

ABSTRACT

Title of Dissertation: STRUCTURAL BIOLOGY OF GROEL
ASSISTED PROTEIN FOLDING

Xue Fei, Doctor of Philosophy, 2014

Dissertation Directed By: Professor George H. Lorimer
Department of Chemistry and Biochemistry
Biophysics Graduate Program

GroEL/ES is the classical example of molecular chaperone that assists the re-folding of many misfolded proteins (SP). Recent kinetic analyses revealed a new paradigm of how GroEL/ES uses ATP to assist protein folding. Following these pioneering biochemical studies, I address two fundamental questions related to GroEL-assisted protein folding using structural biology methods. First, how does GroEL capture SP and how does SP change the kinetics of ADP release? Second, how does GroEL/ES encapsulate SP and control the duration of SP encapsulation?

Chapter 1 summarizes the ATPase cycle of GroEL revealed by systematic biochemical studies, and identifies knowledge gaps in the GroEL-assisted protein folding.

Chapter 2 describes general methods of protein purification and computational approaches, used to analyze conformational differences between two GroEL structures.

Chapter 3 and 4 are focused on the capturing of substrate protein by GroEL.

Crystal structures of GroEL^{D83AR197A}-ADP₁₄ and GroEL^{D83AR197A} show for the first time, ADP binding breaks seven-fold symmetry in the apical and intermediate domains. Such asymmetry provides the structural basis for GroEL to capture heterogeneous SPs and for SP to regulate the release of ADP.

In chapter 5, I described how GroEL/ES encapsulates substrate protein. Two crystal structures of the predominate SP encapsulation complexes: GroEL-GroES₂ “football” complex were reported. One of the complexes is SP free and the other encapsulates two Rubisco molecules simultaneously. From the conformational rearrangement of the inter-ring interface, we proposed “football” complex transmits ATP asymmetry between the rings through an electrostatic interaction between K105 and A109.

Chapter 6 summarized the new knowledge gained by determining these four crystal structures. This chapter ends with a discussion on how chaperonin machine like GroEL promotes the correct folding of various proteins.

STRUCTURAL BIOLOGY OF GROEL ASSISTED PROTEIN FOLDING

By

Xue Fei

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2014

Advisory Committee:
Professor George H. Lorimer, Chair
Professor Dorothy Beckett
Associate Professor Edward Eisenstein
Assistant Professor Nicole A. LaRonde
Professor Devarajan Thirumalai

© Copyright by
Xue Fei
2014

Acknowledgements

First I would like to thank my advisor Dr. George Lorimer, for his patience, encouragement and excellent mentorship. I remembered as if it was yesterday, the first time I walked in to his office and he welcomed me with his characteristic, delightful voice. It is my honor to have worked with you.

I would also like to thank members of the Lorimer lab: Xiang Ye, Dong Yang and Nicolas Corsepius for their help and support. It is wonderful to have such excellent co-workers in the same room who tackled the same problems from different perspectives.

Special thanks to all experts in crystallography I collaborated with, especially Dr. Nicole LaRonde and her lab members. Dr. LaRonde taught me everything about crystallography and generously allowed me to share her instruments and reagents. Even though I am not her student, her door was always open for me and the most important thing I learned from her is a crystallographer's optimism.

Thanks to all members of my dissertation committee. Dr. Eisenstein for serving as the chair of my committee. Dr. Thirumalai for recruiting me into the biophysics program and giving me the opportunity to present my work in his group meeting. Dr. Beckett for teaching me ITC and giving constructive criticism on my two papers.

Thanks to Dr. Wah Chiu, Dr. Steve Ludtke, Dr. Qinfen Zhang and Soung-Hun Roh from Baylor School of Medicine, who put a considerable amount of time in a collaborative project on the cryo-EM structures of the “football” complexes.

Thanks to the Ann. Wylie fellowships, for providing support that allowed me to focus on finishing the last part of my dissertation work.

Finally I would like to thank husband Zhechun Zhang and my parents for their support and encouragement. Being a computational biologist himself, Zhechun has influenced me so much since I was a teenager. I cannot imagine what my life would be like if we never met. I appreciate the understanding from mom and dad, who let their only child travel thousands of miles away to pursue her dream. Any good quality I have must come from them.

Table of Contents

Acknowledgements	ii
List of Tables	vii
List of Figures	viii
Chapter 1: Introduction and Specific Aims	1
1.1 The overall structure of GroEL/GroES	2
1.2 GroEL can be either a two-stroke motor or a parallel processor.	4
1.3 GroEL assists SP folding in steps	6
Capture	7
Unfolding	7
Release	8
1.4 Conformational changes that accompany the turnover of ATP	9
1.5 Specific Aims	12
2.1 Purification of GroEL and GroES	15
2.2 Quantitative structural comparison using a cylindrical coordinate system	15
3.1 Introduction	18
3.2 Methods specific to chapter 3	21
3.2.1 Crystallization	21
3.2.2 Data collection and structural determination	23
3.2.3 Structure details	23
3.3 Results and Discussion	24
3.3.1 Asymmetry in the apical domains of the R-ADP is revealed	24
3.3.2 Crystal structure of R-ADP closely resembles the cryo-EM structure of R-ATP	27

3.3.3 The T to R transition is accompanied by the loss of inter-subunit stabilizing contacts as the origin of asymmetry.	30
3.3.4 R-ADP is distinct from other crystal structures and cryo-EM models of GroEL....	32
3.3.5 Rigid body rotation of the intermediate domain controls nucleotide release.....	37
3.3.6 Rigid body rotation of apical domain enables selective binding to GroES or substrate protein	41
3.4 Summary	45
Chapter 4: The crystal structure of GroEL^{D83A/R197A}	47
4.1 Introduction.....	47
4.2 Methods specific to chapter 4	47
4.2.1 Crystallization.....	47
4.2.2 Data collection and structural determination.....	48
4.3 Results and Discussion.....	48
4.3.1 apo-GroEL ^{D83A/R197A} is in the T state.....	48
4.3.2 A systematic analysis of all inter-subunit salt bridges.....	53
4.3 Summary	56
Chapter 5 Crystal structures of symmetric “football” complexes: [GroEL-(ADPBeF₃)₇-GroES]₂.....	57
5.1 Introduction.....	57
5.2 Methods specific to chapter 5	58
5.2.1 Crystallization.....	58
5.2.2 Data collection and structure determination.....	59
5.2.3 Interface analysis.....	60
5.3 Results and Discussion.....	60

5.3.1 Overall structures of the “football” complexes, with or without SP.....	60
5.3.2 Structural Plasticity of the Football Complexes.....	64
5.3.4 Structural Basis for Inter-Ring Communication.	73
5.4 Summary	74
Chapter 6: Summary and final discussion	76
6.1 What happens during SP capture?	76
6.2 What happens during SP encapsulation?.....	77
6.3 Final discussion	78
Reference	80

List of Tables

Table. 3-1 Data collection and refinement statistics of GroEL^{D83AR197A}-ADP₁₄.	21
Table. 3-2 Possible inter-subunits salt bridges in the R-ADP and T (PDB ID: 1XCK) states.	32
Table. 3-3 Domain averaged b factors for the T(PDB ID: 1XCK), R-ADP and R”(PDB ID: 1AON) states.....	32
Table. 4-1 Data collection and refinement statistics of GroEL^{D83A/R197A}	50
Table. 5-1 Data collection and refinement statistics of MT-football and SP-football.	62

List of Figures

Fig. 1-2 The overall structure of GroES.	4
Fig. 1-3 GroEL's catalytic cycle in the absence of SP.....	5
Fig. 1-4 GroEL's catalytic cycle in the presence of SP.	6
Fig. 1-5 The Overall structures of GroEL-GroES1 "bullet".	11
Fig. 2-1 The cylindrical coordinate system for quantitative structural comparison..	16
Fig. 3-1 Salt bridges perturbed by mutations are ordinarily broken during the T to R transition.	19
Fig. 3-2 Overall structure of the R-ADP and its asymmetry.	20
Fig. 3-3 Crystal of GroEL^{D83AR197A}-ADP₁₄, grown in the presence of MPD.....	22
Fig. 3-4 Quantitative analysis of the R-ADP using cylindrical coordinate system. ..	25
Fig. 3-5 Comparison between symmetry imposed R-ADP to R-ATP structures determined by cryoEM.....	29
Fig. 3-6 Detailed view of model-density fitting in asymmetrical R-ADP and symmetrical R-ATP.	29
Fig. 3-7 Numbers of inter-subunit contacts between two neighboring subunits, in the asymmetric R-ADP state and symmetric T state.	31
Fig. 3-8 Conformation differences between the R-ADP and T or R" state, analyzed using difference "signature" plots.	33

