

Molecular Mechanisms of Chaperonin GroEL–GroES Function

O. Keskin,^{‡,§,⊥} I. Bahar,^{*,||} D. Flatow,[‡] D. G. Covell,[§] and R. L. Jernigan[‡]

Molecular Structure Section, Laboratory of Experimental and Computational Biology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, Computational Technologies Laboratory, Screening Technologies Branch, Developmental Therapeutics Program, National Cancer Institute-Frederick, National Institutes of Health, Frederick, Maryland 21702, and Center for Computational Biology and Bioinformatics and Department of Molecular Genetics and Biochemistry, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania 15213

Received July 3, 2001; Revised Manuscript Received October 3, 2001

ABSTRACT: The dynamics of the GroEL–GroES complex is investigated with a coarse-grained model. This model is one in which single-residue points are connected to other such points, which are nearby, by identical springs, forming a network of interactions. The nature of the most important (slowest) normal modes reveals a wide variety of motions uniquely dependent upon the central cavity of the structure, including opposed torsional rotation of the two GroEL rings accompanied by the alternating compression and expansion of the GroES cap binding region, bending, shear, opposed radial breathing of the *cis* and *trans* rings, and stretching and contraction along the protein assembly's long axis. The intermediate domains of the subunits are bifunctional due to the presence of two hinges, which are alternatively activated or frozen by an ATP-dependent mechanism. ATP binding stabilizes a relatively open conformation (with respect to the central cavity) and hinders the motion of the hinge site connecting the intermediate and equatorial domains, while enhancing the flexibility of the second hinge that sets in motion the apical domains. The relative flexibilities of the hinges are reversed in the nucleotide-free form. Cooperative cross-correlations between subunits provide information about the mechanism of action of the protein. The mechanical motions driven by the different modes provide variable binding surfaces and variable sized cavities in the interior to enable accommodation of a broad range of protein substrates. These modes of motion could be used to manipulate the substrate's conformations.

The activity of the bacterial chaperonin GroEL and its cofactor GroES, and the molecular machinery underlying this activity, have been investigated by several groups in recent years (1–3). The role of GroEL is twofold. First, it binds partially folded or misfolded proteins in its central cavity, thus preventing their aggregation (4, 5). Second, the central cavity works as an Anfinsen cage in which the isolated protein folding intermediate is actively folded, after being unfolded, if necessary, to states more committed toward correct folding (6, 7).

The crystal structure of the GroEL–GroES complex was determined by Sigler and co-workers (7). The space-filling model of the complex is displayed in Figure 1a. GroEL has the form of a double heptameric ring, with a large central cavity in which the substrate protein is bound. The two rings stack back to back. Their (seven) subunits are each comprised of three domains (Figure 1): apical (red), intermediate (green), and equatorial (blue). Two noncontiguous segments at the N- and C-termini of the molecule form the equatorial

domain. These will be designated here as subdomains E1 (residues 1–137) and E2 (residues 411–525). Likewise, the intermediate domain consists of two segments, I1 (residues 138–192) and I2 (residues 376–410), while the apical domain (A) consists of a single sequence connecting the intermediate domains (residues 193–375) (7). The co-chaperonin GroES also exists as a heptamer. It can bind to either GroEL ring to form a cap on the central cavity. In the absence of GroES, the two rings have identical three-dimensional structures, symmetric with respect to the inter-ring interface. In the form complexed with GroES and ADP, on the other hand, they assume different conformations, and they are distinguished as the *cis* (GroES- and ADP-bound) and *trans* (GroES- and ADP-free) rings. See the conformations of the subunits belonging to the *cis* and *trans* rings in panels b and c of Figure 1, respectively.

The individual domains of the subunits have been proposed (8, 9) to engage in specific roles during chaperonin action. The apical domain (A) recognizes the folding intermediates to be sequestered in the central cavity, as well as the flexible loops at the binding interface of GroES. Their interior cavity surface is lined with hydrophobic residues in the polypeptide acceptor state, which can bind non-native polypeptides or help unfold misfolded intermediates that expose hydrophobic surface regions (1, 10, 11); in the protein-folding and protein-release states, on the other hand, the interior cavity lining becomes hydrophilic (1, 7). Substrate sequestration in such a cagelike environment distinct from the bulk solution is an important element of the chaperonin mechanism (12, 13).

* To whom correspondence should be addressed: Center for Computational Biology and Bioinformatics, School of Medicine, University of Pittsburgh, Suite 601, L. Kaufmann Building, 3471 Fifth Ave., Pittsburgh, PA 15213. Phone: (412) 648-6671. Fax: (412) 648-6676. E-mail: bahar@pitt.edu.

‡ Center for Cancer Research, National Cancer Institute, National Institutes of Health.

§ Developmental Therapeutics Program, National Cancer Institute-Frederick, National Institutes of Health.

⊥ Present address: Koc University, College of Arts and Sciences, Sariyer 80910, Istanbul, Turkey.

|| University of Pittsburgh.

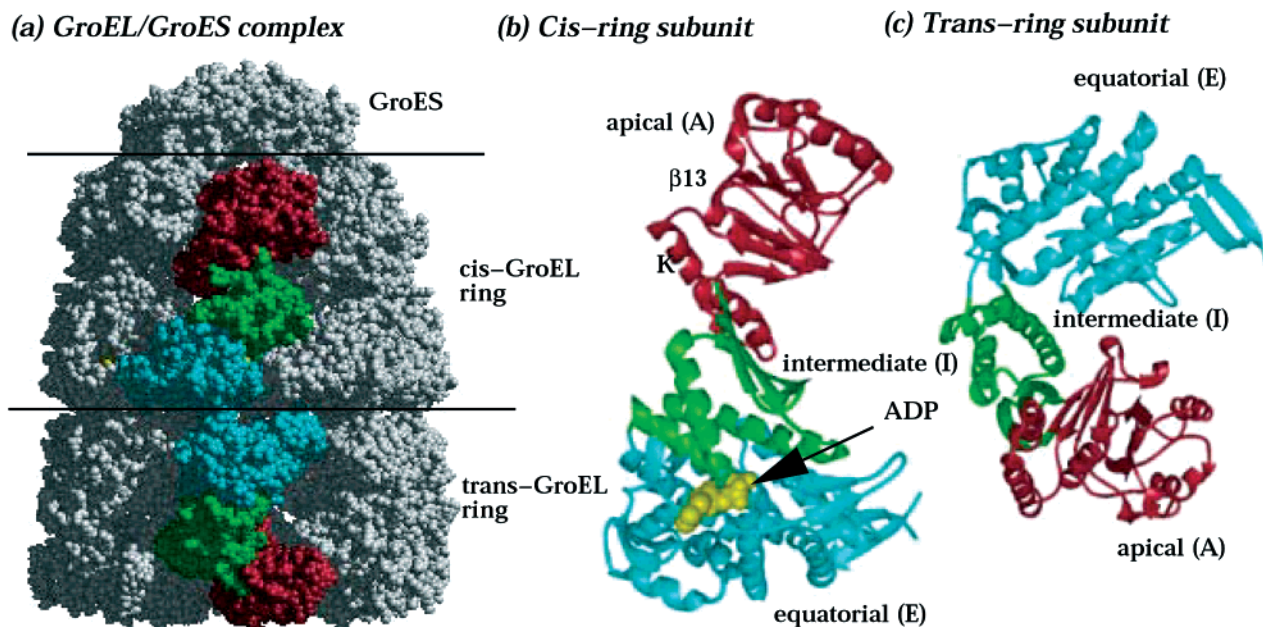


FIGURE 1: Structure of the GroEL–GroES complex. (a) Space filling model from the crystal structure determined by Xu et al. (7) (PDB entry 1aon). The GroEL molecule has a barrel-like shape composed of two rings, termed the *cis* and *trans* rings, depending on the position of the GroES cap. Each ring is comprised of seven subunits, as well as the GroES cap. Panels b and c display structures of the subunits belonging to the *cis* and *trans* rings, respectively. Each subunit consists of three domains: equatorial (blue), intermediate (green), and apical (red). An ADP molecule bound to the equatorial domain of the *cis* ring subunit is displayed in panel b. K and β 13 identify the positions of helix K and β -strand 13, respectively.

