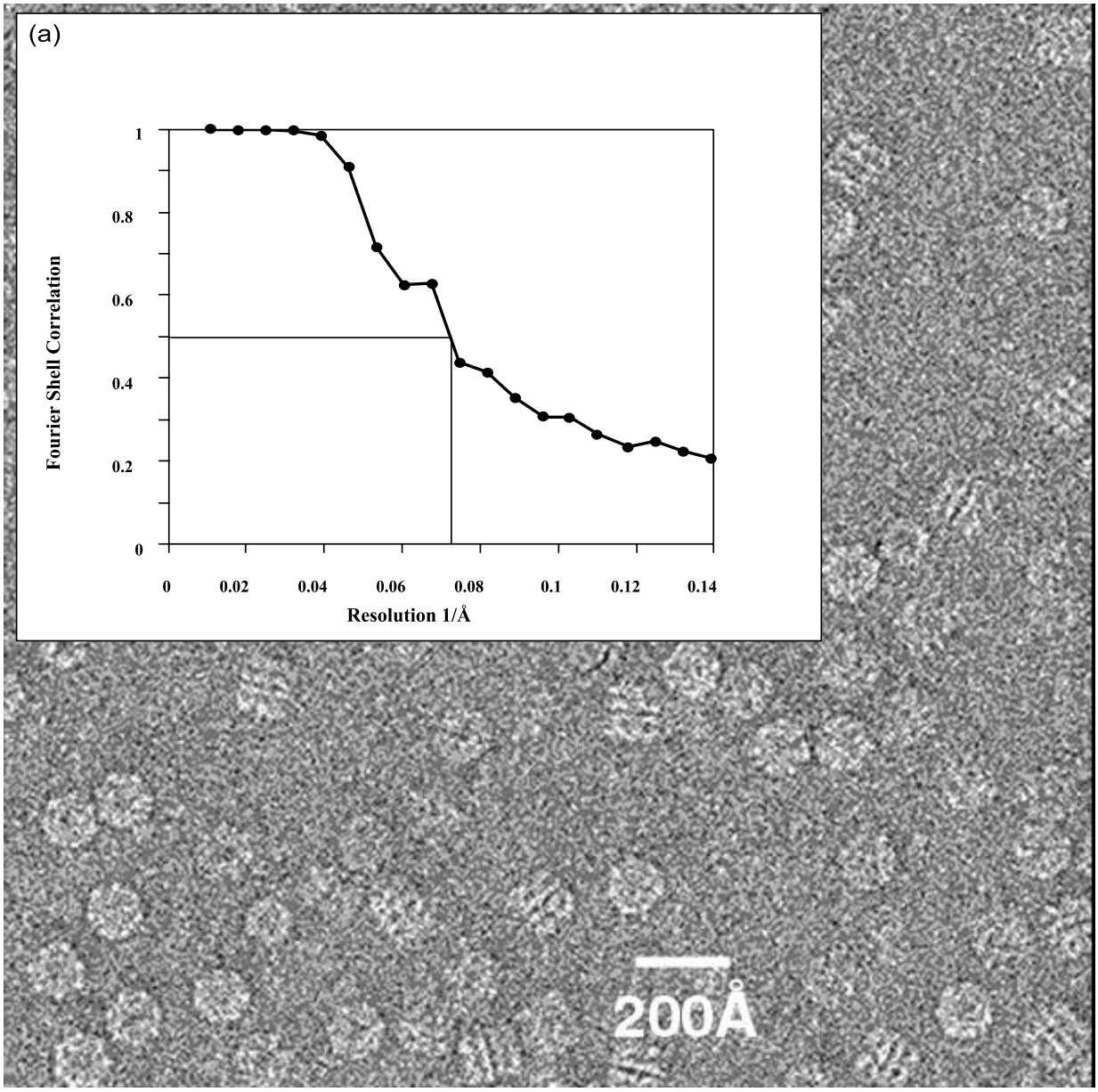
It is evident that both intra-ring and inter-ring allostery are influenced dramatically by protein substrate binding. Unfortunately, very little structural, dynamic or energetic information is available to describe the functional interactions of protein substrates with the chaperonin. To identify possible structural changes that govern long-range substrate protein-induced allosteric communication, we analyzed cryo-electron microscopy (cryo-EM) images to determine the 13 Å structure of the asymmetric GroEL tetradecamer bound to a single large monomer of glutamine synthetase (GS_m) (51 kDa) (Figure 1). Using normal mode flexible fitting analysis,[11] the X-ray crystal coordinates were iteratively modified to fit into the cryo-EM shell to identify the protein structural changes that define the allosteric negative cooperative transition. The observed conformational changes in GroEL induced by binding a large protein substrate (GS_m) are significantly different from those reported recently to be induced in GroEL when binding a 12 amino acid residue peptide.[12] These observations suggest that the allosteric conformational changes in GroEL induced by protein binding are variable, and may depend on properties of the substrate protein.

Results

Our previous analysis of GroEL–GS_m complexes revealed structural changes in the chaperonin caused by the binding of a specific protein substrate in a non-native state.[13] In the work described here, we have improved the resolution and accuracy of our initial study, and provide a molecular approach to interpreting the observed structural effects. The gain in resolution has been accomplished by collecting images of superior optical quality (using a higher-resolution electron microscope), increasing the number of images in our analysis, and changing aspects of our classification and refinement procedures. To reduce the number of optical artifacts, we have integrated images taken at a range of defocus values, in contrast to our previous work at more limited resolution. As in our previous analysis, we have dealt with the heterogeneity of the images that is due to incomplete binding of GS_m by GroEL. At higher resolution, we have identified another possible source of heterogeneity, that in the conformation of the GroEL–GS_m complex. We applied a stringent correlation criterion to maximize the self-consistency of the images included in our analysis, improving both resolution and accuracy of the resulting reconstruction. We eliminated any potential bias toward the X-ray crystal structure by developing models from a classification of the images themselves. Finally, we viewed the structure of the GroEL–GS_m complex at an atomic level using normal mode flexible fitting (NMFF)[11] of the crystallographic atomic coordinates of the unliganded GroEL into the reconstructed complex. This approach provides an energetically sound method of altering protein structures to model conformational changes induced by experimental conditions.