Fig. 3-9 Contact between nucleotides and the nucleotide binding residues in the R-ADP (a, yellow) and GroEL-ATPγS₁₄ (b, grey PDB ID: 1KP8).....	35
Fig. 3-10 Conformation differences between GroEL-ATPγS₁₄ and the T or R-ADP state.....	36
Fig. 3-11 Conformational changes in the intermediate domains, between the T (PDB ID: 1XCK), R-ADP and R” (PDB ID: 1AON) states.....	38
Fig. 3-12 Conformational changes in apical domains, between the T (PDB ID: 1XCK), R-ADP and R” (PDB ID: 1AON) states.....	42
Fig. 4-1 Overall structure of GroEL^{D83A/R197A}.	49
Fig. 4-2 Nucleotide-free GroEL^{D83A/R197A} is in the T state.	51
Fig. 4-3 Analysis of all inter-subunit salt bridges in the T, R, and R” state.....	55
Fig. 5-1 Crystal structures of the “MT-football” complex and the “SP-football” complex with encapsulated Rubisco.....	61
Fig. 5-2 Packing of the “football” complex in the crystal.....	62
Fig. 5-3 Quantitative analysis of the conformational differences between “MT-football” and “SP-football”.....	63
Fig. 5-4 Asymmetry in the “football” complexes.	65
Fig. 5-5 Solvent-exposed residues inside the central chamber are more asymmetric than those outside the chamber.	67
Fig. 5-6 Averaged B factors of the Cα of each residue in the “MT-football”(green) and “SP-football”(orange) complexes.....	67

Fig. 5-7 GroEL-GroES interfaces in the “football” complexes.	68
Fig. 5-8 Switching of electrostatic interactions at the inter-ring interface.	70
Fig. 5-9 Quantitative analysis of the conformational changes at the inter-ring interface during the “football” to the “bullet” transition.	72

Chapter 1: Introduction and Specific Aims

Each protein is a one-dimensional polypeptide chain after translation from mRNA [1]. In order for a protein to carry out its designated function, it must adopt its characteristic three-dimensional structure [2]. This process is called protein folding and the resulting functionally characteristic structure is called the native state. A protein's native structure is coded in its sequence, so in an ideal world, any protein should fold spontaneously [2]. But in reality some proteins fail to reach their native states under physiological conditions and may become "mis-folded".

Misfolded proteins may be dangerous not only because they lack their designated function but also because they tend to aggregate [3, 4]. For example fibrous protein aggregates called amyloids have been associated with more than 20 serious human diseases, including most neurodegenerative disorders [5, 6]. Currently 7 million Americans suffer from such neurodegenerative conditions including Alzheimer's, Parkinson's and Huntington's diseases [7, 8]. Protein aggregates can be toxic for *E. coli*. The formation of insoluble inclusion bodies in recombinant organisms expressing foreign proteins is another manifestation of the aggregation problem [9].

Protein aggregates can be prevented by the chaperonin machine GroEL/ES [9]. GroEL interacts with a large variety of misfolded proteins and together with GroES and ATP, help them to reach their native states [10].

1.1 The overall structure of GroEL/GroES

The first crystal structure of GroEL was reported in 1994 [11]. This structure of nucleotide-free GroEL (apo-GroEL) showed it consists of fourteen identical subunits, arranged into two heptametrical rings. These two rings are stacked back to back and there is an open chamber in each ring (Fig. 1-1). Although the two chambers appear to be connected, C terminal tails invisible to crystallography but not to cryo-EM (residue 526-548) may plug the hole between rings [12].

Each GroEL subunit has a molecular weight of about 57KD and contains three domains: the equatorial domain, the intermediate domain and the apical domain (Fig. 1-1) [11]. The equatorial domain is the most rigid domain. It mediates inter-GroEL ring communications and contains the nucleotide-binding pocket that accommodates one ADP/ATP-Mg²⁺-K⁺. The intermediate domain covers the nucleotide-binding pocket like a “lid” and contains a residue crucial for ATP hydrolysis, D398. The apical domain is the most flexible domain. It binds both GroES and misfolded proteins (SP), with the same hydrophobic groove between helices H and I [13].

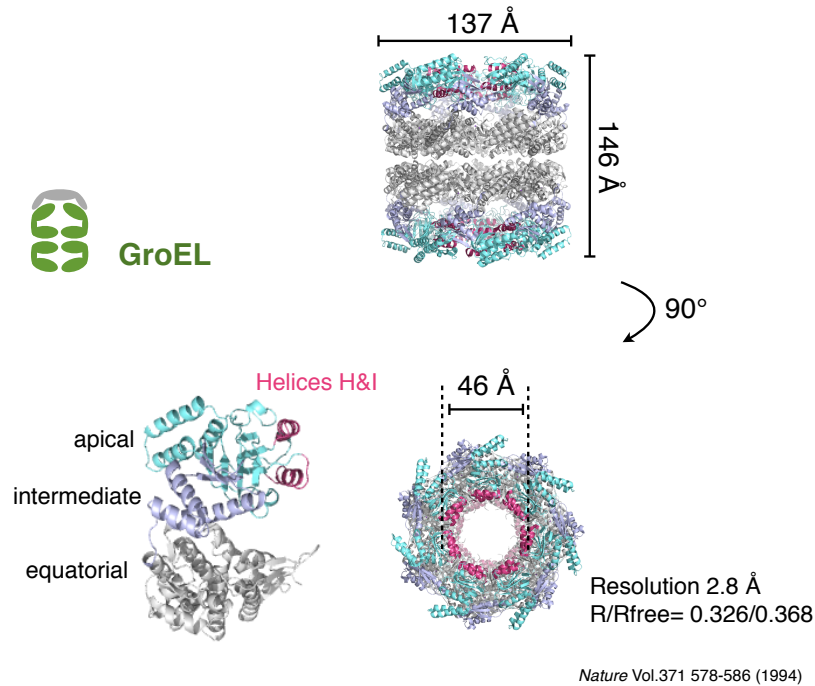


Fig. 1-1 The overall structure of GroEL.

The equatorial, intermediate and apical domains are colored in grey, light blue and cyan respectively. Helices H and I are colored in red. (PDB ID: 1OEL, 1GRL, 1XCK)

GroES contains seven identical subunits, approximately 10kD each [14]. The seven GroES subunits arrange into a dome shaped “lid” (Fig. 1-2). GroES interacts with helices H and I of GroEL through seven so-called “mobile-loops”, each containing 28 residues (residue 16-33). Before GroEL binds, these loops are unstructured.

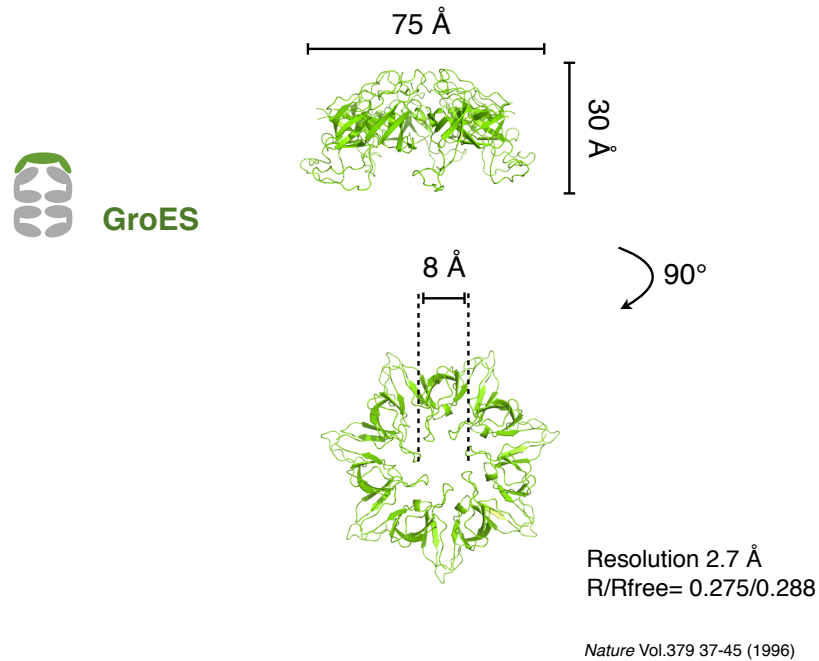


Fig. 1-2 The overall structure of GroES.

1.2 GroEL can be either a two-stroke motor or a parallel processor.

GroEL/ES assisted protein folding depends on the turnover of ATP [15]. The classical view of this ATP-driven machine is a two-stroke motor [16, 17]. In the two-stroke motor model: (1) the two GroEL rings are always 180 degree out of phase of one another so that they hydrolyze ATP and bind GroES alternatively and (2) the “bullet”-shaped GroEL-GroES1 is the protein folding functional form (Fig. 1-3).