The equatorial domain (E) binds and hydrolyzes ATP. It plays a major role in mediating the overall chaperonin activity by an ATP-dependent allosteric mechanism, and the intermediate domain (I) serves as a functional bridge between the apical and equatorial domains (9).

A schematic description of one cycle of GroEL–GroES action is presented in Figure 2 (2, 14–16). First, the GroEL binds a partially folded or misfolded intermediate at its apical domain (step 1). Binding is facilitated by the hydrophobic interactions between the exposed hydrophobic residues of the substrate protein and those of the apical domain as well as electrostatic interactions (17). Then, seven ATP molecules bind, and the molecule proceeds to bind the cofactor GroES in such a way that the GroEL and GroES molecules form a capsule around the refolding protein (step 2). GroEL is unable to bind the GroES cap in the absence of ATP. GroES binding is accompanied by a large conformational change at the apical domain of the *cis* ring leading to a 2-fold enlargement of the central binding cavity (7). The protein substrate is accommodated inside the cavity by this conformational change, where, protected against binding to other proteins, it will be able to undergo conformational changes toward the folded state. The next step (step 3) is the hydrolysis of the ATPs in the *cis* ring. The energy released during this reaction is likely to be absorbed by the substrate protein to surmount the energy barriers to unfold and fold correctly. This process is succeeded by binding of ATPs to the opposite (*trans*) ring, which drives the release of the GroES cap and the substrate protein (step 4). The opposite, originally *trans*, ring thus becomes the folding-active *cis* ring for the next round of the chaperonin cycle (18). The released protein goes through several cycles until it is correctly folded (2, 15, 19–21). Large conformational fluctuations during the chaperonin cycle, and the accompanying changes in the volume and hydrophobicity of the central cavity, are crucially

important for the efficiency of chaperonin-assisted folding. These conformational changes exemplify the relevance of flexibility to biological activity (14).

Computational studies by two different groups (22–25) have provided important insights into the allosteric mechanism of the chaperonin GroEL. Karplus and co-workers (24) analyzed the conformational transitions between the closed and open states of the individual subunits using targeted molecular dynamics simulations. The opening of the subunits was shown to be accommodated by the heptameric ring only if accompanied by a concerted twisting and upward displacement of the apical domains. Ma and Karplus (25) performed a normal-mode analysis (NMA) of a construct of three equatorial domains (one in the *cis* ring and two in the *trans* ring) and reported a positive cooperativity in one ring and a negative cooperativity between the two rings. Berendsen and co-workers (22), on the other hand, considered the fluctuations in the residue positions of five different, crystallographically determined, ring structures. The associated covariance matrix was diagonalized for obtaining a set of conformations sampled by the essential modes of motion. This analysis indicated the role of the fluctuations near the GroES and nucleotide binding sites in establishing the inter-ring communication, as well as the importance of the double-ring structure for the allosteric coupling between the nucleotide-induced conformational changes and the GroES-mediated transitions (22).

In the study presented here, we analyze the collective dynamics of the entire GroEL–GroES complex using a simple, but rigorous, mathematical model based on statistical mechanical principles. Previous theoretical studies of chaperonin dynamics have explored either the individual subunits (or groups of subunits) (23–25) or statistical subsets of the complete conformational space (22). Our approach deviates

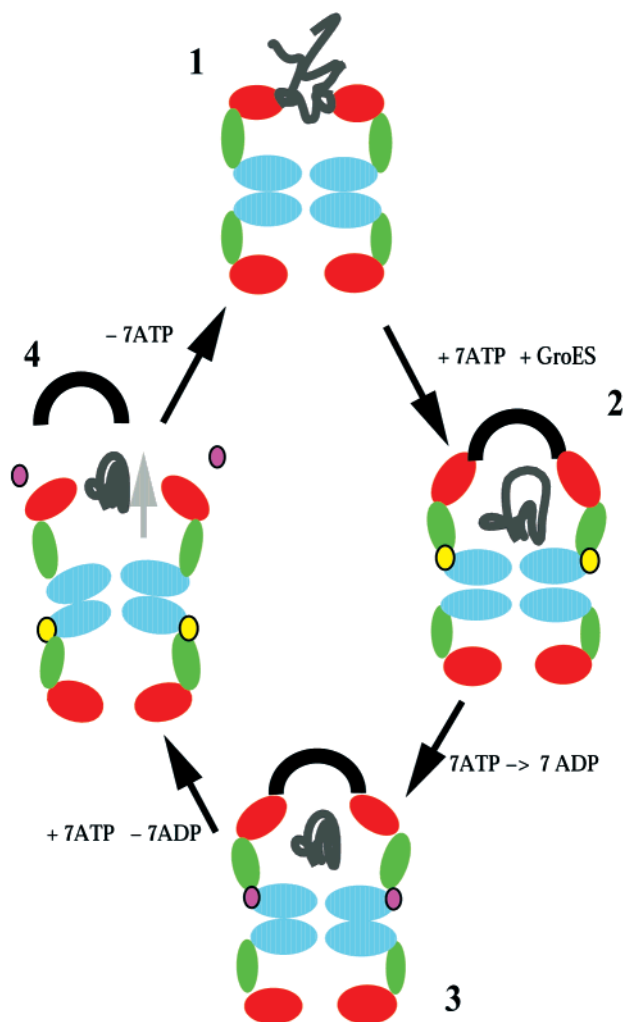


FIGURE 2: Schematic description of the chaperonin activity cycle. Blue, green, and red regions designate the equatorial, intermediate, and apical domains of the subunits, respectively, consistent with Figure 1. The co-chaperonin GroES and the substrate are colored black. ATP is shown in yellow and ADP in magenta. The four stages are as follows: (step 1) binding of a partially folded or misfolded polypeptide to the apical domain; (step 2) binding of seven ATP molecules to the equatorial domains of the *cis* ring subunits, induced enlargement of the *cis* ring cavity that encapsulates the substrate polypeptide, and binding of the GroES cap; (step 3) hydrolysis of ATP and the folding and/or processing of the substrate; and (step 4) binding of ATP to the *trans* ring, followed by dissociation of the GroES cap and release of the folded (or partially folded) protein.

from simulations in that it lends itself to a deterministic (or unique) solution, for the mechanisms of motions of a given structure. This information about the *global* machinery of the intact oligomer (of 21 subunits) is extracted at the cost of specificity loss at the *local* (atomic) scale. Our approach is to model the overall structure as a network of interacting residues, following the Gaussian network model (GNM) (26–29) of proteins and its anisotropic network extension model (ANM) that has recently been introduced (38, 39). We will extract the most cooperative collective motions of the overall structure after a decomposition of the spectrum of modes.

Our results indicate that the complex can undergo a diversity of molecular mechanisms that underlie its chaperone action. The intermediate domains play a pivotal role in

coordinating the motions of the complex, being coupled either to the equatorial domains (for the *cis* ring, i.e., ATP- and GroES-bound form of subunits) or to the apical domains (*trans* ring). This dual character is ensured by two hinge sites, originally proposed by Xu et al. (7), one near the apical domain (hinge 1, centered at Gly192 and Gly375) and the other near the equatorial domain (hinge 2, at Pro137 and Gly410), which are alternatively activated and frozen by a switch mechanism controlled by ATP binding and hydrolysis. Nucleotide binding locks *hinge 2* at the interface between domains I and E, and stabilizes an “open” conformation of the ring, open from the perspective of the interior of the GroES molecule, or closed with respect to the nucleotide binding site. This has two effects. (i) The central cavity enclosed by the ATP-bound ring is enlarged, and (ii) the apical domains of the same ring exhibit an enhanced mobility supported by active motions about hinge 1 (connecting domains I and A). The former effect ensures either the accommodation of the substrate [for ATP molecules bound to the *cis* ring (step 2)] or the release of the folded protein from the *cis* ring [for ATPs bound to the *trans* ring (step 4)]. And the latter effect either facilitates the recognition and binding of the GroES cap by exposing and moving the apical domains (step 2) or promotes the release of GroES by dissociating the *cis* ring apical domains (step 4).