Classification of the heterogeneous image set

As documented previously, the images used to determine the GroEL–GS_m structure at 13 Å were heterogeneous due to incomplete loading of GroEL with GS_m.[13,14] Thus, these data required classification to separate unliganded GroEL from those particles with GS_m bound. We began this process with 2D correspondence analysis procedures,[14] which we performed on the entire set of 7200 side-view images with the exception that multiple eigenvectors, as opposed to a single one, were used for hierarchical classification. Our separation of



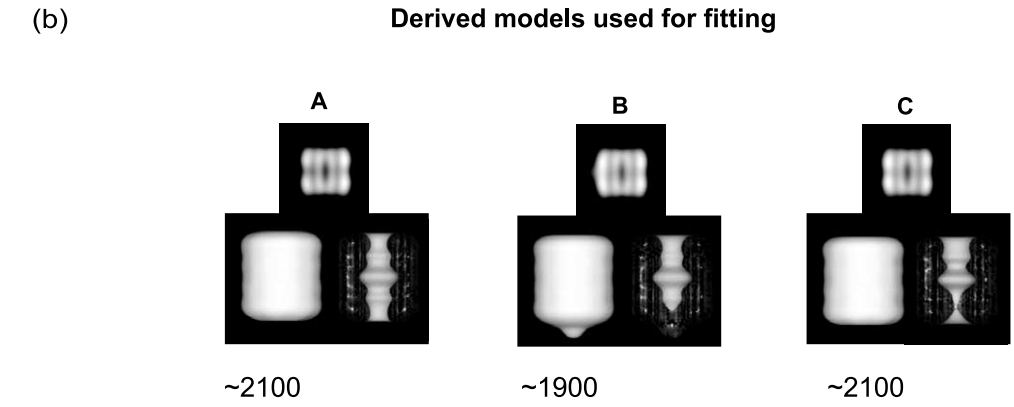


Figure 1. (a) Left: cryo-EM image of a GroEL–GS_m specimen taken with a field emission gun (FEG) EM showing numerous particles embedded in the ice. The percentage of side-views in this field is around 36%. Inset: Fourier shell correlation curve. (b) The numbers of particles and the averaged structures that were used for the three initial models were derived from classification analysis and cylindrically averaged.

side-view images into substrate-bound and unbound categories allowed us to build starting models, which were then used in competing model classification and 3D reconstruction. In this initial step, three major classes emerged from classification of the images (discarding 15% of them), two of which appeared to contain density in addition to that of GroEL. Low-resolution 3D models (Models A, B and C in Figure 1(b)) were calculated by sevenfold back projection of the averaged images of each class. Two of these models (B and C) displayed the additional density within the binding cavity of one ring, while the remaining model (A) has no density other than that corresponding to GroEL.

All 7200 images were then subjected to a multiple-model competing classification in 3D, using all three starting models. Rather than relying on crystal structure reference models throughout the classification and reconstruction scheme as outlined previously,[14] the use of the models built from correspondence analysis classification reduced initial model bias. By a fifth round of refinement, image migration among the models was minimized (<5% image migration between models) and the Fourier shell correlation for each of the reconstructions reached a plateau. Two of the three resulting structures, models B (2100 images) and C (1500 images), had distinct protein density within the binding cavity of one ring and yielded resolutions of ~18 Å (as determined by the Fourier shell correlation curve at a value of 0.5). The third model (~3500 images) had no protein density in either of the two cavities, refined to ~13 Å resolution, and represented the unliganded GroEL fraction of the image data set.

To optimize the resolution of the GroEL–GS_m reconstruction, the two substrate-bound classes (B and C) were combined to make a single class of 3600 images and refined using model B. Knowing that variability existed in this class of images, we selected the images that were optimally consistent with a single 3D model. For each round of alignment and reconstruction, only images whose correlation coefficients against that round's model projection exceeded the mean were used in reconstruction. However, all images were reintroduced into the alignment and classification procedure for each subsequent round. Although this classification refinement and stringent correlation cut-off procedure eventually discarded about half of the initial 3600 images, the remaining GroEL–GS_m images (1800 self-consistent images) refined to 13 Å resolution. When an identical procedure was applied to the unliganded GroEL image class (model A), half the images were retained and refined to 11 Å. We could discard 15% of the images (2600 images) and still obtain a structure that refined to ~11 Å. To test for potential model bias in the final structures, models were exchanged for the two final image classes (i.e. using the unliganded structure as a starting model for the GroEL–GS_m images and the liganded structure for unliganded ones). This test resulted in two structures that were identical with our final classified structures; the GroEL–GS_m class gained density even though none corresponding to substrate were initially included in the model; the unliganded class lost substrate density despite its presence in the starting model. This outcome strongly supports our contention that our classification procedure successfully separated GroEL–GS_m from unliganded GroEL populations.

The correlation based selection procedure used to generate a self-consistent GroEL–GS_m reconstruction left an equal number of "rejected" images. We re-examined these by subjecting them in a parallel reconstruction and refinement (starting with the unliganded structure). This procedure yielded a GroEL structure with a smaller additional substrate density (not shown), which may represent one or more alternative conformational states of GroEL–GS_m. The refinement of the rejected images resulted in a structure that resolved to ~18 Å. If this particular GroEL class represents a heterogeneous population, future classification and refinement of this separate image class will require us to increase our initial image population to better define the structural nature of the observed heterogeneity.

New structural features of the intermediate-resolution GroEL–GS_m

A comparison of the 13 Å GroEL–GS_m structure with our previous one[14] shows that the major structural effects of binding GS_m are verified, specifically the protrusion of substrate from the *cis* end of GroEL, partial constriction of the *trans* ring and a slight bulge of the equatorial domain. However, the magnitude of these differences is better defined and is actually smaller in the more accurate, higher-resolution reconstruction. The outer surface of GroEL–GS_m differs from that of unliganded GroEL by 3–7 Å at these points of maximum deviation (see the overlay in Figure 2(a) and (b)). Each of the major difference features cited above is altered in comparison with that seen in our prior reports of GroEL–GS_m at lower resolution. The additional density attributed to the protein substrate is reduced in size and location to a cylinder at the opening of the *cis* apical cavity, presumably due to the effects of sevenfold averaging at higher resolution. The *trans* apical opening, while displaying a constriction compared to the unliganded GroEL structure, is less restricted than in the lower-resolution reconstruction, and the central bulge of GroEL–GS_m appears to be limited mainly to the *trans* equatorial domain. These differences between our previous and current GroEL–GS_m reconstructions are the result of improvements in both the quality of our data and the methods used for its analysis, most prominently the more stringent classification procedures that allowed the selection of the maximally consistent molecules of the major conformation of the complex present in our preparations.