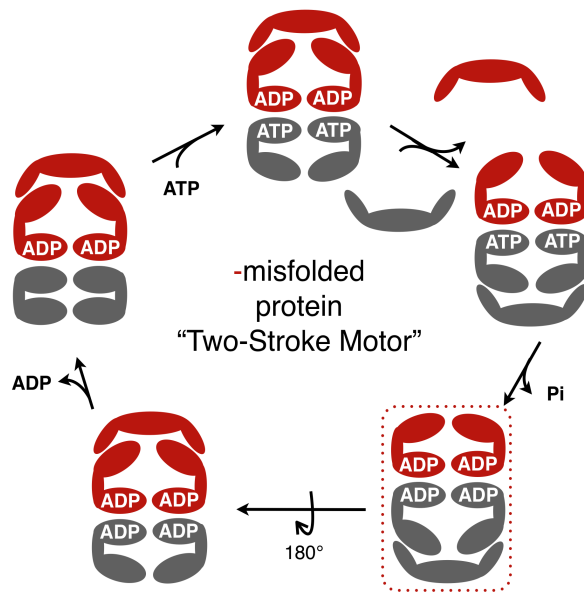


Fig. 1-3 GroEL's catalytic cycle in the absence of SP.

According to the two stroke motor model, two GroEL rings work alternatively. In each half cycle, only one of two rings binds ATP and GroES (the grey ring) while the other ring idles (the red ring).

But the two-stroke motor model is only accurate in the absence of SP. Using multiple spectroscopic probes, our lab recently showed that in the presence of misfolded proteins, GroEL/ES works as a parallel processor [18, 19]. (1) The two GroEL rings are only slightly out-of phase so that they hydrolyze ATP and bind GroES simultaneously and (2) the “football”-shaped GroEL-GroES₂ is the predominate species (Fig. 1-4).

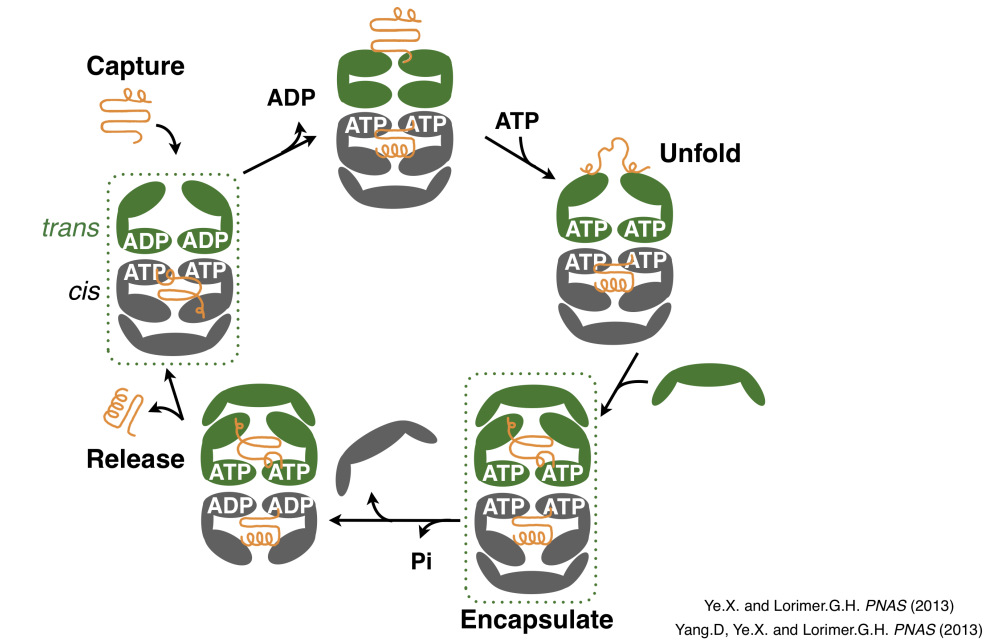


Fig. 1-4 GroEL's catalytic cycle in the presence of SP.

SP changes the kinetics of ATP turnover, so that two GroEL rings bind ATP, GroES and SP simultaneously.

1.3 GroEL assists SP folding in steps

As a parallel processor, each GroEL ring can be treated as a separate functional unit.

A misfolded protein is treated by GroEL ring through multiple steps. These steps include (i) capture, (ii) unfolding and (iii) release (Fig. 1-4).

Capture

GroEL captures a large variety of cytosolic proteins in *E.coli*, when these proteins are in misfolded states [10]. In general GroEL does not interact with the native states of SPs, since the recognition motifs are buried in the native states. Biochemical and structural biology studies both showed GroEL capture SP with hydrophobic interactions. Each hydrophobic groove between apical domain helices H and I is able to capture a SP mimicking hydrophobic polypeptide [13]. These hydrophobic GroEL recognition motifs are usually buried in native states and are only exposed when denatured so that GroEL may distinguish between native proteins and misfolded proteins [20].

In order to be captured efficiently, a SP contains at least 3 hydrophobic GroEL binding motifs on average [21]. To bind GroEL with high affinity, SP has to occupy at least 3 out of seven binding grooves between helices H and I simultaneously [22]. Once captured, SP binds to GroEL tightly, the equilibrium dissociation constant is nanomolar [15, 23]. But how GroEL captures a variety of SP different in size and shape is yet to be determined.

Unfolding

Two kinds of partial SP unfolding have been observed: passive unfolding and forced unfolding. Passive unfolding accompanies SP binding: GroEL prefers to capture SP in extended conformations that are in fast equilibrium with compact

states [24]. Local unfolding may also occur as SP interacts successively with an increasing number of binding sites [25].

Forced unfolding is triggered by ATP induced conformational changes [25, 26]. ATP binding causes GroEL's apical domains to rotate and separate from each other [27, 28]. Because captured SP binds to at least 3 apical domains of different subunits, when the apical domains move apart, SP will be stretched and partially unfolded. Although forced unfolding was observed using different model substrates, including SP mimicking stack-dyes [29], whether GroEL assisted protein folding requires forced unfolding is still controversial [25, 26].

With the addition of GroES following ATP induced unfolding, SP is displaced from its binding sites on apical domains into a closed chamber [25]. It is unclear if this closed GroEL-GroES chamber guides SP folding actively or only passively prevents aggregation like an "Anfinsen's Cage" [30, 31]. Recent structural and biochemical evidence favors the active chamber model slightly [30, 32]. But for some SPs, encapsulation after unfolding might not be necessary; a partially unfolded SP will continue to fold after released.

Release

The duration of encapsulation is controlled by the rate of ATP hydrolysis. SP is encapsulated until sufficient nucleotide asymmetry has developed (i.e. the difference in number of ATP hydrolyzed between two GroEL rings) when the ring with less ATP releases its GroES [18, 33]. Presumably, the nucleotide

asymmetry is sensed by the inter-ring interface and eventually triggers the release of one GroES “lid” and the encapsulated SP. The average half time of encapsulation is 1.5 seconds, after which SP will be released regardless of whether it has folded or not [18].

After each round of capture, unfolding and encapsulation, less than <1% of SPs are released in folded states [18]. Therefore each SP goes through the chaperonin cycle many times before it reaches its native state; this process is referred as iterative annealing [34].

Previous studies have attempted to answer how GroEL assisted protein folding using a single turnover mutant (GroEL^{SR1}), where iterative annealing is prohibited. Once SP is captured by GroEL^{SR1}, it can be unfolded and encapsulated but never released [35]. However the mechanism of SP folding under infinite encapsulation by GroEL^{SR1} very likely differs from the mechanism of SP folding by GroEL^{WT}. For example, in the cavity GroEL^{SR1}/ES, the half time of SP folding is about 10 minutes but under physiological conditions, one SP only remains encapsulated no longer than 2 seconds per turnover [18,19,35].

1.4 Conformational changes that accompany the turnover of ATP

To optimize SP folding in separate steps, each GroEL ring undergoes conformational changes [36, 37]. The conformation of a GroEL ring is coupled to the turnover of ATP and GroES [27, 28, 33, 38, 39]. In the absence of ATP

and GroES, a GroEL ring resides in the **T** (Taut) state. The **T** state has lower affinity for nucleotide and higher affinity for SP [11, 40].

ATP or ADP binding induces a conformational change and drives GroEL from the **T** state to the **R** (Relaxed) state [27, 39]. The **R** state has higher affinity for nucleotide and lower affinity for SP. AFM and crosslinking studies suggested **R-ATP** and **R-ADP** rings are both taller and wider than the **T** state ring [28, 41]. Cryo-EM at 8-9 Å resolution showed nucleotide-induced ring elevation and expansion is mainly caused by the conformational changes of the apical domains [27]. ATP binding causes the apical domain to tilt up to 35 degrees and twist up to 20 degrees around a hinge connecting the intermediate and apical domains. Interestingly a total of six **R-ATP** conformations were observed but due to imposed seven-fold symmetry, subunits of the same synthetic ring are identical.