MATERIALS AND METHODS

An important issue in the analysis of protein dynamics is the identification of the mechanism of slow, large-amplitude motions, also termed global motions. These are conceived to be essential for protein function (30–34). One standard technique for studying protein dynamics, and in particular low-frequency domain motions, is NMA (35–37). The anisotropic network model (ANM) (38, 39) presently assumes that a protein in its folded state is equivalent to a network, in analogy to the elasticity theory of random polymer networks (40). This is an extension of the previously developed Gaussian network model (GNM) (26, 27) to include directional anisotropic effects on fluctuation dynamics. In principle, the GNM is equivalent to NMA with one additional simplifying assumption, that being that there are no directional effects, i.e., all fluctuations are isotropic. The GNM thus predicts the relative sizes of fluctuations but not their directionalities. The ANM does predict the directionalities, and permits us to decompose the molecular motions into a series of $3N - 6$ modes ($N = 8015$ for the GroEL–GroES complex).

In general, the intramolecular potential V of a protein of N residues may be expressed as a series expansion in the fluctuations ΔR_i of individual residue positions, and it can be written in terms of the Hessian matrix (\mathbf{H}) of the second derivatives of V as

$$V = \frac{1}{2} \sum_i \sum_j (\partial^2 V / \partial \Delta R_i \partial \Delta R_j)_0 \Delta R_i \Delta R_j = \frac{1}{2} [\Delta \mathbf{R}]^T \mathbf{H} [\Delta \mathbf{R}] \quad (1)$$

where $\Delta \mathbf{R}$ represents the $3N$ -dimensional array composed of the N fluctuation vectors ΔR_i and the superscript T designates the transpose. The conformational potential of the structure takes the form

$$V = \frac{1}{2}\gamma \sum_i \sum_j (\Delta R_i - \Delta R_j) \cdot (\Delta R_i - \Delta R_j) h(R_c - R_{ij}) \quad (2)$$

where \cdot designates the dot product, γ is the spring constant for all inter-residue interactions, and $h(R_c - R_{ij})$ is the Heaviside step function, which takes the value 1 if the argument is positive (if the inter-residue distance R_{ij} remains below the cutoff distance R_c) and 0 otherwise. \mathbf{H} can be conceived as an $N \times N$ matrix of 3×3 superelements defined as $\mathbf{H}_{ij} = \partial^2 V / \partial \Delta R_i \partial \Delta R_j$. Using the classical theory of networks (40), it can be shown that the mean-square (ms) fluctuations of junctions scale with the inverse of \mathbf{H} as

$$\langle (\Delta R_i)^2 \rangle = kT \text{trace}[\mathbf{H}^{-1}]_{ii} \quad (3)$$

where $[\mathbf{H}^{-1}]_{ii}$ is the i th diagonal superelement of the inverse \mathbf{H}^{-1} and trace designates the sum of the diagonal elements of this 3×3 matrix. Equation 3 will be used in this study for calculating the ms fluctuations of individual residues, and the global motions will be extracted by a mode analysis of \mathbf{H} . To this end, \mathbf{H} is rewritten as the product of three matrices, the diagonal matrix Λ of its non-zero eigenvalues λ_m ($1 \leq m \leq 3N - 6$), the matrix \mathbf{U} of the corresponding eigenvectors \mathbf{u}_m , and the transpose of \mathbf{U}

$$\mathbf{H} = \mathbf{U}\Lambda\mathbf{U}^{-1} = [\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3 \ \dots \ \mathbf{u}_{3N-6}] \text{diag}(\lambda_1 \ \lambda_2 \ \lambda_3 \ \dots \ \lambda_{3N-6})[\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3 \ \dots \ \mathbf{u}_{3N-6}]^T \quad (4)$$

where \mathbf{U} is an orthonormal matrix, i.e., $\mathbf{U}^T = \mathbf{U}^{-1}$, and the inverse of \mathbf{H} can be readily written as $\mathbf{H}^{-1} = \mathbf{U}\Lambda^{-1}\mathbf{U}^{-1} = \mathbf{H}^{-1} = \sum_m [\lambda_m^{-1} \mathbf{u}_m \mathbf{u}_m^T]$, such that the ms fluctuation of residue i driven by the m th mode of motion becomes

$$[(\Delta R_i)^2]_m = kT \text{trace}[\lambda_m^{-1} \mathbf{u}_m \mathbf{u}_m^T]_{ii} \quad (5)$$

The orientational cross-correlations between the fluctuations of residues i and j are given by

$$C(i,j) = \frac{\langle \Delta R_i \cdot \Delta R_j \rangle}{(\langle \Delta R_i \cdot \Delta R_i \rangle \langle \Delta R_j \cdot \Delta R_j \rangle)^{1/2}} \quad (6)$$

where $\langle \Delta R_i \cdot \Delta R_j \rangle$ is calculated from the equation $\langle \Delta R_i \cdot \Delta R_j \rangle = kT \text{trace}[\mathbf{H}^{-1}]_{ij}$ (eq 3). The two limits of $C(i,j)$ are 1 and -1 , and correspond to pairs of residues exhibiting fully correlated (same direction, same sense) and fully anticorrelated (same direction, opposite sense) motions, respectively. Zero refers to uncorrelated or orthogonal motions.

Equation 6 cannot distinguish between pairs of residues subject to random (uncorrelated) motions and pairs undergoing correlated but orthogonal fluctuations. Other measures of describing correlated domain motions have been developed, which take account of the couplings of orthogonal rotational motions (35, 41, 42). We use eq 6 here because we focus on the functionally important cross-correlations that occur between the intermediate domains and the neighboring apical or equatorial domains, i.e., between pairs of domains located at a given rotational angle about the cylindrical axis.

RESULTS AND DISCUSSION

Temperature Factors

Figure 3a compares the temperature factors (B -factors) predicted by the ANM theory (—) with the experimental

results (7) found from the X-ray crystallography (\cdots). The 7-fold periodicity consistent with the heptameric structure can easily be discerned. Panel b gives a more detailed view of the dynamic behavior for representative monomers belonging to the *trans* and *cis* rings, and to the GroES cap. The apical domains (residues 193–375) are observed to be the most mobile parts of the subunits, along with the GroES molecule. These form the maxima on the curve. The equatorial domains (residues 1–137 and 411–523) form the minima. The high flexibility of the apical and intermediate domains was also observed by Ma and Karplus (25). The *trans* ring enjoys a higher mobility than the *cis* ring, especially near the peptide segment of residues 331–360 of the subunits, which includes β -strand 13, α -helix K, and the loop connecting them. These are solvent-exposed elements in the apical domains of the *trans* subunits (see Figure 1c). Unfolding studies of GroEL suggest that the apical and intermediate domains unfold first, while the equatorial domains continue to hold the two rings together (43). This is consistent with the presently observed high stability (small fluctuations) of the equatorial domains.

The curve in Figure 3a has been obtained using the first 40 dominant (lowest-frequency) modes elucidated by the ANM. The correlation coefficient between the theoretical and experimental data is 0.9. We obtain approximately the same level of agreement with experimental data even using the first 10 dominant modes, which indicates the dominant role of a small subset of low-frequency modes in determining the observed dynamics. Panel c in the same figure shows the cumulative contribution of the $3N$ eigenvalues to the observed dynamics. The first 10 modes account for $\sim 60\%$ of the overall dynamics, and the first 40 modes amount to $\sim 75\%$. We focus on the effect of the individual lowest-frequency modes in the following subsection.

Mechanisms of the Dominant Modes of Motion and Their Relevance to Function

The six panels in Figure 4 illustrate the fluctuating conformations visited by the action of the dominant modes of motions, some of them related by radial symmetry being combined.

Mode 1: Torsion about the Cylindrical Axis of symmetry. In this mode, the equatorial domains serve as support regions about which the apical domains and part of the intermediate domains rotate. The directions of rotation of the two rings oppose one another; i.e., this is a global back-and-forth torsional motion involving the overall cylindrical structure, with the interfacial plane between the two rings being almost fixed in space. The ms fluctuation distribution in Figure 5 (top panel) shows the suppressed mobilities of equatorial domains E1 and E2 of both rings. The apical domains of the *cis* ring subunits, on the other hand, undergo large-amplitude coherent fluctuations consistent with previous observations (7, 44). The high mobilities of the *cis* ring apical domains are likely to be a prerequisite for the efficient recognition of the substrate and the cofactor. The structural deformation driven by mode 1 also induces an expansion at the GroES binding interface. See the movie of this global mode at <http://www.cccb.pitt.edu/groel.avi>.