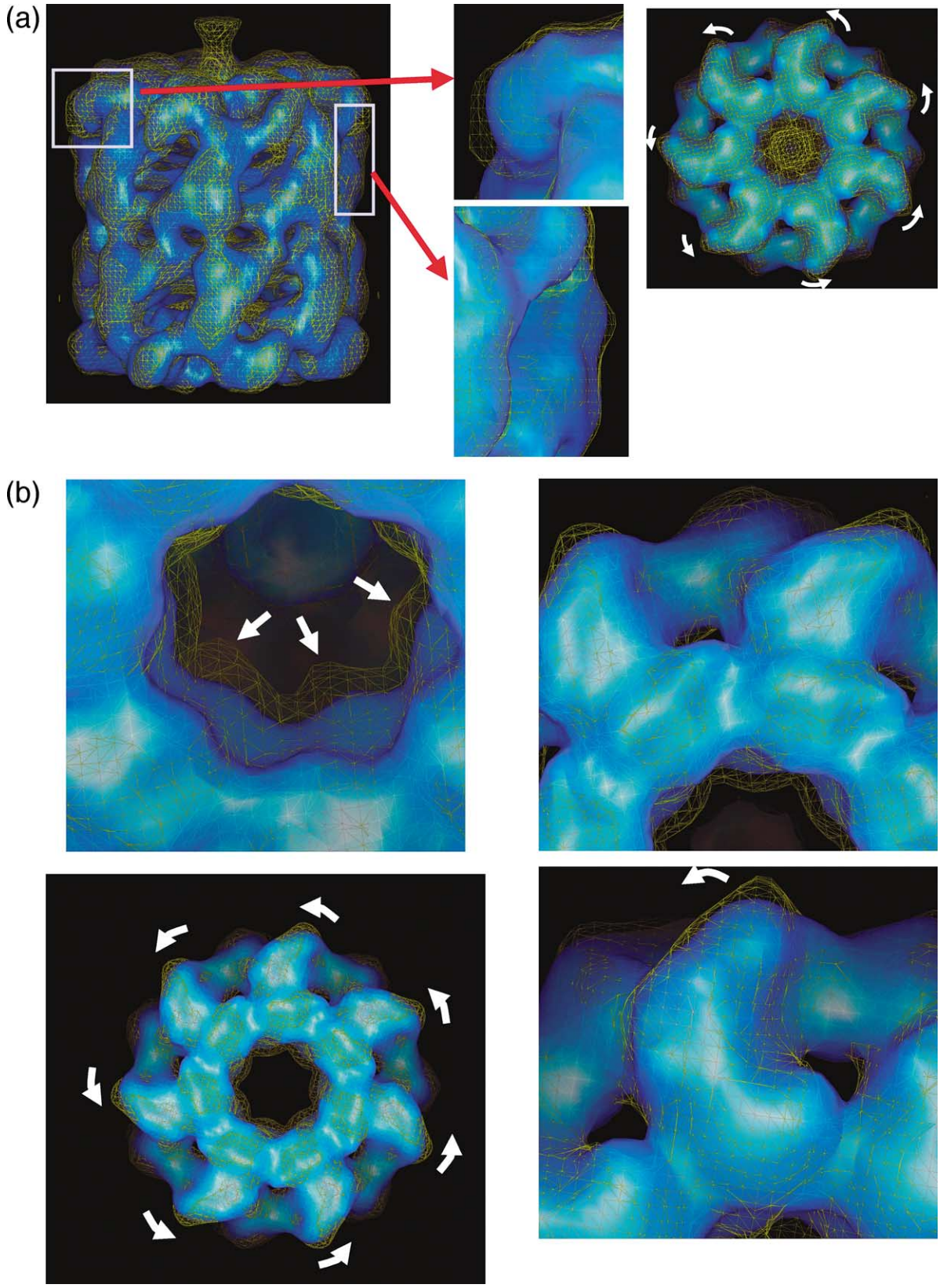


Figure 2. (a) Overlaid side and (b) *trans* side surface view structures of the GS_m-bound GroEL (yellow mesh) and substrate-free GroEL (blue solid surface) shows the differences and rotations between the different protein species surfaces in more detail. Both molecular surfaces were generated using the molecular program Chimera (UCSF).

Normal mode analysis of GroEL and GroEL–GS_m

The combined cryo-EM structure and the 2.8 Å X-ray crystallographic coordinates[15] were used to perform dynamic fitting using normal mode analysis.[16] A preliminary rigid fit of the GroEL structure (Figure 2), PDB code 1OEL, into the cryo-EM maps of GroEL and the GroEL–GS_m complex revealed that conformational changes were necessary to accommodate the X-ray structure into the EM map of the complex. The initial correlation coefficient between the computed and measured electron densities for both GroEL and GroEL–GS_m complexes was 0.79. Flexible fitting of the high-resolution structure into the cryo-EM map of the unliganded GroEL and GroEL–substrate protein complex was performed using the NMFF procedure (Figure 3).[11] The NMFF procedure uses a linear combination of low-frequency normal modes in an iterative manner to deform the structure optimally to conform to the low-resolution structure. Normal modes are used for the flexible fitting because they represent the large conformational changes observed in biological systems.[16,17] One advantage of NMFF analysis over the independent fitting of disconnected domains[18,19] is that it incorporates the structural constraints of the connected hinge regions, thereby restraining domain separation to energetically reasonable limits.

The NMFF algorithm was adapted to take into account the 7-fold symmetry of the GroEL structure. Thus, during the refinement procedure, only the lowest frequency normal modes that display the 7-fold symmetry are selected as search directions. The correlation coefficients after interactive fits into the electron density maps are 0.90 and 0.88 for GroEL and GroEL–GS_m, respectively. The RMS difference between the entire unliganded and GroEL–GS_m structures is 2 Å.

Structure of unliganded GroEL

The unliganded GroEL heptameric rings were asymmetric (Figure 2(a), top row), in agreement with previous intermediate-resolution (10–15 Å) cryo-EM solution structures of unliganded GroEL.[19,20] The normal mode fit of two GroEL heptamers also showed asymmetric structure between the opposing rings. The RMS difference of the C^{α} backbones between the two heptamers was calculated to be 1.2 Å.