When both ATP/ADP and GroES bind, the GroEL ring further elevates and expands [33, 38, 42]. The apical domains rotate 100 degrees and as the result the volume of GroEL cavity expands. These conformations are called **R'**(GroEL₇-ATP₇-GroES) state and **R''** (GroEL₇-ADP₇-GroES) state [33, 38] (Fig. 1-5). Crystal structures showed the **R'** state and the **R''** state are almost identical [33, 38, 42], indicating it might not be necessary to hydrolyze all seven ATPs of the **R'** state to proceed in the chaperonin cycle.

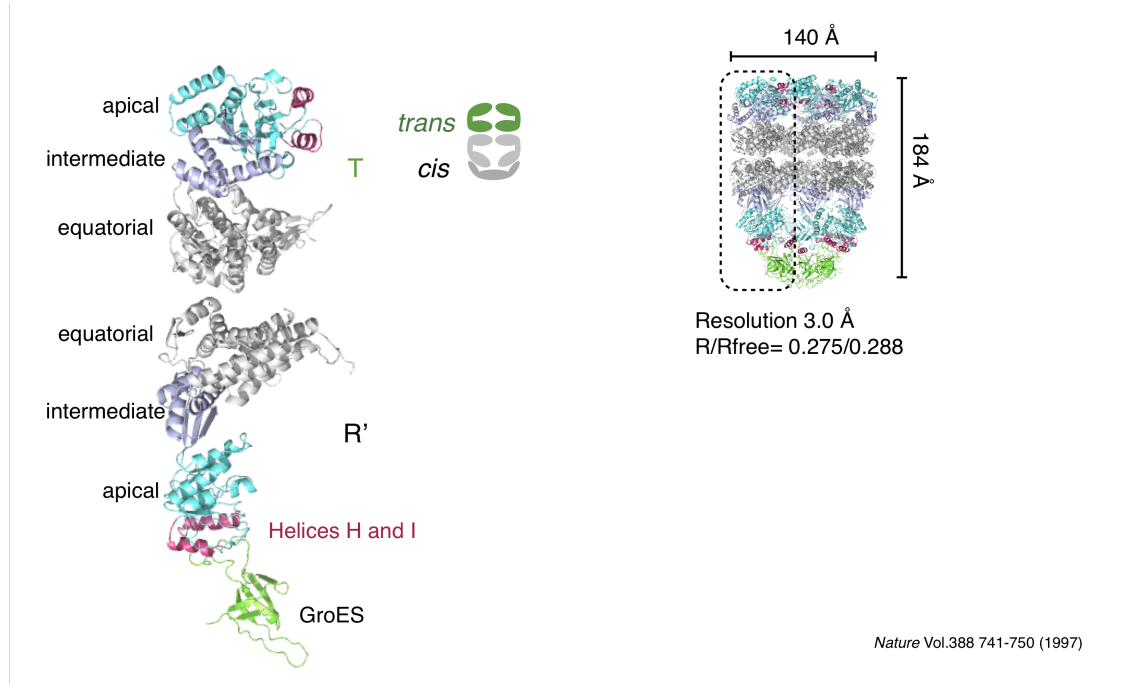


Fig. 1-5 The Overall structures of GroEL-GroES1 “bullet”.

Nucleotide and GroES binding induced dramatic conformational changes so that the cis-ring looks different from the trans-ring (PDB ID: 1AON (R'') and 1PCQ (R')).

Even though many combinations of conformations between rings are possible, only a few turn out to be functionally relevant [38, 42]. High-resolution structures were only available for two complexes involved in the functional cycle: the $[\text{cisGroEL-GroES}_1\text{-ADP}_7]\text{-}[\text{transGroEL}]$ complex and $[\text{cisGroEL-GroES}_1\text{-(ADPAIF}_3)_7]\text{-}[\text{transGroEL}]$ complex. In both structures GroES bound cis rings showed expanded **R'** or **R''** conformation, while the ligand free trans rings in these two structures looked like **T**-like. To distinguish two rings in different

conformations, we name the GroEL bound ring the cis-ring and the GroES free ring the trans-ring.

1.5 Specific Aims

The goal of my work is to gain a better understanding of GroEL assisted protein folding using structural methods, particularly X-ray crystallography. I am interested in two aspects of GroEL assisted protein folding: SP capture and SP encapsulation.

(1) SP capture: Previous work in the lab suggested SP is captured by the ADP bound trans-ring [18, 19]. Once captured, SP accelerates ADP release so the chaperonin machine switches from energy efficient “resting” state in which it operates the energy conserving asymmetric cycle, to the active mode, which turns over ATP much faster [17, 19]. The questions are:

- (i) How does the ADP bound trans-ring captures various, structurally diverse SP?
- (ii) How does SP bound to the apical domain accelerate ADP release from the distant equatorial domains?

To answer these questions, I aimed to determine the structure of the resting state capture complex $[\text{cisGroEL-GroES}_1\text{-ADP}_7]\text{-}[\text{transGroEL-ADP}_7]$. However crystals of this complex never diffracted well enough to pursue structural studies so we made two compromises. First, I concentrated on the **R-ADP** ring responsible for SP capture, assuming GroEL binding to one ring will not alter the other ring's

conformation dramatically. Second I used a GroEL double mutant D83A/R197A, which has a higher affinity for ADP but otherwise behaves very similar to the wild-type [43]. By removing two salt bridges (D83-K327 and E386-R197), which ordinarily break during **T** to **R** transitions, the equilibrium was shifted towards the **R** state. By doing so, we determined the crystal structure of a GroEL^{D83A/R197A}-ADP₁₄ complex, both rings of which are in the **R-ADP** state. This is also the first high-resolution structure of GroEL in the **R** state.

In chapter 3, the structure of **R-ADP** GroEL will be analyzed in detail, highlighting dramatic conformational changes of the apical and intermediate domains. In chapter 4, the crystal structure of nucleotide free GroEL^{D83A/R197A} is reported. As we expected, the conformational changes we observed in double mutant **R-ADP** are not caused by mutations but are caused instead by nucleotide binding.

(2) SP encapsulation: According to the chaperonin dogma, SP is only encapsulated by the cis-ring of the “bullet”-shaped complexes [16, 17]. However recent discoveries by a few other labs suggested that “football”-shaped complex (GroEL-GroES₂) could be crucial in assisting SP folding [44-50]. Ye, Yang and Lorimer demonstrated that due to changes in GroEL’s kinetic mechanism, in the presence of SP the “football” complex become the predominate species [18, 19]. Designated for SP encapsulation, the “football” complex has two chambers and may encapsulate two molecules of SPs simultaneously [18, 19, 33].

To understand what happens during SP encapsulation, I addressed the following questions:

- (i) What are the differences between the “football” complex and the “bullet” complex?
- (ii) How do the cavities of “football” complex interact with encapsulated SP.
- (iii) How is ATP hydrolysis coupled to the release of one of GroES “lids” from the “football” complex?

To address these questions, I determined two crystal structures of the symmetric GroEL:GroES₂ “football” complexes; one is SP free and another contains a substrate Rubisco in each cavity. These two structures are analyzed in chapter 5 and a model of transmitting the signal of ATP hydrolysis through equatorial domains is proposed.

Chapter 2 General methods and experimental procedures

2.1 Purification of GroEL and GroES

GroEL and GroES were over-expressed in *E. coli*. BL21 and purified using an existing protocol [17, 51]. After several steps of purification including anion exchange chromatography, gel filtration and acetone precipitation, we estimated GroEL's purity using tryptophan fluorescence method [17, 51]. If more than 10% of GroEL rings were SP contaminated, the acetone precipitation step was repeated. One preparation from 6L of over-night culture typically yields 120 mg GroEL or 100 mg GroES. Because freezing affects crystallization, purified GroEL and GroES were stored at 4 °C, in the buffer containing 50 mM Tris-HCl pH 7.5, 20 mM MgCl₂, 200 mM KCl and 0.03% (w/v) NaN₃.

2.2 Quantitative structural comparison using a cylindrical coordinate system

Each GroEL contains ~7300 residues and a GroEL-GroES₁ complex contains ~8000 residues. The size of GroEL and GroEL/ES complex make it difficult to identify conformational difference between two structures by visual inspection. Traditional structural alignment method, such as the one implemented in Pymol, rotates and translates one structure respect to another, to minimize the sum of squared distances [52]. Although this method may work well for smaller proteins, it is not ideal for large protein complexes, such as GroEL, which contains multiple domains and subunits.

To quantitatively compare two GroEL structures and capture conformational differences at single residue resolution, we developed a cylindrical coordinate system to analyze the conformation of each GroEL subunit [39]. R is the distance between the $C\alpha$ of a given residue and the 7-fold symmetry axis (defined by the symmetrical equatorial domains). H is the height of the $C\alpha$ of a given residue over the 2-fold plane of symmetry (between two rings). θ is the angle between two vectors; one connects the $C\alpha$ of residue i to the axis of 7-fold symmetry and the other connects the $C\alpha$ of residue i in the neighboring subunit to the axis of 7-fold symmetry (Fig. 2-1). Usually, we plotted R , H and θ as a function of residue number in a subunit. In this way, these three variables together fully depict the conformation of each subunit. Therefore we refer the R , H and θ plots as “signature plots” to describe their cursive nature.