The ribbon diagrams displayed in Figure 4 for mode 1 show the top view of two fluctuating conformations of the

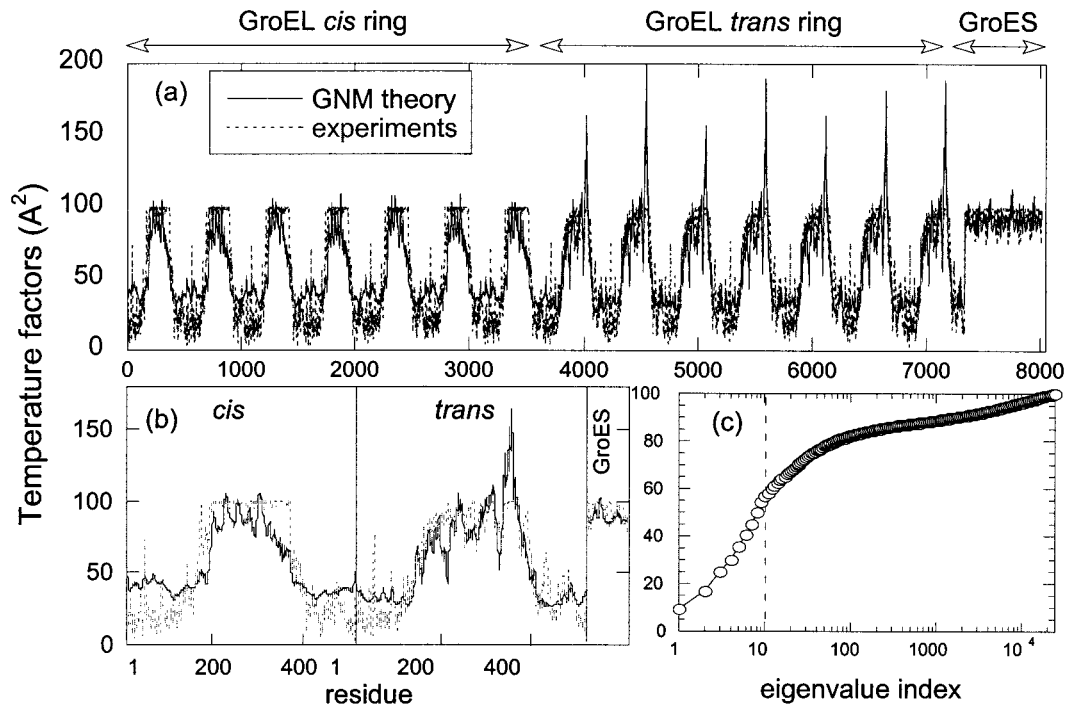


FIGURE 3: Temperature factors. (a) Comparison of the theoretical (solid curve) and experimental (dotted curve) temperature factors for the GroEL–GroES complex. Results in panel a are shown for all 21 subunits of the chaperonin complex. Note the 7-fold repeat pattern induced by the heptameric structure of the rings and the cap. Panel b shows the same curves enlarged for individual monomers, belonging to the *cis* and *trans* subunits, and to the GroES cap. A subset of low-frequency modes (40 global modes) that dominate the fluctuation spectrum are included in the calculations. The correlation coefficient between theoretical and experimental results is found by least-squares fit to be 0.90. (c) Cumulative contribution of all modes ($3N$ modes) to the observed dynamics. The abscissa is the eigenvalue index, and the cumulative percent contribution of all eigenvalues at the lower-frequency end of the spectrum, for a given index.

trans ring, in red and blue (on the left), and the side views of two adjacent subunits, in *cis* and *trans* rings, in the two fluctuating conformations (pink and green, on the right). The arrows indicate the displacement directions for the apical domains. Ma and Karplus (24) found that in the lowest-frequency mode, domains A and I in *trans* rings rotate together with respect to domain E. This type of twisting motion is very similar to our first mode and resembles the structural changes described by Xu et al. (7).

A closer examination of mode 1 reveals, on the other hand, significant differences in the behavior of the *cis* and *trans* rings. The bottom curve in Figure 5 displays the change in the fluctuations of *cis* ring residues, relative to those of the *trans* ring (bottom part). The most significant difference is the suppression of the mobilities of the intermediate domains (and the C-terminal portion of the apical domains) in the *cis* ring, relative to those of the *trans* ring. We also see an enhancement in the mobility of certain residues in the apical domains of the *cis* ring. Ma and Karplus (25) compared the fluctuations of the two subunits in an ATP- and GroES-free structure and in an ATP-bound GroES-free structure considering the cumulative effect of the 10 slowest modes. They found that the equatorial domain is more rigid, and both the intermediate and apical domains are more flexible in the ATP-bound (*cis*) subunits. The lower flexibility of the *cis* intermediate domains is found here by focusing on mode 1 for the intact complex of 21 subunits. This restriction in mobility is induced by the coupling of the *cis* ring intermediate domains to the almost rigid equatorial domains of the same ring. Hinge 2 in these subunits is indeed frozen by ADP binding. On the other hand, the higher flexibility of the apical domains of the *cis* ring subunits compared to their

counterparts in the *trans* ring is imparted by the active hinge 1 present in the ADP-bound form. The opposite situation occurs in the *trans* ring subunit, where mainly hinge 2 is operative and hinge 1 is inactive. Thus, in the absence of ADP, the molecule flexes at the interface between the equatorial and intermediate domains.

In a sense, the intermediate domains are coupled either to the equatorial domains (*cis* ring) or to the apical domains (*trans* ring) of the same ring. This fundamental difference between the dynamics of the intermediate domains in the two rings, which is an allosteric effect imparted by ADP binding, emerges by considering the intact complex, and becomes more evident upon direct examination of the cross-correlations between domain motions (see below).

Modes 2 and 3: Bending of the Complex at the Interface between the Cis Ring and the GroES Molecule, Accompanied by a Slight Rotation of the Trans Ring. The right top panel in Figure 4 displays the conformational fluctuations driven by orthogonal modes 2 and 3. Two of the subunits, belonging to the respective *cis* and *trans* rings, are colored green and yellow, to illustrate the alternative conformations of the same subunits. Calculations show that residues 189–195, 363–375, 171–173, 327, and 328 in the *cis* ring are almost fixed in space during this mode (not shown). These include hinge 1 crossover residues Gly192 and Gly375 previously shown (7) to link the apical and intermediate domains, suggesting that hinge 1 is controlled and/or activated by these modes. GroES is the most flexible part of the complex in this mode. See Figure 6 which shows alternative conformations of a *cis* subunit, which are sampled by the action of modes 2 and 3. The equatorial domain and the neighboring part of the intermediate domain, shown in blue and green, respec-

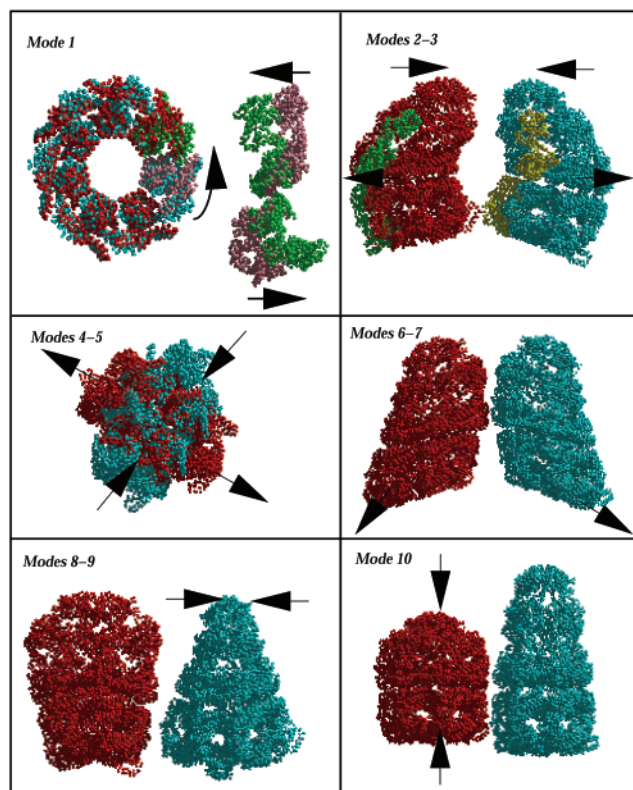


FIGURE 4: Alternative conformations visited by the action of the global modes. The associated mechanisms of motions are torsion about the cylindrical axis of symmetry (mode 1), bending near the GroEL–GroES interface (modes 2 and 3), compression and/or extension of the rings (*trans* ring, in particular) into oblate forms (modes 4 and 5), global deformation along the diagonal axis (modes 6 and 7), alternating radial breathing of the apical domains of the *cis* and *trans* rings (modes 8 and 9), and elongation and/or contraction of the complex near the GroES cap (mode 10).

tively, exhibit little change, while the apical domain enjoys a large motion. The high mobility of the apical domains is presumably functional in nature for facilitating an optimal interaction with the GroES cap together with substrate proteins.