The overall surface structure of the GroEL–GS_m complex

The 13 Å GroEL–GS_m map (Figure 3) shows a distinct additional density within the cavity of one ring. This density, attributed to the GS_m substrate protein, is next to the apical domain cavity, near the helices H and I (identified in Figure 4 by yellow arrows), and extends far enough into the cavity to potentially interact with the loop separating GroEL β strands 6 and 7.[15,21] This positioning of GS_m within the binding cavity is fully consistent with mutational analysis, indicating that the majority of

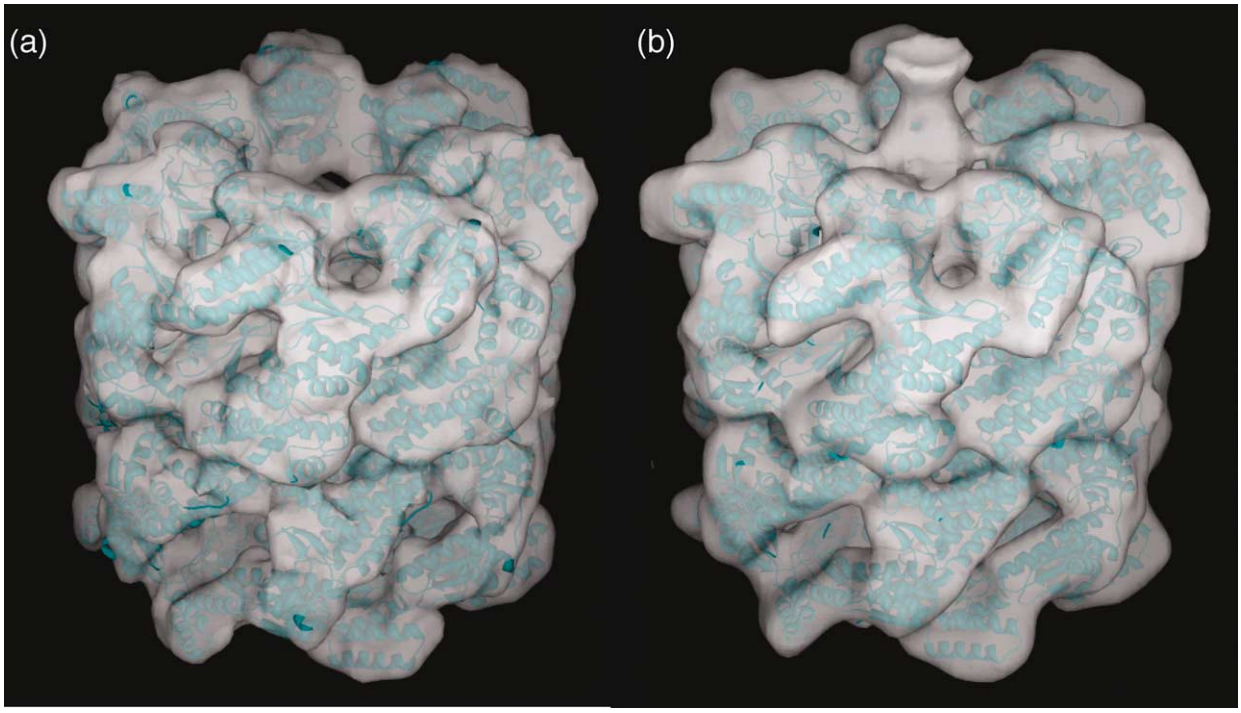


Figure 3. Normal mode flexible fitting of the X-ray coordinates of GroEL 1-OEL PDB inside the defined EM protein density shell for (a) GroEL (A) and (b) GroEL–GS_m. The correlation coefficient for both fits was high (0.90 for substrate-free GroEL and 0.88 for GroEL–GS_m. The ribbon (cyan) and solid surface (white) representations were generated using the molecular program Chimera (UCSF).

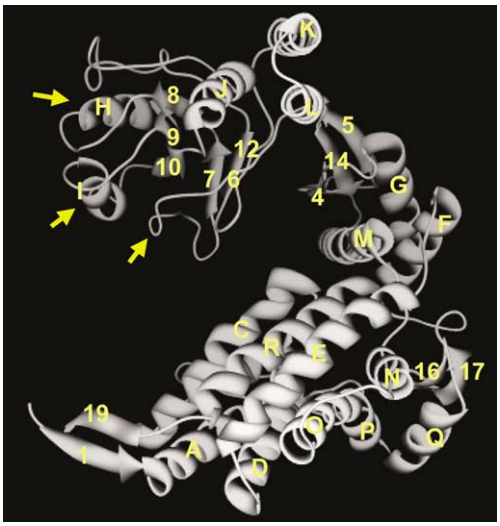


Figure 4. Ribbon model of a single GroEL subunit with labeled strands and helices for identification (adapted from Xu *et al.*).[21] The yellow arrows depict the general secondary structural regions of GroEL that are implicated in binding a hydrophobic substrate protein.[22]

substrate–protein binding sites lie on two parallel helices (H and I) in the apical domains.[22] The orientation and structure of the bound substrate cannot be interpreted directly for two reasons. First, application of 7-fold symmetry in the reconstruction smears the density of the asymmetric GS_m. Secondly, the orientation and position of GS_m may be variable, resulting in an "average" position for GS_m. When the GroEL–GS_m structure is thresholded to include more than 100% of the molecular volume, the most visible increase in protein density of the structure occurs around the bound GS_m suggesting that some GS_m volume is lost during the reconstruction.

The higher-resolution GroEL–GS_m map contains a novel feature in a centrally located strand of density connecting the equatorial domains of opposing rings (not shown). This feature is unlikely to be attributable to the bound GS_m. Interestingly, the calculated normal mode analysis fitting of the GroEL subunits into the reconstruction positions shows that the unresolved equatorial residues of the N terminus (five amino acid residues) and C terminus (24 amino acid residues) end up at the beginning of the new inter-ring density connections, making it possible that the central density connecting the two rings of the GroEL–GS_m map is composed of these residues.

Comparing substrate-free and GS_m-bound cyro-EM GroEL structures by NMFF

Specific secondary structure element changes that are modeled by NMFF are identified by labels and