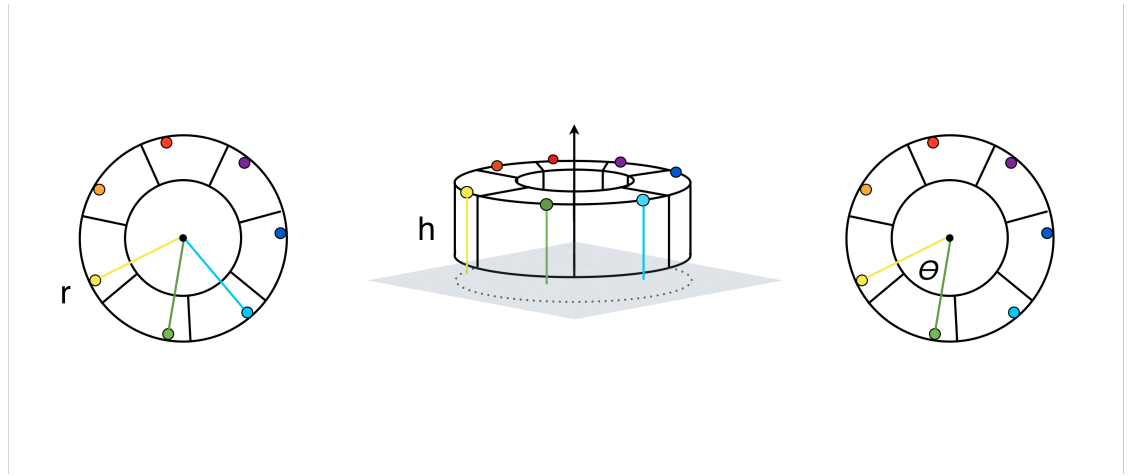


Fig. 2-1 The cylindrical coordinate system for quantitative structural comparison.

Three cylindrical coordinates (R , H , θ) replace cartesian coordinates (x , y , z).

The “signature plots” can be used to quantify the structural difference (i) between subunits in a ring and (ii) between two rings. First, the R , H and θ plots of each individual subunit in one structure capture any variation or asymmetry between subunits. If the seven subunits in one ring show perfect seven-fold symmetry, for a given residue R , H and θ are expected to be the same for all seven subunits within a ring and the “signatures” of R , H and θ of different subunits should overlap (for example, θ plots overlap at $360/7 = 51.43^\circ$). However, if seven subunits in a ring are not identical, the “signature” plots of each subunit will be different. By comparing these “signature” plots, one can easily identify domains that show most structural variation between subunits.

Second, we can also quantify the difference between two GroEL rings, by comparing (R, H, θ) . For example, one can calculate $\Delta R = \langle R' \rangle - \langle R'' \rangle$ and $\Delta H = \langle H' \rangle - \langle H'' \rangle$, where $\langle ' \rangle$ is the average over one ring and $\langle '' \rangle$ is the average over the other. If two GroEL rings have identical averaged conformation, for any residue $\Delta \langle R \rangle$ and $\Delta \langle H \rangle$ are expected to be below a certain threshold. To define the threshold, we calculate $\Delta \langle R \rangle$ and $\Delta \langle H \rangle$ between two T state structures solved independently in different space groups [53, 54].

3.1 Introduction

To assist protein folding the GroEL/GroES chaperonin-machine cycles through a series of conformational states in response to ligand binding [34, 36, 55]. Two states, **T** (taut) and **R** (relaxed) are populated in the absence of GroES. The compact **T** state is favored in absence of nucleotide, while after ADP or ATP binding the equilibrium between the **T** and **R** states shifts to favors the extended **R** state. Another two states, **R'** and **R''**, that exist before and after ATP hydrolysis respectively, are populated upon GroES binding to the **R** state [34, 36, 55]. Crystal structures of the **T**(1OEL) [11], (1XCK) [53], the **R'** (1PCQ) [42] and the **R''** (1AON) [38] states are available (Fig.1-1 and Fig. 1-5).

As described by the theory of nested cooperativity, the conformational transitions within each heptameric ring are positively cooperative, while the transitions between the rings are negatively cooperative [37, 56]. Multiple salt-bridges, some within a subunit, others between subunits, stabilize these various conformational states. In such a dynamic system, these salt-bridges must continually be broken and re-formed at different points during the chaperonin cycle. Two salt-bridges, an intra-subunit one D83-K327 and an inter-subunit one R197-E386, stabilize the **T** state are ordinarily broken during the **T** to **R** transition that follows the binding of ATP to the wild type (Fig. 3-1). Individually, the R197A mutation destabilizes the **T** state [57, 58] while replacing the D83-K327 salt-bridge with a disulfide bond locks the ensemble in

the T state [40, 59]. Here, two mutations (D83A, R197A) were introduced to remove these salt bridges (D83-K327 and R197-E386) so that the T state was destabilized. Steady state and pre-steady-state analyses of GroEL^{D83A/R197A} reveals that it populates a T-like state in the absence of nucleotide but more readily transitions to the R-state upon nucleotide binding than the wild-type GroEL [39, 43]. We also report the crystal structure at 2.7Å resolution of this GroEL^{D83A/R197A} containing bound ADP in the R conformational state (Fig. 3-2a and Table. 3-1). The crystal structure of the R state is distinct from previously reported crystal structures [11, 38]. It is similar in some respects to the cryo-EM (PDB ID 4AAS, etc) structures of GroEL-ATP [27]. However, the structure reported here differs by being strikingly asymmetric in the apical domain.

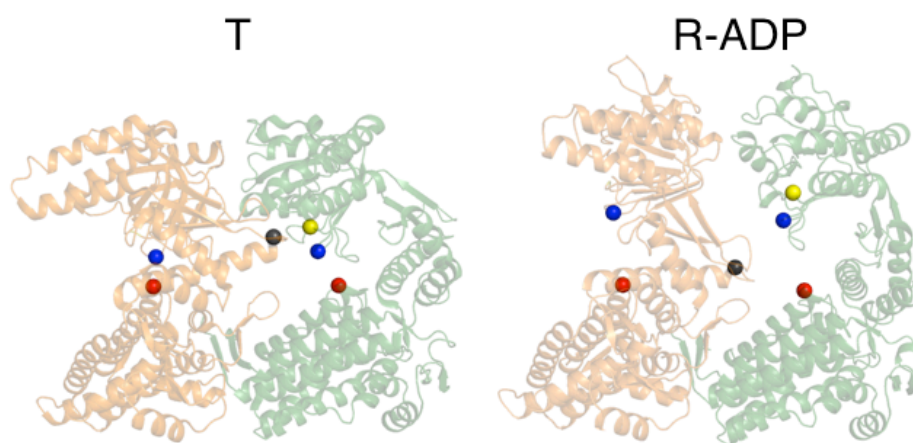


Fig. 3-1 Salt bridges perturbed by mutations are ordinarily broken during the T to R transition.

The positions of the inter-domain, intra-subunit D83-K327 (red and blue spheres) and inter-domain, inter-subunit R197-E386 (black and yellow spheres) salt-bridges before and after the T to R allosteric transition.