Modes 4 and 5: Lateral Expansion and Compression of the Rings. Here the circular symmetry of the rings is manifested by a combination of two orthogonal modes that differ by 90° in their symmetry axes. In the figure, we displayed only the top view of the superimposed fluctuating conformations of the *trans* ring.

Modes 6 and 7: Shear Motion with Extension along a Diagonal Longitudinal Axis. The molecule fluctuates between two conformations distorted along a diagonal axis, as illustrated in the figure. This mechanism of motion does not involve a net change in the volume of the internal cavity but allows for cooperative geometric distortions that likely affect interactions with the encapsulated protein.

Modes 8 and 9: Alternating Breathing of the Apical Domains. This mode involves an expansion of the apical domains of the *trans* ring, accompanied by a contraction of *cis* ring apical domains, followed by the opposite process. A substantial increase in the volume of the expanded ring occurs. The internal association of the apical domains is necessarily weakened during the expansion stage. We recall that the GroEL–GroES complex is very stable in the presence of ATP and cannot dissociate until the ATP bound

by the *cis* ring is hydrolyzed. Hydrolysis weakens the interaction between the GroES molecule and the GroEL apical domains of the *cis* ring and primes the *cis* chamber for dissociation. Binding of ATP to the *trans* ring further triggers the disassembly of the weakened *cis* ring apical domains, releasing the GroES cap, the ADP molecules, and the substrate protein, and completes a single round of the folding cycle (18). The global motion observed in this mode, along with mode 1, may contribute to the opening or closing of the entrance of the central cavity, required both for the binding (*cis* ring) and release (*trans* ring) of substrate. Binding of denatured glutamine synthetase monomers induces structural alterations in GroEL heptamers such as inward movements of apical domains of both rings and narrowing of the *trans* ring (45, 46). De Groot et al. (22) also observed inward and outward motions along their first two principal components in their calculations, which could be attributed to these modes.

Mode 10: Longitudinal Elongation and Compression of the Entire Complex. A substantial change in the overall shape of the complex occurs in this mode, which can be described as an overall stretching and/or contraction along the cylindrical axis of symmetry. The GroES cap undergoes large displacements during this mode. The low packing density at the interface between GroES and GroEL favored by this mode suggests its possible involvement in the dissociation of the cap, or perhaps a syringelike intake of substrate.

Cross-Correlations between Domain Motions

Figure 7 shows the correlations between residue fluctuations for a *cis* subunit (top map), for a *trans* subunit (middle map), and for the overall complex together with GroES (bottom map). A large subset of modes (41 of them) is considered, for a clear visualization of the types of cooperativity between the motions of the individual domains. We note that the top two maps represent only small portions (the $7 + 7$ diagonal squares) of the overall correlation map displayed at the bottom. The labels indicate the different domains of the subunits, and the color codes on the right side indicate the type and strength of cross-correlations between the fluctuations of the individual residues. Red regions refer to positively correlated pairs of residues, i.e., those undergoing concerted motions in the same direction. Blue regions, on the other hand, refer to pairs of anticorrelated residues, or residues undergoing opposite direction fluctuations. Both extreme cases correspond to strongly coupled, concerted motions, either in the same direction or in opposite directions, whereas yellow-green regions are uncorrelated pairs. The transitions between the red and blue regions indicate the loci of hinge-bending sites. These are highlighted by the white lines.

The first observation in the top two maps is the coupled, concerted motions of the individual domains, indicated by the red blocks along the diagonal. The off-diagonal red blocks also refer to the noncontiguous segments of the same domains, consistent with the coherent motion of the individual domains. The most important difference between the maps for the *cis* and *trans* subunits resides in the cross-correlations involving the intermediate domains. See the block corresponding to the correlations between the E1 and I1 subdomains, as well as those referring to the I1, I2, and A domains.

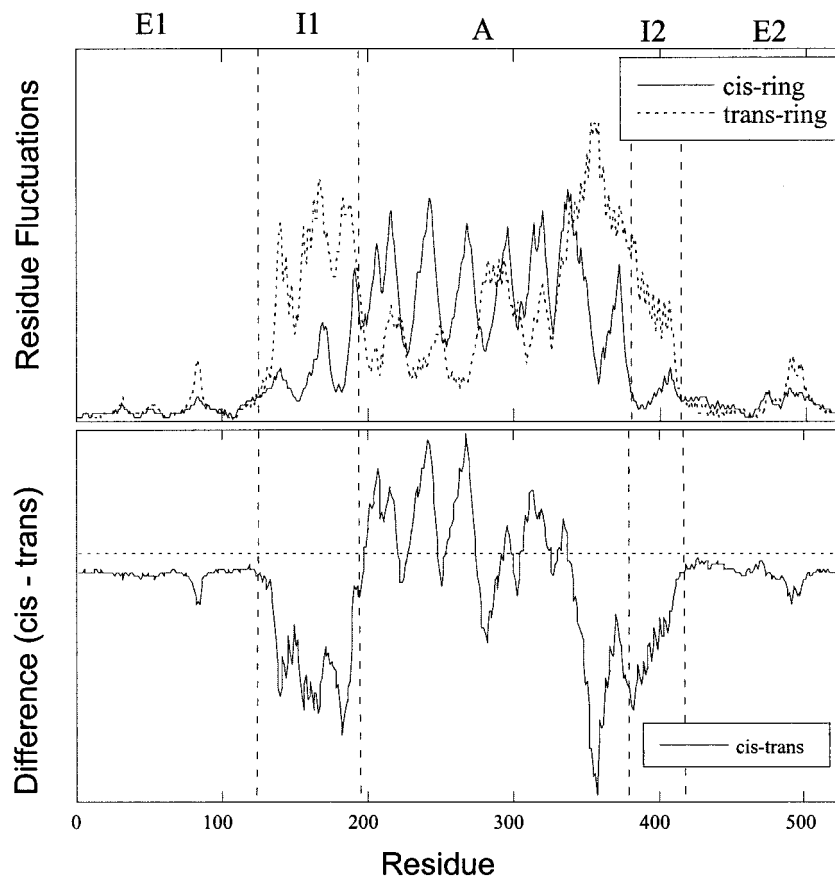


FIGURE 5: Differences between the residue fluctuations of *cis* and *trans* rings. Distribution of residue fluctuations in the *cis* ring (—) and *trans* ring (···) subunits. Note that the equatorial domains, consisting of two separate sequences, E1 and E2, are severely confined in both rings, while the intermediate and apical domains enjoy relatively high flexibilities in the *trans* and *cis* rings, respectively. The bottom curve indicates the difference in the flexibility of the *cis* ring subunit, relative to that of a *trans* ring subunit. The reduced flexibility of the apical domain in the *trans* ring is attributed to its coupling to the intermediate domain, while the same (apical) domain of the *cis* ring is decoupled from the other parts of the subunit.

In general, the intermediate domain acts to communicate between the apical and equatorial domains. Yet in the *cis* subunit, it is positively correlated with the equatorial domain, whereas in the *trans* subunit, it is positively correlated with the apical domain. Its coupling to the equatorial domain in the *cis* ring is an ATP-dependent effect, already noted above in the examination of the first global mode. In the absence of ATP (the *trans* ring in the structure presented here), on the other hand, it has a tendency to move in concert with the apical domain.