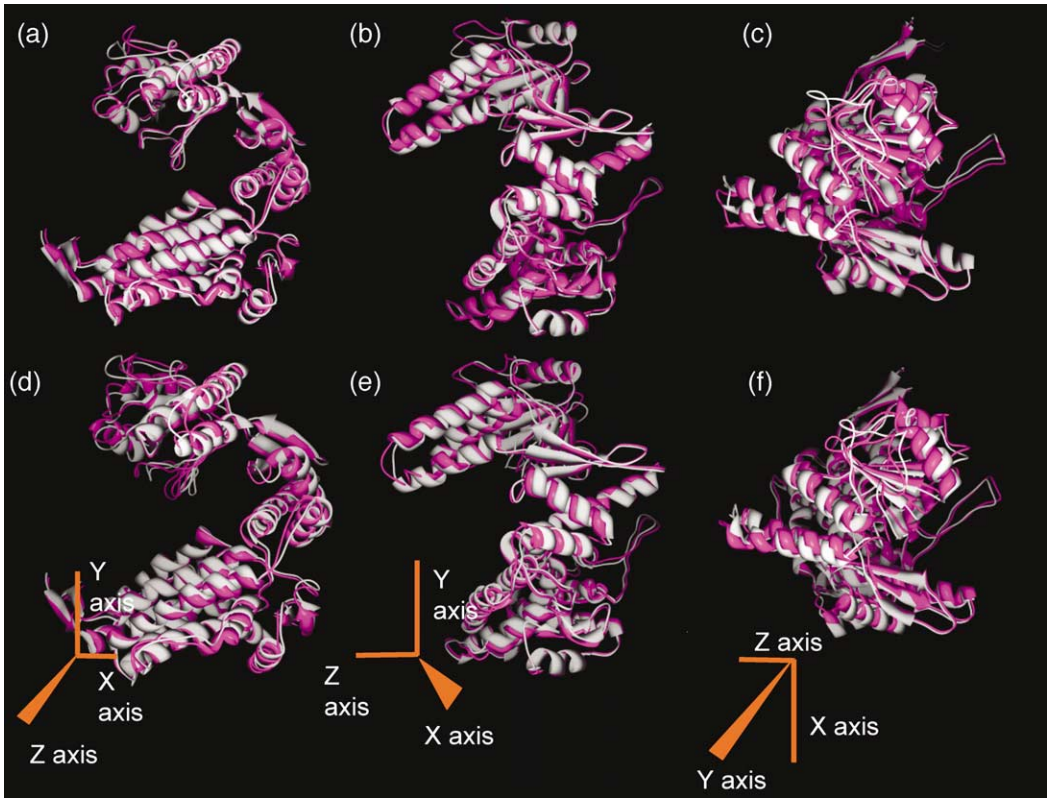


Figure 5. Comparison of the structural rearrangements for GroEL–GS_m *versus* unliganded GroEL obtained from normal mode analysis (NMA) fits. Overlays of monomeric subunits comparing the substrate-free (white) and GS_m-bound (purple) forms of GroEL shows major movements about assigned *Z*, *X* and *Y* (7-fold) axes with either (a)–(c) the *cis* or (d)–(f) the *trans* subunits.

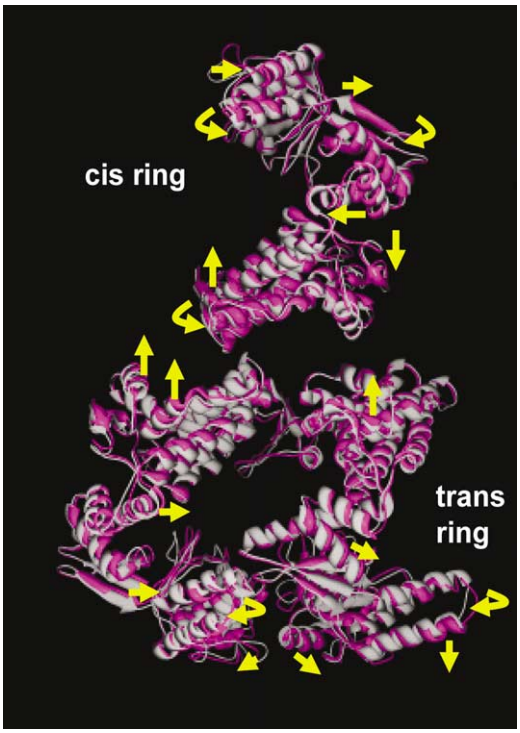


Figure 6. Compiled movements with directional arrows showing the position changes of three side-views of representative subunits (one *cis* and two *trans* subunits). The *cis* subunit is located within the GS_m-binding ring. The movements (yellow arrows) show concerted upward movements and rocking motions in the *cis* and *trans* equatorial domains. The movements are transmitted throughout the structure, with the largest differences eventually being manifested within the apical domains of the *trans* ring. The rotations of both *cis* and *trans* apical domains are counterclockwise.

are displayed in Figure 4.[21] Most of the GS_m-induced movements discussed in the following sections are presented as overlaid ribbon structures (Figures 5 and 6). Since normal mode analysis uses vibrational modes centered on the C^α backbones, the resulting structures are represented as ribbon diagrams (Figures 4–6). In Figures 5 and 6, the white ribbon structure represents the substrate-free or unliganded GroEL, while the overlaid GS_m-induced changes are indicated by the presence of the purple ribbon.

GS_m-induced changes in the *cis* and *trans* apical domains

When GS_m binds to GroEL, the EM surface maps suggest that the apical domains in both the *cis* (GS_m-bound) and *trans* rings move counterclockwise about the 7-fold axis (Figure 2(a) and (b)). This rotation is defined by normal mode analysis fits for the unliganded and GroEL–GS_m structure (Figures 5 and 6). Interestingly, in both surface views and NMFF ribbon diagrams, the GS_m-induced changes in the apical domains of the *trans* ring are larger than the movements observed at the *cis* apical domains, even though the *cis* apical domain contacts GS_m directly. Since there is no clearly defined boundary between the interior binding cavity of the *cis* subunits of GroEL and the GS_m protein density, the separate movements of the C^α backbone in this region are less well defined. However, the NMFF-derived structures that fit the exterior EM-defined shell furthest from the central cavity show that the prominent super secondary structural motif defined by the K and L helices in both the *trans* and *cis* apical domains rotates counterclockwise with respect to the 7-fold axis. The large mobile loop regions (297–322) in both the *cis* and *trans* apical domains appear to rotate in the same direction. A comparison of the unliganded GroEL with the GroEL–GS_m structures reveals that some of these mobile loop regions appear to move as much as 6–7 Å when GS_m binds to GroEL (e.g. C^α carbon distances of Ile305). The movements modeled by the best NMFF analysis at the exterior of the *cis* apical domain show that both of the H and I helices of the *cis* apical domain move toward the interior of the cavity (see Figure 5; and see the Supplementary Data).

The NMFF analysis identifies the movements of the H and I helices that define the internal cavity surface at the *trans* ring. Specifically, the fits for the *trans* apical domains show that the H helix must move back and slightly upward, while the I helix rotates both upward and outward with its C-terminal end extending into the cavity (Figure 5(b)). Furthermore, movements of the mobile loop located below the I helix centered at Thr210 moves a total of 7 Å in the same outward direction as the I helix. Thus, the inward movement of the I helix and the lower mobile loop toward the substrate-binding cavity represent the protein secondary structural elements that cause structural constriction of the *trans* cavity (Figures 2 and 5(c) and (f)). As stated above, the apical L and K helices undergo significant movements in both the *cis* and the *trans* rings. However, these helix-turn-helix motifs rotate counterclockwise within the *trans* ring and pivot slightly upward about the apical-intermediate hinge region. The modeled apical *trans* domain movements are not en bloc, because the H helix moves upward with respect to the 2-fold axis.