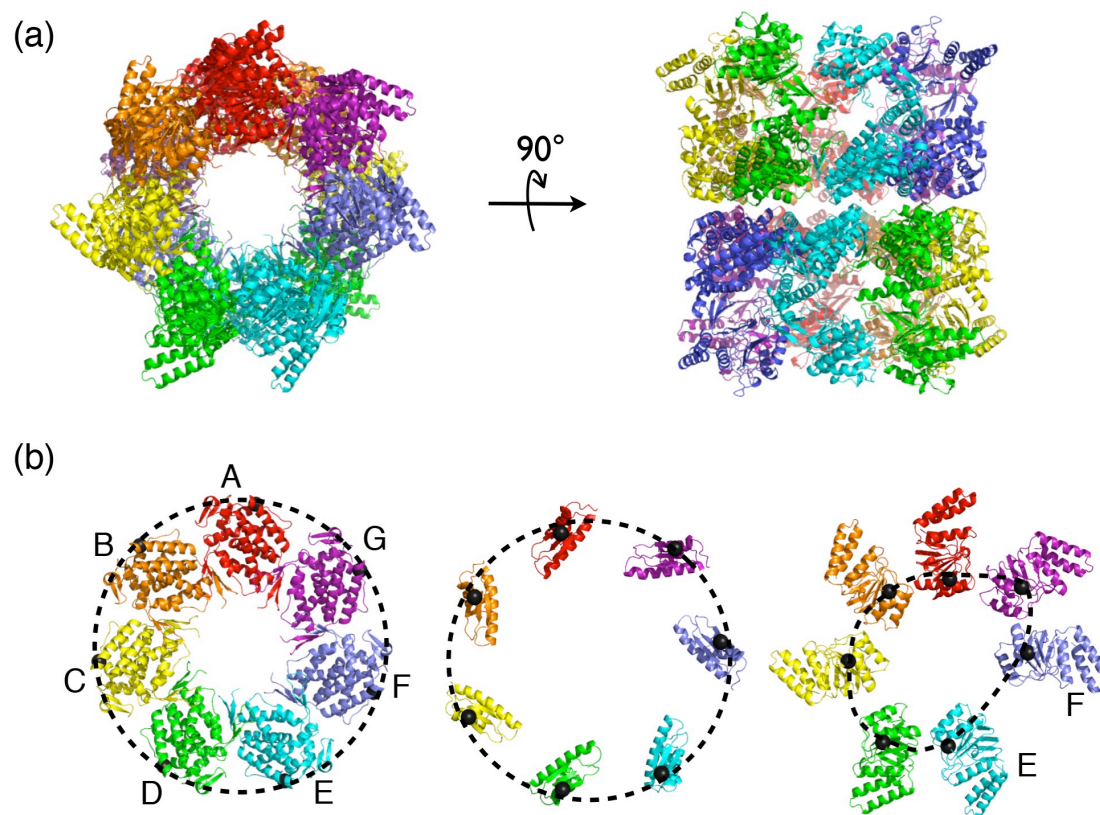


Fig. 3-2 Overall structure of the R-ADP and its asymmetry.

(a) Overall structure of the **R-ADP** state viewed from the top (left) and the side (right). Seven subunits in one ring are colored in red, orange, yellow, green, cyan, blue, and purple, respectively. (b) Arrangement of equatorial (left), intermediate (middle) and apical (right) domains in the ring. E130 (in equatorial domain), V190 (in intermediate domain) and E310 (in apical domain) are connected (C α s as black spheres and black dash lines) to highlight the degree of asymmetry in different domains.

Space group	C121	
Unit cell a, b, c (Å)	235.222/141.655/156.693	
Resolution range (Å)	46.17-2.722(2.819-2.722)	
Rmerge (%)	5.2(41.1)	
Total number of reflections	120262(10710)	
Redundancy	3.1(3.1)	
Completeness (%)	95.62(85.28)	
I/σ (I)	13.98(3.13)	
Number of residues/asymmetry unit		3668
Number of protein atoms		27012
Number of ligand atoms		368
Number of water atoms		597
R (%)	16.63 (23.10)	
Rfree (%)	20.30 (28.34)	
Test set size (%), selection	5, radom	
RMSD		
bond lengths (Å)		0.005
bond angles (degree)		0.95
Ramachandran plot		
Most favored (%)		97.1
allowed(%)		2.71
outlier(%)		0.11
Average B factors (Å²)		
Protein		88.2
Ligand		81.8
Water		63.2

Table. 3-1 Data collection and refinement statistics of GroEL^{D83A/R197A}-

ADP₁₄.

3.2 Methods specific to chapter 3

3.2.1 Crystallization

The **R-ADP** state GroEL was made by mixing 190 μM GroEL^{D83A/R197A} and 5 mM ATP in 50:20:200 buffer (50 mM Tris-Acetate pH 7.5, 20 mM MgCl₂, 200 mM KCl). GroEL and ATP mixture was incubated in room temperature until ATP was all converted to ADP. The **R-ADP** samples were screened for crystallization condition by using sitting drop vapor diffusion method (drop: 0.5

μL well solution: 100 μL). Diffraction quality crystals were grown at the screen condition with well solution consists of 34% MPD (v/v), 0.1 M acetic acid pH 5.5 and 20 mM CaCl_2 . Crystals were optimized by hanging-drop vapor diffusion method (drop: 6 μL well solution: 1 mL). Cubic shaped crystals began to show up after 3-4 days in drops containing 2:1 mixture of protein sample and well solution (Fig. 3-3). One week after the crystallization trays were set up, single crystals were removed from mother liquor and soaked in well solution for 5s before mounted and then flash-frozen by liquid nitrogen.

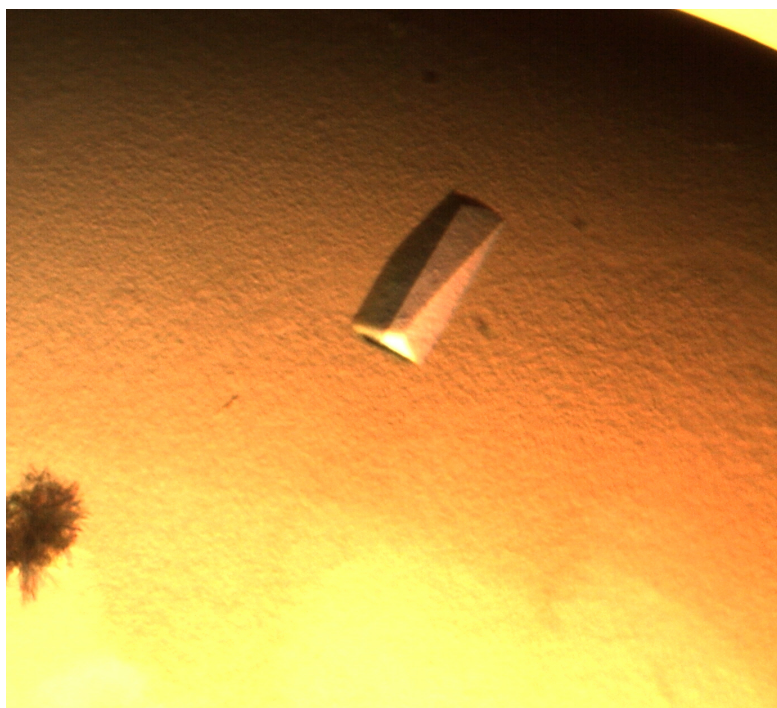


Fig. 3-3 Crystal of GroEL^{D83AR197A}-ADP₁₄, grown in the presence of MPD.

3.2.2 Data collection and structural determination

Diffraction data were collected at the NE-CAT beamline 24-ID-C located at the Advanced Photon Source, Argonne National Laboratory (Argonne, IL, USA). 300 frames with 0.5° oscillation were collected at 100K. The data were indexed and integrated using HKL2000. The structure of the **R-ADP** was solved by segmented molecular replacement using apo GroEL(PDB ID: 1XCK) as the search model. Three domains in the apo GroEL were dissected and fitted sequentially. The first search ensemble consisted of 7 equatorial domains and 7 intermediate domains. Then search ensembles 2 to 8 were add one at a time, each ensemble contains one apical domain. In structure refinement, NCS (non crystallographic symmetry) was turned off and each subunit was refined independently. Molecular replacement and structure refinement were carried out using AutoMR, Refine and Coot in the Phenix suites [60]. The coordinates after refinement were deposited in the PDB as 4KI8.

3.2.3 Structure details

Two heptametrical rings in the **R-ADP** are related by crystallographic symmetry and the asymmetrical unit contains one of two rings. Electron density is weak from residue 221 to residue 226 in chain C and chain G due to high flexibility in those regions. Data collection and refinement statistics are show in Table 1. The figures that depict structures of the **R-ADP** were prepared using Pymol. The analysis of R , H and θ were plotted using Prism.

3.3 Results and Discussion

3.3.1 Asymmetry in the apical domains of the R-ADP is revealed

In contrast to previously solved GroEL structures in other conformational states, the **R-ADP** shows a striking asymmetry. While the equatorial domains and most the intermediate domains are almost perfectly symmetrical, the apical domains break the 7-fold symmetry (Fig. 3-2b). A huge gap (~ 1 nm, Fig. 3-2b right) exists between the apical domains of the chain E and chain F.

To quantify such asymmetry, we used the cylindrical coordinate system to analyze the conformation of each GroEL subunit, as described in chapter 2 (Fig. 2-1). Three cylindrical coordinates R , H and θ together fully depict the unique conformation of each subunit, including possible asymmetry in three different directions: radial (R), vertical (H) and rotational (θ). If the seven subunits show perfectly 7-fold symmetry, for a given residue R , H and θ are expected to be the same for all seven subunits within a ring and the “signature” plots of R , H and θ of different subunits should overlap (θ plots overlap at 51.43°).

Our quantitative analysis confirms the asymmetry, showing apparent heterogeneity in R , H and θ (Fig. 3-4). This heterogeneity or asymmetry is confined to the intermediate and apical domains. For a given residue R can differ by up to 20 Å and H by up to 10 Å (indicated by arrows in Fig. 3-4). For a given pair of residues, θ deviates from 51.43° (as in perfect 7-fold symmetry) between $+30^\circ$ (between chain E and chain F) and -20° (between chain D and chain E).