The lowermost map shows the cross-correlations for the overall structure. In general, each subunit in the rings is highly correlated with its neighboring subunits in the same ring. The red positive correlations between the *cis* and *trans* rings, on the other hand, refer to adjacent subunits of the two rings. The regions of the GroEL making contact with the GroES cap (i.e., the M-loops in the apical domain of the *cis* rings) are positively correlated with the GroES. GroES motions are generally negatively correlated with those of the *trans* ring. Finally, GroES subunits are highly correlated (lowermost right square), suggesting that the cap undergoes cooperative motions *en bloc*.

Figure 8 displays a schematic view of the two types of coupling of the intermediate domains, revealed by the top two maps in Figure 7. Panel a shows a subunit of the *cis* ring and panel b one from the *trans* ring. In the *cis* ring, the apical domain is shown in darker gray, whereas the equatorial

and intermediate domains, which are found to be engaged in coupled correlated fluctuations, are shown in lighter gray. The ADP molecule is shown in black. The darker and lighter regions thus represent two anticorrelated substructures, A and I–E. The location of hinge 1 at residues Gly192 and Gly375 separating the two anticorrelated substructures is shown on the left. In the *trans* ring, on the other hand, the apical domain is coupled to the intermediate domain (both shown in darker gray). Hinge 2 residues Gly410 and Pro137 facilitate the anticorrelated motions of the two substructures, A–I and E, in this case.

These figures illustrate how the hinge site changes upon ADP and GroES binding. In the *cis* ring, as described in the study by Xu et al. (7), the intermediate domain swings down toward the equatorial domain, pivoting approximately 25°, around Pro137 and Gly410. This movement closes the nucleotide binding site, and generates new interactions with the equatorial domain, both within the same subunit and in the neighboring subunits. The present correlated motions indicate that the presence of a bound nucleotide is sufficient to join the equatorial and intermediate domains into a united substructure. This reduction in entropy is however counterbalanced by the enhanced mobility of the apical domain, ensured by the activation of hinge site 1. Indeed, it has been pointed out that the apical domain swings up 60° relative to the equator and twists around the long axis of the domain by 90°, leading to an interaction with the mobile loop (M-

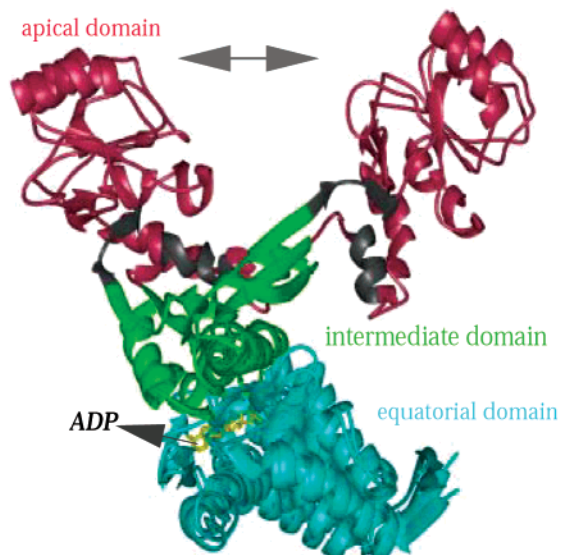


FIGURE 6: Fluctuating conformations of *cis* subunits in modes 2 and 3. Equatorial domains are superimposed to show the displacements of the intermediate and apical domains. ADP is shown in yellow. A significant change in the position of the apical domain occurs, leading to an overall bending of the complex (see the top right of Figure 4). The hinge-bending sites are identified from the minima of the mode shape curve in Figure 3b as the stretches of residues 363–375 and 189–195. These regions are colored gray.

loop) of the GroES (7). Ma and Karplus (25) also indicate that structural changes induced by ATP binding in the individual subunits appear to be a primary source of the positive cooperativity within a ring.

Fluctuations of Hydrophobicity of the Cavity

Previous analyses using simple models suggested that a moderately hydrophobic confinement assists the folding process by lowering barriers to folding (48). However, in the case of confinement in an overly hydrophobic cavity, folding could become inefficient because of many strong substrate–cavity interactions (14). Furthermore, a cavity with periodically varying hydrophobicity has been shown to reduce the folding time (14). A change in the hydrophobicity of the walls lining the cavity can occur either due to periodic changes in conformation during the chaperonin cycle or due to conformational fluctuations driven by the cooperative modes of motion of the molecule. Here, we explore the changes in the hydrophobicity associated with the individual global modes of the GroEL–GroES complex.

Figure 9 is analogous to Figure 4, but instead shows views of the central interior cavity colored by the residues' hydrophobicities, for similar fluctuating conformations as in Figure 4. Residues Val, Leu, Ile, Met, Phe, Trp, Pro, Ala, and Tyr, which can be classified as hydrophobic residues, are colored yellow; Gly is colored gray, and the remainder (polar and charged residues) are colored cyan. The front three subunits have been removed to facilitate viewing of the interior cavity surfaces. Note the yellow patches in the *trans* ring and cyan patches in the *cis* ring of the X-ray structure (top left of Figure 9). The fluctuating conformations driven by the global modes are then shown in the different parts numbered by mode numbers. Also, the substantial changes in volume of the interior of the *cis* and *trans* rings are readily visible in this figure.

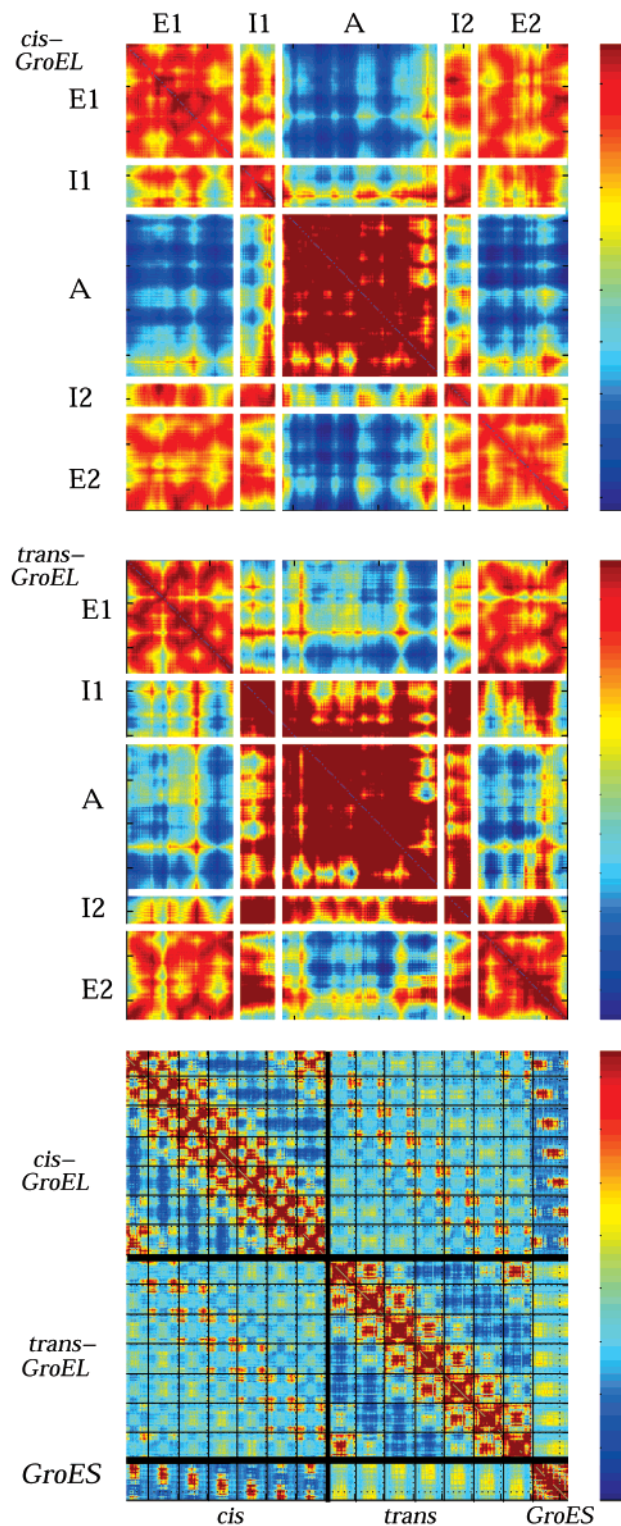


FIGURE 7: Correlation maps. Correlation maps for residue fluctuations in the *cis* subunit, the *trans* subunit, and the overall GroEL–GroES complex, based on the dominant global motions of the complex. The abscissa and ordinate refer to residue numbers, and the red and blue regions correspond to strongly correlated (+) and strongly anticorrelated (–) pairs of residues, respectively. Correlated pairs undergo concerted same direction, same sense fluctuations; anticorrelated pairs, on the other hand, undergo same direction, opposite sense motions. Note the differences in the types of correlations exhibited by the intermediate domains in the two rings.