Intermediate domain movements induced by GS_m binding

As viewed from the side of a representative GroEL monomer (Figure 5(a) and (d)), the normal mode analysis fits show that when GS_m binds GroEL, the intermediate domain, particularly in the β strands at the apical G192/G375 hinge region, move in opposing directions in the *cis* and *trans* rings. For instance, the hinge region of the *cis* apical and intermediate interfaces moves outward, pushing β strands 5 and 14 away from the cavity. In addition, the *cis* intermediate domain moves clockwise (with respect to the 7-fold axis) in an en bloc

fashion, with a slight downward movement toward the equatorial domain (Figures 5(b) and 6). The prominent helices of the *trans* intermediate domain also move slightly toward the equatorial domain (Figure 5(d)) but in contrast to the observed movements in the *cis* intermediate domain, the two β strands next to the *trans* ring apical/intermediate G192/G375 hinge region move toward the internal binding cavity. These coupled movements presumably drive the inward movement of *trans* apical domain toward the interior of the cavity, resulting in the observed EM surface ring constriction. The NMFF analysis also indicates that the prominent intermediate helices (M, F and G) in the *trans* ring do not move en bloc like the *cis* domain. Also unlike the GS_m binding-induced en bloc movements of the *cis* intermediate domain, the movement of the *trans* intermediate domains, as viewed from the outside surfaces of the G and F helices (view down the *X* axis; see Figure 5(f)), shows that these two helices move away from the cavity and in the direction opposite that of the M helix. The *trans* intermediate domain appears to pivot near the Pro137 hinge region, bringing the M helix closer to the equatorial domains (Figure 5(d)).[21]

GS_m-induced equatorial domain movements

Due to constraints imposed by the quaternary contacts, the equatorial domains of the *cis* and *trans* rings do not undergo very large rotational movements compared with the apical or intermediate domains. However, the movements of these equatorial domains that are modeled by NMFF are critical because they are the structural manifestation of the transmitted binding energies across the GroEL heptamer interfaces. The predominant structural changes of both *cis* and *trans* equatorial domains show coupled concerted motions where the axial outward movement (away from the 2-fold axis) of the *cis* equatorial domains is accompanied by a compensatory upward movement (towards the 2-fold axis) of the *trans* equatorial domains (Figures 6 and 7). Only a small portion of the *cis* equatorial domain, the C terminal portions of the Q and N helices rock slightly downward toward the opposite ring. The specific C^{α} distances between the interfaces in both the unliganded GroEL and GroEL–GS_m remain the same, suggesting that the movement of the *trans* equatorial domain toward the 2-fold heptamer interface maintains the contact distances and hence non-covalent bonding between the interacting residues (illustrated by Figure 7). In addition to the concerted interface movements between the equatorial domains, there are subtle rotations changes that occur when substrate protein binds. Like the en bloc movements of the *cis* intermediate domain, the equatorial domain of the *cis* ring moves in an en bloc fashion in tandem with the apical domain movements (slight counterclockwise rotation about the 7-fold axis) (for a view down the 7-fold axis, see Figure 5 (b)). In contrast, the equatorial domain in the *trans* ring does not move en bloc (compare the views down the *X* axis overlays in the equatorial regions of Figure 5(b) (*cis*) with (e) (*trans*)). It is apparent that the long-range communication between rings is governed primarily by the upward shift and slight rotation of the *cis* equatorial domain and the maintenance of the integrity of the heptamer interface distance between both equatorial domains.

Imposing 7-fold symmetry upon the GroEL–GS_m complex

Our previous low-resolution GroEL–GS_m structure indicated that GS_m becomes partially buried within the binding cavity.[13] In its folded compact state, GS_m is roughly spherical, with a mean diameter of ~60 Å. Given its size, a partially folded GS_m may not be able to enter the GroEL binding cavity (~55 Å) without inducing some movement in and interacting with the multiple apical domains. Since the interactions between GroEL and GS_m are distorted due to the 7-fold symmetry operation, it is still possible that GS_m interacts with only a subset of the monomeric binding sites in the heptameric ring. Although intra-subunit cross-linking studies show that the internal domain conformational changes are concerted during nucleotide binding and

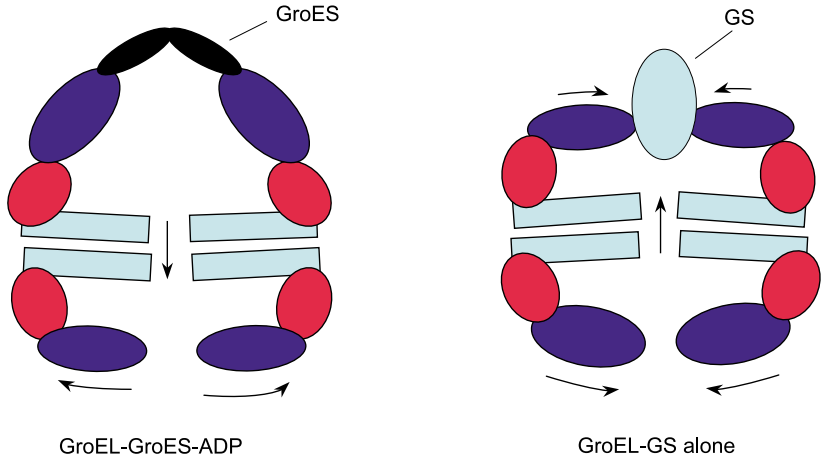


Figure 7. The block diagrams of a side-view of the GroEL–GroES–ADP complex (modified from Xu *et al.*)[21] and the GroEL–GS_m complex shows that some of the prominent movements induced by binding a large protein substrate such as GS_m are in the direction opposite that of those depicted in the GroEL–GroES–ADP complex.

hydrolysis,[1] it is not known if the substrate protein-induced changes in GroEL are concerted or sequential. If sequential rather than concerted changes occur, the application of 7-fold symmetry during reconstruction would yield an inaccurate depiction of the substrate protein-induced structural changes in GroEL. To determine how a partial or asymmetric sequential conformational change in GroEL structure is altered if 7-fold symmetrized, we modeled a modest asymmetric conformational change: 15° upward rotation of three adjacent apical domains in the GroEL heptamer. When this test asymmetric structure was 7-fold symmetrized, we found that there was no appreciable loss of GroEL protein density. However, the final positions of all seven apical domains within the modeled ring were still large, tilting upward by about 7°. Because the GroEL–GS_m complex was symmetrized during reconstruction (to maximize resolution), we cannot be certain that the GS_m-induced conformational changes within the heptamer result are concerted allosteric structural transitions. Reconstructions carried out without using 7-fold symmetries resulted in very noisy reconstructions due to limited numbers of images. To test if the substrate protein-induced conformation change in GroEL is concerted, and to avoid imposing 7-fold symmetry, we are presently collecting and classifying more images.