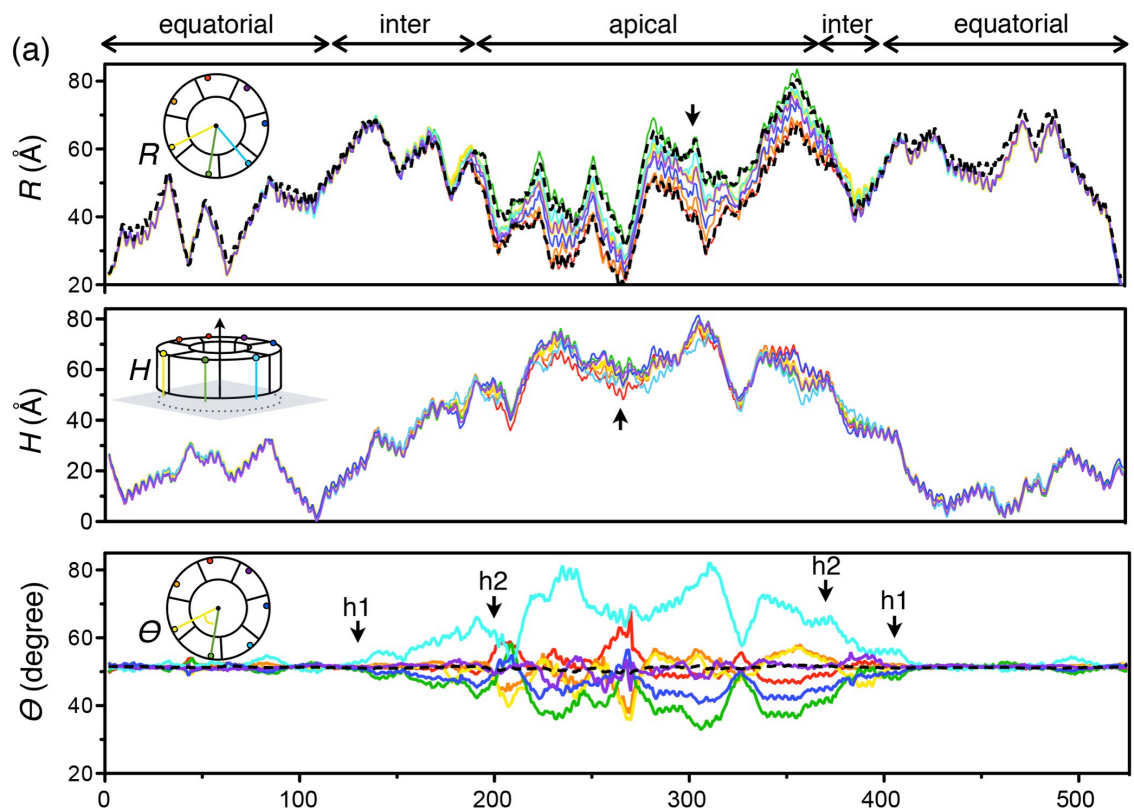


Fig. 3-4 Quantitative analysis of the R-ADP using cylindrical coordinate system.

Asymmetry between the subunits in the **R-ADP** is dissected (inserts in upper and lower panels) and quantified using R , H and θ as described in chapter 2. The colors of traces (from red to purple) reflect the arrangement of subunit and are consistent with that in Figure 3-2(b). Two black arrows in R and H panels indicate the largest degree of asymmetry in radial (R) or vertical (H) direction. Four black arrows in θ panel marked with h1 (P137 & G410) and h2 (G192 & G375) are hinge residues connecting equatorial and intermediate domains and intermediate and apical domains respectively. For comparison, two R-ATP states determined by cryo-EM Rd2 (lower black dash, PDB ID: 4AB2) and Rdopen

(upper black dash, PDB ID: 4AB3), are analyzed and plotted with the asymmetry of R in the R-ADP state. The **T** state with 7-fold symmetry (PDB ID: 1XCK) is also analyzed and plotted with the asymmetry of θ in the R-ADP state (black dash).

The signature plots show three features. First, the departure from symmetry that is evident in the intermediate and apical domains, begins and ends at the two hinge points, hinge1 (P137 and G410) and hinge 2 (G192 and G375), which adjoin the intermediate and equatorial and the apical and intermediate domains respectively (black arrows in Fig. 3-4, θ plot). This suggests that the asymmetry arises by small differences in the rigid body motion that accompanies the **T** to **R** transition. Consistent with this idea, the plots corresponding to different subunits are parallel to one another (Fig. 3-4, R plot), indicating that the secondary structure is conserved among subunits. Secondly, no two signature plots overlap at the apical domain (Fig. 3-4), meaning none of the two subunits adopts identical conformation. Finally, for a given residue the difference in R is smallest between neighboring subunits (Fig. 3-4, R plot), indicating that the conformation of a given subunit influences the conformation of the neighboring subunits.

3.3.2 Crystal structure of R-ADP closely resembles the cryo-EM structure of R-ATP

The flexibility of nucleotide-bound GroEL has recently been observed in the **R-ATP** by cryoEM [27]. After imposing 7-fold symmetry it was concluded that six different symmetrical conformations of **R-ATP** existed in solution. To compare the variation of the **R-ATP** to **R-ADP**, we plot the residue-dependent R and H of the two most distinct conformations of the **R-ATP** state (two dotted lines in Fig. 3-4). For >95% of the residues, the seven R and H signature plots of the asymmetric conformations in the **R-ADP** state (colored solid lines) are within the boundaries set by the Rd2 (lower dotted line) and the Rd-open (upper dotted line) conformations of the **R-ATP** state [27]. Therefore, the flexibility of the **R-ADP** in the crystal structure is consistent with the variation of the **R-ATP** conformations observed previously in solution by cryo-EM. However, due to the imposition of 7-fold symmetry upon each of the cryo-EM structures at 8.5 Å, the asymmetry evident in the crystal structure at 2.7 Å is lost.

To underscore the similarity of the **R-ADP** and **R-ATP** structures and their common asymmetry, we imposed 7-fold symmetry on each of the sub-units A-G to create synthetic **R-ADP-A⁷** to **R-ADP-G⁷** rings (Fig. 3-5). This affords us the opportunity to compare these crystal-based synthetic ensembles with the synthetic cryo-EM-based ensembles. Despite the differences in resolution (2.7 Å versus 8-9 Å) and correlation coefficients (Fig. 3-6 0.89 versus 0.69) for the x-ray and cryo-EM structures respectively, it is obvious that the synthetic “x-ray”

rings resemble very closely the synthetic “cryo-EM” rings, irrespective of the nature of the ligand (ADP or ATP) in the nucleotide-binding site. The synthetic **R-ADP** rings can readily be ranked from the most “closed” structure (smallest value for d1, Fig. 3-5) to the most “open” structure (largest value for d1). Thus, the synthetic **R-ADP-A**⁷ corresponds to the cryo-EM based structure Rd2. This is especially evident when the two models for the apical domain are superimposed upon their respective density maps (Fig. 3-5, left). Similarly, the synthetic **R-ADP-D**⁷ corresponds to the most open cryo-EM based Rdopen (Fig. 3-5, Fig. 3-4 right). We conclude that, were it not for the imposition of 7-fold symmetry upon the cryo-EM **R-ATP** structure, the asymmetry observed in the crystal structure of **R-ADP** would be found in solution in the cryo-EM structure of **R-ATP**.

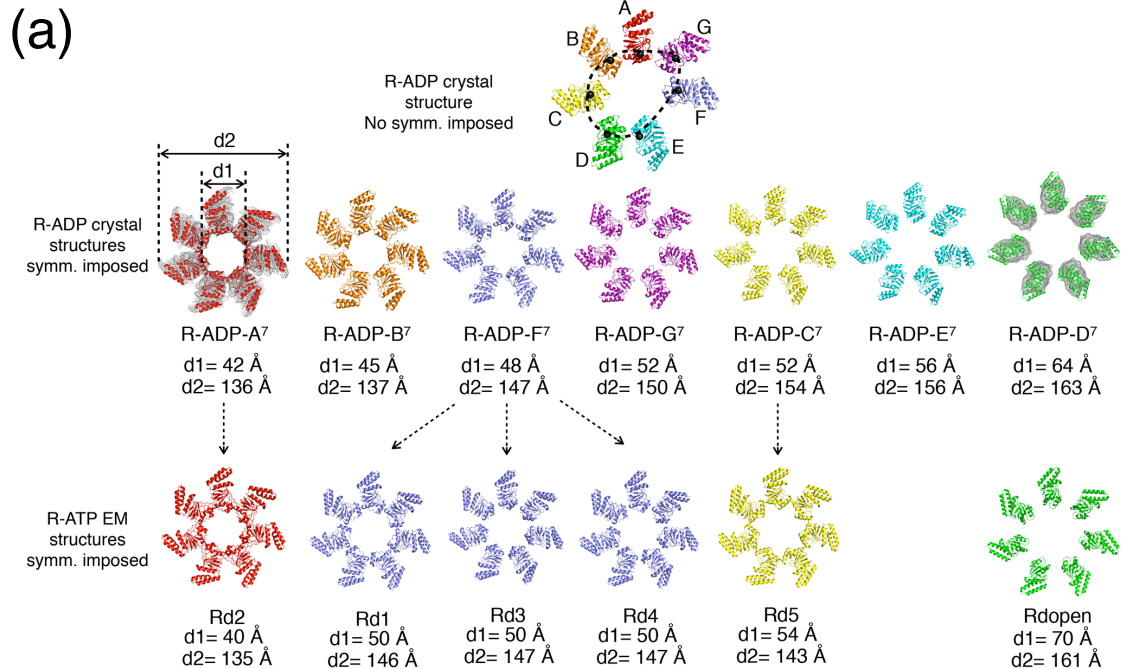


Fig. 3-5 Comparison between symmetry imposed R-ADP to R-ATP structures determined by cryoEM.