In the first global mode, there is no visible change in the hydrophobicity of the rings, but their relative positions do

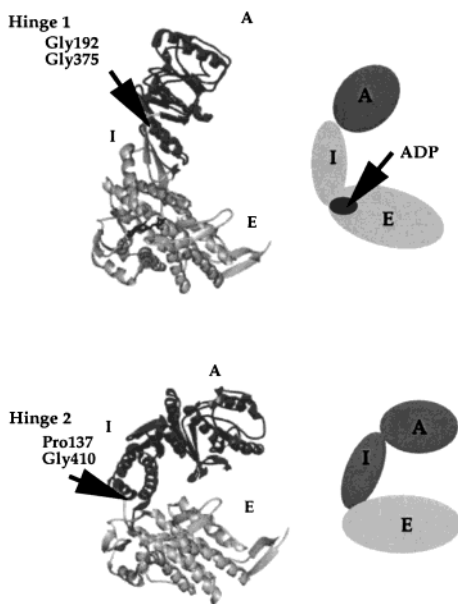


FIGURE 8: Coupling of the three domains. Schematic diagram of the type of coupling exhibited by the intermediate domains of the *cis* (a) and *trans* (b) subunits. The substructures subjected to strongly coupled *en bloc* motions are shown in either dark or light gray. Note the shift in the coupling of the intermediate domain, an allosteric effect induced by nucleotide binding.

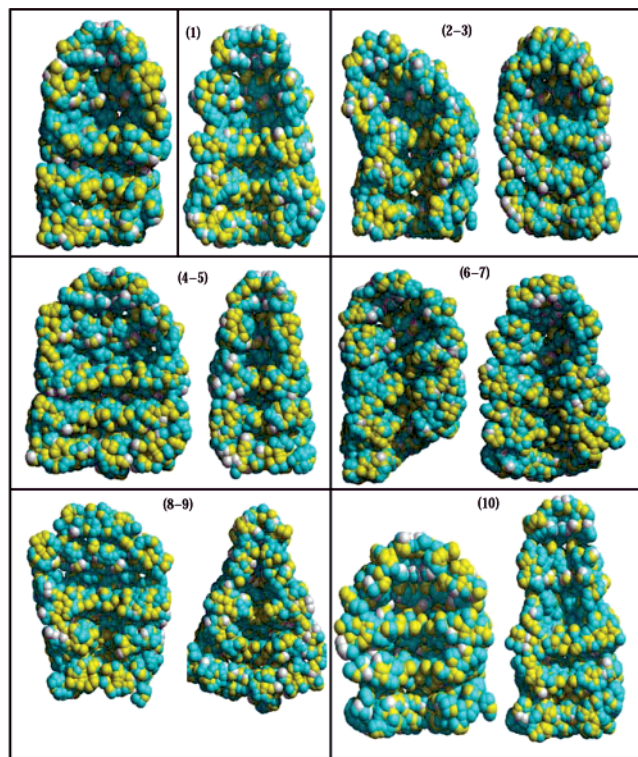


FIGURE 9: Cavity hydrophobicity profiles. The inner walls of the cavity encapsulating the substrate protein in the native conformation, and in the fluctuating conformations driven by global modes presented in Figure 4, are colored by the residues' hydrophobicities. Residues Val, Leu, Ile, Met, Phe, Trp, Pro, Ala, and Tyr, here classified as hydrophobic residues, are colored yellow; Gly is colored gray, and the remainder (polar and charged residues) are colored cyan.

change. This is consistent with providing the essential binding sites for the substrate protein. In the second and third modes (top right box, two side views of the same conforma-

tion), the polarity around the wall of the cavity changes such that the surface becomes more hydrophilic (bluer) on the walls where the bending occurs (see the right panel). This change in polarity might trigger the dissociation of the non-native polypeptide from the wall of the cavity. The released polypeptide would then be free to reinitiate folding as an isolated molecule in the *cis* cavity that has a nonhomogeneous hydrophilic lining conducive to self-burial of its hydrophobic residues.

Fluctuations in the hydrophobicity of the cavity are observed in modes 4 and 5. The hydrophobic residues seem to be more clustered in one form (right) of the fluctuating conformations of these modes. In modes 8 and 9, as the ring breathes, the cavity exposes more hydrophobic residues, and as it shrinks, the polar residues cluster. The density of hydrophobic residues in the GroES binding interface is also subjected to significant changes in these important global modes.

Conserved Residues

The most evolutionarily conserved residues are those located at the ATP and/or ADP binding sites. Hydrophobic residues that contribute to substrate binding are also significantly conserved (49). All the conserved binding regions (Thr30–Pro33, Lys51–Gly53, Asp87–Thr91, Ile150, Ser151, Asp398, Gly414–Gly416, Ile454, Tyr478–Ala481, Ile493, and Asp495) occupy lowermost minima regions in the ms fluctuation curve (Figure 3b), which can be verified by quantitative examination of the results. These residues play a critical role in coordinating the global hinge-bending mechanism near the nucleotide-binding site. They are therefore conserved due to functional requirements.

On the other hand, most of proposed substrate recognition regions (10, 50) are highly conserved. These are located in the apical domain of the GroEL molecule. Recognition sites have been seen in several of our previous fluctuation studies to be located at the maxima of the global mode shapes, their high mobilities presumably permitting more efficient recognition of, and optimal interaction with, substrate. And, consistent with previous studies for other substrate-binding proteins (29), in the GroEL residues involved in the recognition of the substrate are distinguished here by their high mobility. Further studies of the relationship between sequence conservation and structural location vis-à-vis the hinges in individual modes might indicate the relative importance of the individual modes.

CONCLUSION

The wide range of changes in the shape of the internal cavity during the GroEL cycle of activity is believed to play an important role in unfolding and folding the substrate. These are also coupled to ATP binding, to hydrolysis, and to interactions with the co-chaperonin GroES.

In the study presented here, we have explored the switch role of the intermediate domain in establishing communication between the other two domains, apical and equatorial, of the individual subunits, in the presence of the coupling of all 21 subunits forming the GroEL–GroES complex. Kawata et al. (9) showed that for certain proteins, the binding of nucleotides to the GroEL equatorial domain results in the release of refolding protein intermediates from the GroEL

central cavity. This suggested that binding of nucleotides induces a signal that is transferred from the equatorial domain to the apical domain.

We find that the intermediate domain is bifunctional; it can assist either the equatorial domain (as in the *cis* ring) or the apical domain (as in the *trans* ring). The existence of two hinge-bending sites on either side of the intermediate domain is a requirement for this dual function.

In the *cis* subunit, domains I and E form a unified block (Figure 8). Their coupling is induced by nucleotide binding near hinge 2, which hinders the conformational flexibility at the connection between I and E, and stabilizes a relatively open conformation at the E–I interface. The only operative hinge bending site is hinge 1, which triggers the bending of the A domains. In the absence of ATP (the *trans* ring), on the other hand, domain I is decoupled from E. It can sustain a wide range of conformational changes relative to E, and moves in concert with domain A. These periodic changes in domain couplings, from I–E to I–A coupling and vice versa, presumably play an important role in transmitting the allosteric effects induced by nucleotide binding during the chaperonin cycle.