Discussion

The normal mode analysis fitting indicates that GS_m binding induces structural rearrangements throughout the entire GroEL oligomer. This reciprocal long-range structural effect is observed upon GroES binding,[21] resulting in large changes in the opposite unoccupied ring. Our GS_m–GroEL structures also show substantial structural changes with the unoccupied *trans* heptameric ring.

Traditionally, the structural descriptions of the transitions between R and T states within the heptamer rings focuses primarily on the movements of the apical domain as it rotates about the 7-fold symmetry axis. A recent structure of a GroEL–oligopeptide complex shows that a small, high-affinity oligopeptide (SBP) induces small rotations of the apical domain that are clockwise about the 7 fold symmetry axis.[12] These movements agree with the proposed substrate protein-induced structural changes that favor a more taut (T) state.[4] However, the data presented here shows that the *trans* and *cis* apical domains of the GroEL–GS_m viewed from their separate ends, rotate counterclockwise, which is the direction that is observed in the proposed ATP-bound R or relaxed state.[19] The apical domain regions of the GroEL–GS_m complex are in an intermediate conformation between previously proposed T and R states. Since the GroEL apical domain has a high degree of flexibility, we suggest that it is more relevant to focus on the movements between the rings (within the equatorial interface regions) to define the structural elements that control the functionally relevant interring negative allosteric effects. The equatorial domains show a rocking and concerted upward motions toward the substrate-bound ring when GS_m substrate binds to GroEL (Figures 5–7). Interestingly, the observed movements of the apical and intermediate domains for GroEL–SBP[12] and GroEL–GS_m rotate in opposite directions, suggesting that the substrate protein-induced changes are probably dependent on undefined substrate protein properties.

The most important observation of this work is that the GS_m-induced conformational changes in the *trans* ring are significantly larger than those observed within the *cis* ring. Although it is feasible that some of these same movements can contribute to the functionally relevant global negative cooperative effects that govern accelerated GroES dissociation and ejection of the enclosed substrate protein (Figure 7), a better understanding of the structural origin of the substrate-induced allosteric negative cooperativity in the chaperonin cycle will be realized when we compare the ADP–GroEL–GroES complex to the substrate–protein-bound ADP–GroEL–GroES forms.

Experimentally, it is observed frequently that substrate proteins bind to GroEL with a 1:1 stoichiometry,[13] even though there are two possible binding sites on each end of the GroEL tetradecamer. In our reconstruction, the large improvement in resolution of the GroEL–GS_m complex and the NMFF analysis have allowed us to obtain a more detailed view of the structural/conformational changes in GroEL induced by binding a large substrate protein, such as partially folded GS_m. The counterclockwise rotations of the *cis* (GS_m-bound) and *trans* apical domains, movements f the apical substrate-binding helices and loop regions, the corresponding constriction of the *trans* apical opening and the slight upward movement of the *trans* H and I helices describe adequately the observed surface changes in the *trans* ring. However, the best definitive measure to quantify this anecdotal antagonistic effect requires a determination of the binding energetics of the substrate protein onto GroEL.

Previous hybrid crystallography work with cryo-EM single-particle reconstructions of the GroEL and GroEL–ATP or GroEL–ADP–GroES states have employed procedures to fit optimally clipped X-ray crystallographic domains of substrate-free GroEL as en bloc movements inside the EM shell.[19] This approach provides a reasonable approximation of the global movements for the large changes in GroEL that occur during binding of GroES with ATP or ADP. However, our normal mode analysis indicates that global movements may actually be significantly flexible, thus making it difficult to describe the structures of various allosteric states of GroEL by rigid-body changes. Indeed, our NMFF analysis of the smaller changes in GroEL induced by the binding of substrate

protein shows that all of the domains may not move in a strict en bloc fashion. Furthermore, it is apparent that the flexibility that we see in all three of the major subunit domains will complicate efforts to equate particular structural states within the paradigm of a two-state allosteric structural transition.

The thermodynamic and kinetic ramifications of the polypeptide-induced structural changes are unknown. Future work will include defining the coupling free energies and dynamics of nucleotide binding to the substrate bound and unliganded *trans* rings along with a more detailed structural analysis. The protein-induced changes in GroEL appear to vary depending on the substrate, and may provide some functional basis for the recently observed weak substrate specificity among various chaperonins.

Methods

Electron microscopy and image processing

GroEL–GS_m complexes were prepared for cryo-EM as described.[13] Images of frozen-hydrated GroEL–GS_m complexes were recorded on a Philips CM200 field emission gun (FEG) electron microscope operated at 120 kV at the Scripps Research Institute. Images were collected using minimal-dose methods at a magnification of 66,000× with a defocus of 0.5–2.0 μm. A total of 80 micrographs were chosen as visually acceptable (on the basis of particle clarity and lack of visible charging and/or specimen drift effects) and digitized using a Eurocore Hi Scan densitometer with a step size of 17.4 μm, corresponding to 2.7 Å pixels. Power spectra were used as a more rigorous test of micrograph quality. Digitized micrographs with indications of drift, charging and/or astigmatism (determined by a comparison of the presence of Thon rings in either the power spectrum or the rotational average that are absent from the other) were discarded, leaving ~50 micrographs, from which 7200 side-view orientations were selected. Because orientations representing rotation about the long axis of GroEL cover the full structural sphere, side-view images alone (particles exhibiting the characteristic four stripes of density) can be used for image analysis and reconstruction.[19]

Classification, reconstruction and refinement were performed as described,[14] but with the changes described here. Cross Transfer Function (CTF) minima were determined by comparing the rotationally averaged power spectrum of each micrograph to a series of calculated CTFs and the image set was corrected for phase inversion prior to classification and reconstruction. All image analysis was performed with SPIDER.[23] Following initial three-dimensional reconstruction, one-dimensional structure factors were calculated from the GroEL crystal structure amplitudes,[15] and were used to correct for amplitude-dampening effects of the CTF. Because unliganded GroEL and the GroEL–GS_m complex differ in protein density by only ~6%, the crystal structure amplitudes present a reasonable approximation for correcting both the unliganded GroEL and GroEL–GS_m maps. The Fourier amplitudes of unliganded GroEL and GroEL–GS_m cryo-EM maps were scaled to the amplitudes of the crystal structure and a three variable polynomial enhancement curve was calculated. The filter was applied to both maps followed by band-pass filtration (between 270 Å and 11 Å for GroEL and 270 Å and 13 Å for GroEL–GS_m). The observed structural differences between unliganded GroEL and GroEL–GS_m were apparent prior to amplitude correction. Also, final volumes of both GroEL and the GroEL–GS_m complex were thresholded to 100% of the molecular volume as calculated by the amino acid sequence (with the addition of one GS monomer for GroEL–GS_m). Importantly, all structural features of the GroEL–GS_m complex remain visible even when the threshold level is decreased to 95% or 90% of the molecular volume (only the bound GS monomer suffers sufficient density loss to result in a visible structural difference).

Normal mode flexible fitting (NMFF) analysis

Normal modes were computed using a simplified potential[24] and the RTB method.[25] It has been demonstrated that NMFF can be performed on reduced representations of biological systems such as C^{α}-based models,[11] which allow faster calculation while giving results similar to those achieved with an all-atoms analysis. Thus, in the refinement, only the C^{α} atoms were taken into account. An all-atoms structure for the final C^{α}-based model was reconstructed by energy minimization of the original all-atoms X-ray structure (1OEL) using the C^{α}-based model obtained from NMFF as a restraint and was followed by energy minimization using 7-fold symmetry of the simulation package CHARMM.[26] The molecular graphics images shown as ribbon representations and surface views illustrated in Figures 2–6 were produced using the UCSF Chimera package from the Computer Graphics Laboratory, University of California, San Francisco (supported by NIH P41 RR-01081)†.[27]

Acknowledgements

This work was supported by the KUMC Biomedical Research Grant to S.F and M.T.F., Lied Research Grant RI-G1803820 to M.T.F, NIH GM57403 to E.P.G. and Multi-scale Modeling Tools for Structural Biology (MMTSB) funded by NIH RR12255 to F.T. and C.B. The authors thank Ron Milligan for generously allowing us to collect images on the Philips CM200 (FEG) electron microscope at CIMBIO/TSRI.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2005.02.027

References

1. Horovitz, A., Fridmann, Y., Kafri, G. & Yifrach, O. (2001). Review: allostery in chaperonins. *J. Struct. Biol.* **135**, 104–114.

† http://www.cgl.ucsf.edu/chimera/

2. Todd, M. J., Viitanen, P. V. & Lorimer, G. H. (1994). Dynamics of the chaperonin ATPase cycle: implications for facilitated protein folding. *Science*, **265**, 659–666.
3. Yifrach, O. & Horovitz, A. (1995). Nested cooperativity in the ATPase activity of the oligomeric chaperonin GroEL. *Biochemistry*, **34**, 5303–5308.
4. Yifrach, O. & Horovitz, A. (1996). Allosteric control by ATP of non-folded protein binding to GroEL. *J. Mol. Biol.* **255**, 356–361.
5. Saibil, H. R., Horwich, A. L. & Fenton, W. A. (2001). Allostery and protein substrate conformational change during GroEL/GroES-mediated protein folding. *Advan. Protein Chem.* **19**, 45–72.
6. Rye, H. S., Roseman, A. M., 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.
7. Inbar, E. & Horovitz, A. (1997). GroES promotes the T to R transition of the GroEL ring distal to GroES in the GroEL–GroES complex. *Biochemistry*, **36**, 12276–12281.
8. 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.
9. 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.
10. Chaudhry, C., Farr, G. W., Todd, M. J., Rye, H. S., Brunger, A. T., Adams, P. D. *et al*. (2003). Role of the gamma-phosphate of ATP in triggering protein folding by GroEL–GroES: function, structure and energetics. *EMBO J.* **22**, 4877–4887.
11. Tama, F., Miyashita, O. & Brooks, C. L., 3rd (2004). Flexible multi-scale fitting of atomic structures into low-resolution electron density maps with elastic network normal mode analysis. *J. Mol. Biol.* **337**, 985–999.
12. Wang, J. & Chen, L. (2003). Domain motions in GroEL upon binding of an oligopeptide. *J. Mol. Biol.* **334**, 489–499.
13. Falke, S. F., Fisher, M. T. & Gogol, E. G. (2001). Analysis of GroEL-glutamine synthetase complexes by electron microscopy. *J. Mol. Biol.* **308**, 569–577.
14. Falke, S. F., Fisher, M. T. & Gogol, E. G. (2001). Classification and reconstruction of a heterogeneous set of electron microscopic images: a case study of GroEL-substrate complexes. *J. Struct. Biol.* **133**, 203–213.
15. Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D., Joahimiak, A., Horwich, A. L. & Sigler, P. B. (1994). Crystal structure of GroEL at 2.8 Å. *Nature*, **371**, 578–586.
16. Tama, F., Valle, M., Frank, J. & Brooks, C. L., 3rd (2003). Dynamic reorganization of the functionally active ribosome explored by normal mode analysis and cryo-electron microscopy. *Proc. Natl Acad. Sci. USA*, **100**, 9319–9323.
17. Tama, F. & Sanejouand, Y. H. (2001). Conformational change of proteins arising from normal mode calculations. *Protein Eng.* **14**, 1–6.
18. Roseman, A. M., Ranson, N. A., Gowen, B., Fuller, S. D. & Saibil, H. R. (2001). Structures of unliganded and ATP-bound states of the *Escherichia coli* chaperonin GroEL by cryoelectron microscopy. *J. Struct. Biol.* **135**, 115–125.
19. Ranson, N. A., Farr, G. W., Roseman, A. M., Gowen, B., Fenton, W. A., Horwich, A. L. & Saibil, H. R. (2001). ATP-bound states of GroEL captured by cryo-electron microscopy. *Cell*, **107**, 869–879.
20. Lutke, S. J., Jakana, J., Song, J.-L., Chuang, D. L. & Chiu, W. (2001). A 11.5 Å single particle reconstruction of GroEL using EMAN. *J. Mol. Biol.* **314**, 253–262.
21. 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.
22. Fenton, W. A., Kashi, Y., Furtak, K. & Horwich, A. L. (1994). Functional analysis of the chaperonin GroEL: identification of residues involved in polypeptide binding and release. *Nature*, **371**, 614–619.
23. Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M. & Leith, A. (1996). SPIDER and WEB: processing and visualization of images in 3D electron microscopy and other fields. *J. Struct. Biol.* **116**, 190–199.
24. Tirion, M. M. (1996). Large amplitude elastic motions in proteins from a single parameter, atomic analysis. *Phys. Rev. Letters*, **77**, 1905–1908.
25. Tama, F., Gadea, F. X., Marques, O. & Sanejouand, Y. H. (2000). Building-block approach for determining low-frequency normal modes of macromolecules. *Proteins: Struct. Funct. Genet.* **41**, 1–7.
26. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983). CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.* **4**, 187–217.
27. Huang, C. C., Couch, G. S., Pettersen, E. F. & Ferrin, T. E. (1996). Chimera: an extensible molecular modeling application constructed using standard components. *Pacific Symp. Biocomput.* **1**, 724–728.

Edited by W. Baumeister

(Received 13 August 2004; received in revised form 25 January 2005; accepted 3 February 2005)