The most extraordinary behavior of the GroEL–GroES chaperone system is its wide range of fluctuations, or deformability of the entire structure, near the highly stable native structure. These motions can contribute to deforming or partially unfolding substructures in bound proteins. The interaction between the substrate protein and the GroEL molecule is usually assisted by the changes in the volume and hydrophobicity of the cavity encapsulating the protein. It is likely that the protein is stretched, or pulled apart, by the uniaxial extension of the chaperonin as revealed in mode 10, or the radial expansion of the apical domains of the *cis* rings as driven by modes 1, 8, and 9. The unfoldase activity of GroEL is required to give a misfolded protein a new opportunity to fold productively. The expansion of GroEL as it binds to protein substrates can also be a way to accommodate variable size substrates. Falke et al. (45, 46) obtained EM images that indicate that GroEL opens upon binding of a large protein. Another possibility is suggested by the behavior of the helicases, where only one strand of the DNA double helix is taken into the central cavity and the other remains on the outside of the protein (47); it is conceivable that only part of a protein is inserted into the central cavity. Perhaps proteins are taken, in terms of the domain, into the interior of the chaperone complex.

While this study has uncovered a broad range of these motions, it is not possible to correlate these with specific unfolding or processing mechanisms of the chaperonins. Further study might uncover some systematic characteristics that would permit constructing a more detailed mechanism of action. Simulations including substrate protein are required for this purpose. A recent experimental study suggests that breaking hydrogen bonds between the encapsulated protein and the cavity can be more important than forced mechanical unfolding (51). One might argue that the presence of a protein of only a few hundreds of residue could not significantly alter the accessible mechanisms of collective motions of the ~8000-residue complex. Yet given that the binding of a nucleotide can exert an allosteric effect in the interdomain cooperativity, which is made possible when the binding site affects a global hinge-bending site, it may be important to

unravel the possible changes in the dynamics of the system, when the cavity of GroEL is occupied with a protein.

ACKNOWLEDGMENT

We are thankful to the staff of the Advanced Biomedical Computing Center (ABCC), FCRDC, for access to their computer facilities. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

REFERENCES

- Horovitz, A. (1998) *Curr. Opin. Struct. Biol.* 8, 93–100.
- Sigler, P. B., Xu, Z. H., Rye, H. S., Burston, S. G., Fenton, W. A., and Horwich, A. L. (1998) *Annu. Rev. Biochem.* 67, 581–608.
- Thirumalai, D., and Lorimer, G. H. (2001) *Annu. Rev. Biophys. Biomol. Struct.* 30, 245–269.
- Corrales, F. J., and Fersht, A. R. (1996) *Proc. Natl. Acad. Sci. U.S.A.* 93, 4509–4512.
- Clarke, A. R. (1996) *Curr. Opin. Struct. Biol.* 6, 43–50.
- Weissman, J. S., Rye, H. S., Fenton, W. A., Beechem, J. M., and Horwich, A. L. (1996) *Cell* 84, 481–490.
- Xu, Z. H., Horwich, A. L., and Sigler, P. B. (1997) *Nature* 388, 741–750.
- Braig, K., Otwinowski, Z., Hedge, R., Boisvert, D. C., Joachimiak, A., Horwich, A. L., and Sigler, P. B. (1994) *Nature* 371, 578–586.
- Kawata, Y., Kawagoe, M., Hongo, K., Miyazaki, T., Higurashi, T., Mizobata, T., and Nagai, J. (1999) *Biochemistry* 38, 15731–15740.
- Buckle, A. M., Zahn, R., and Fersht, A. R. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94, 3571–3575.
- Wang, Z. L., Feng, H. P., Landry, S. J., Maxwell, J., and Gierasch, L. M. (1999) *Biochemistry* 38, 12537–12546.
- Wang, J. D., Michelitsch, M. D., and Weissman, J. S. (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95, 12163–12168.
- Weber, F., Keppel, F., Georgopoulos, C., Hayer-Hartl, M. K., and Hartl, F. U. (1998) *Nat. Struct. Biol.* 5, 977–985.
- Betancourt, M. R., and Thirumalai, D. (1999) *J. Mol. Biol.* 287, 627–644.
- Ranson, N. A., White, H. E., and Saibil, H. R. (1998) *Biochem. J.* 333, 233–242.
- Rye, H. S., Burston, S. G., Fenton, W. A., Beechem, J. M., Xu, Z. H., Sigler, P. B., and Horwich, A. L. (1997) *Nature* 388, 792–798.
- Perrett, S., Zahn, R., Stenberg, G., and Fersht, A. R. (1997) *J. Mol. Biol.* 269, 892–901.
- Rye, H. S., Roseman, A. M., Chen, S., Furtak, K., Fenton, W. A., Saibil, H. R., and Horwich, A. L. (1999) *Cell* 97, 325–338.
- Shtilerman, M., Lorimer, G. H., and Englander, S. W. (1999) *Science* 284, 822–825.
- Sigler, P. B., and Horwich, A. L. (1995) *Philos. Trans. R. Soc. London, Ser. B* 348, 113–119.
- Weissman, J. S., Kashi, Y., Fenton, W. A., and Horwich, A. L. (1994) *Cell* 78, 693–702.
- de Groot, B. L., Vriend, G., and Berendsen, H. J. C. (1999) *J. Mol. Biol.* 286, 1241–1249.
- Ma, J., and Karplus, M. (1999) *Biophys. J.* 76, A118.
- Ma, J., Sigler, P. B., Xu, Z., and Karplus, M. (2000) *J. Mol. Biol.* 302, 303–313.
- Ma, J. P., and Karplus, M. (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95, 8502–8507.
- Bahar, I., Atilgan, A. R., and Erman, B. (1997) *Folding Des.* 2, 173–181.
- Haliloglu, T., Bahar, I., and Erman, B. (1997) *Phys. Rev. Lett.* 79, 3090–3093.
- Bahar, I., Erman, B., Jernigan, R. L., Atilgan, A. R., and Covell, D. G. (1999) *J. Mol. Biol.* 285, 1023–1037.

29. Keskin, O., Jernigan, R. L., and Bahar, I. (2000) *Biophys. J.* 78, 2093–2106.
30. Amadei, A., Linssen, A. B., and Berendsen, H. J. (1993) *Proteins* 17, 412–425.
31. Berendsen, H. J. C., and Hayward, S. (2000) *Curr. Opin. Struct. Biol.* 10, 165–169.
32. de Groot, B. L., Hayward, S., van Aalten, D. M., Amadei, A., and Berendsen, H. J. (1998) *Proteins* 31, 116–127.
33. Hinsen, K. (1998) *Proteins* 33, 417–429.
34. Kitao, A., and Go, N. (1999) *Curr. Opin. Struct. Biol.* 9, 164–169.
35. Hinsen, K., Thomas, A., and Field, M. J. (1999) *Proteins* 34, 369–382.
36. Case, D. A. (1994) *Curr. Opin. Struct. Biol.* 4, 285–290.
37. Tama, F., and Sanejouand, Y. H. (2001) *Protein Eng.* 14, 1–6.
38. Atilgan, A. R., Durell, S. R., Jernigan, R. L., Demirel, M. C., Keskin, O., and Bahar, I. (2001) *Biophys. J.* 80, 505–515.
39. Doruker, P., Atilgan, A. R., and Bahar, I. (2000) *Proteins* 40, 512–524.
40. Flory, P. J. (1976) *Proc. R. Soc. London, Ser. A* 351, 351–380.
41. Wriggers, W., and Schulten, K. (1997) *Proteins* 29, 1–14.
42. Hayward, S., and Berendsen, H. J. C. (1998) *Proteins* 30, 144–154.
43. Chen, J., and Smith, D. L. (2000) *Biochemistry* 39, 4250–4258.
44. Tanaka, N., and Fersht, A. R. (1999) *J. Mol. Biol.* 292, 173–180.
45. Falke, S., Fisher, M. T., and Gogol, E. P. (1999) *Biophys. J.* 76, A23.
46. Falke, S., Fisher, M. T., and Gogol, E. P. (2001) *J. Mol. Biol.* 308, 569–577.
47. Marians, K. J. (2000) *Structure* 8, R227–R235.
48. Chan, H. S., and Dill, K. A. (1996) *Proteins* 24, 345–351.
49. Brocchieri, L., and Karlin, S. (2000) *Protein Sci.* 9, 476–486.
50. Fenton, W. A., Kashi, Y., Furtak, K., and Horwich, A. L. (1994) *Nature* 371, 614–619.
51. Chen, J., Walter, S., Horwich, A. L., and Smith, D. L. (2001) *Nat. Struct. Biol.* 8, 721–728.

BI011393X