Each subunit in asymmetrical **R-ADP** (top) is isolated and symmetrized to generate synthetic **R-ADP** with 7-fold symmetry (upper line). These seven synthetic **R-ADP** structures are arranged from the least open one (**R-ADP-A⁷**) to the most open one (**R-ADP-D⁷**) according to their diameters (d1 and d2). Seven symmetrized **R-ADP** structures closely resembles a series of cryoEM **R-ATP** structures (lower line) in diameters and fit well to cryoEM electron densities (upper line left and right, **R-ADP-A⁷** vs. Rd2 and **R-ADP-D⁷** vs. Rdopen).

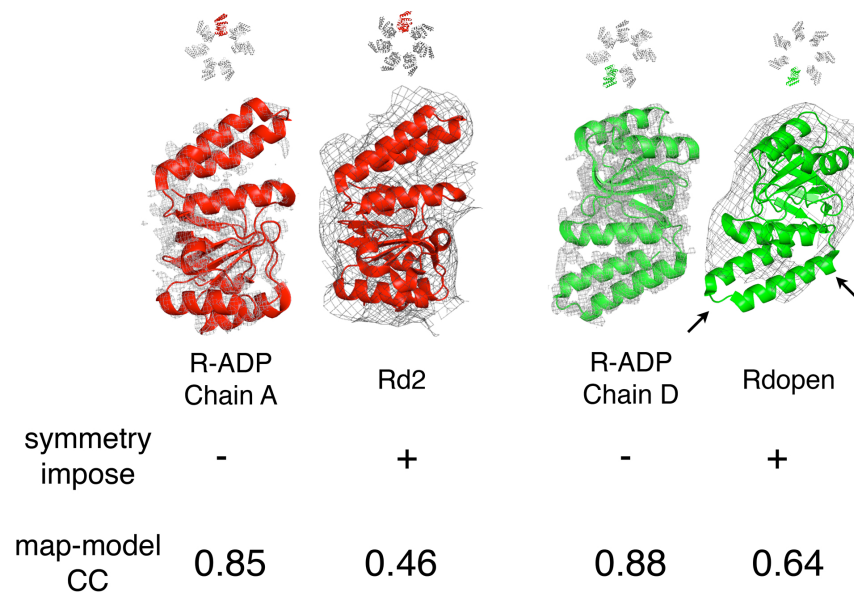


Fig. 3-6 Detailed view of model-density fitting in asymmetrical R-ADP and symmetrical R-ATP.

At the cost of imposing symmetry, cryoEM R-ATP models show rather poor correlation coefficient with maps and parts of the secondary structures show no electron density (black arrows). The apical domains are shown as cartoon. The

colors for symmetrized **R-ADP** are consistent with the asymmetrical **R-ADP** (top) and reflects the arrangement of subunits in a ring. **R-ATP** structures are colored to the symmetrized **R-ADP** with smallest RMSD. Electron density maps are contoured at 1.0σ .

3.3.3 The T to R transition is accompanied by the loss of inter-subunit stabilizing contacts as the origin of asymmetry.

The flexibility of intermediate and apical domains in the **R-ADP** is correlated with the loss of inter-subunit stabilizing contacts. In the **T** state, the intermediate and apical domains are stabilized by 42 inter-subunit hydrogen bonds or salt bridges per ring, 6 at each interface [11, 51]. However, in the crystal structure of **R-ADP**, these inter-subunit contacts involving intermediate and apical domains are absent (Fig. 3-7 and Table. 3-2). As a result, the intermediate and apical domains in **R-ADP** can rotate relatively freely with respect to the equatorial domains, leading to the loss of seven-fold symmetry. Thus, the cause of flexibility and asymmetry in **R-ADP** (loss of inter-subunit contacts) differs to that observed in another case [61]. In fact the asymmetric arrangement we observed in this crystal structure is only one possible conformation from an ensemble of flexible conformations the **R-ADP** may access. An alternate source of asymmetry concerns the intra-subunit salt-bridge between D155 and R395, both in the intermediate domain. The mutation D155A leads to an ATP-induced breakage of symmetry [62]. However, since this particular salt bridge is maintained in **R-ADP**, it cannot be the source of the asymmetry observed here.

Compared with the **T** or **R** state, the dramatically increased b-factors for intermediate and apical domains in the **R-ADP** support this argument (Table. 3-3) A computational study tracking the trajectories of the individual subunits of GroEL during the **T** to **R** transition reported this dynamic flexibility in the apical domain [63].

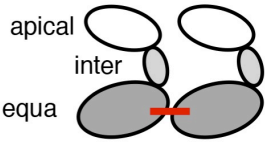
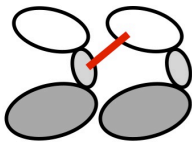
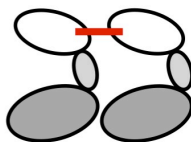
	equa-equa	inter-apical	apical-apical
	 <p>Diagram of the T state showing two subunits with equatorial domains (equa) in contact, indicated by a red line. Labels 'apical', 'inter', and 'equa' are present on the left subunit.</p>	 <p>Diagram of the T state showing two subunits with an intermediate domain (inter) and an apical domain (apical) in contact, indicated by a red line.</p>	 <p>Diagram of the T state showing two subunits with apical domains (apical) in contact, indicated by a red line.</p>
T	12	5	1
R-ADP	13	0	0

Fig. 3-7 Numbers of inter-subunit contacts between two neighboring subunits, in the asymmetric R-ADP state and symmetric T state.

These contacts include salt-bridges ($<5 \text{ \AA}$) and hydrogen bonds ($<3.5 \text{ \AA}$). Based on their locations, the contacts are categorized into three groups: between equatorial domains, between apical domains and between intermediate-apical domains.

Each interface in R-ADP									
	AB	BC	CD	DE	EF	FG	GA	R-ADP Average	T
Arg197-C β Glu386-C δ	16.6	17.3	20.4	16.8	24.2	20.4	17.3	19.0	4.6
Lys80-N ζ Glu386-C δ	12.2	13.6	14.8	16.9	19.7	13.2	13.7	14.9	17.8
Glu255-C δ Lys207-N ζ	17.3	17.2	28.8	17.7	28.6	22.9	12.8	20.8	4.0
Glu255-C δ Lys245-N ζ	4.2	7.7	16.7	10.8	22.9	6.3	12.1	11.5	16.4
Glu257-C δ Lys245-N ζ	14.7	9.8	12.6	9.1	25.1	14.6	20.8	15.2	17.5
Glu257-C δ Lys242-N ζ	6.2	12.6	15.0	16.5	25.2	3.1	20.0	14.1	12.5

Table. 3-2 Possible inter-subunits salt bridges in the R-ADP and T (PDB ID: 1XCK) states.

	R-ADP	T	R''
equatorial	41.4	29.2	22.5
intermediate	85.0	42.7	55.9
apical	131.2	60.5	98.9

Table. 3-3 Domain averaged B factors for the T(PDB ID: 1XCK), R-ADP and R''(PDB ID: 1AON) states.

3.3.4 R-ADP is distinct from other crystal structures and cryo-EM models of GroEL.

We quantified the averaged difference between the **R-ADP** state and previously solved crystal structures in the **T** [11] and **R''** [38] states. We used the same cylindrical coordinate system as described above. For each residue, we first calculate the average R and H over seven subunits within a ring. ΔR and ΔH is

the difference in R and H , respectively, between the **R-ADP** state and a previously solved state. While the difference between **T** and **R-ADP** states is larger than 10 Å for some residues (red and blue in Fig. 3-8), the difference between the **R''** and **R-ADP** states exceeds 20 Å (pink and cyan in Fig. 3-8). These differences are confined to the intermediate and apical domains.

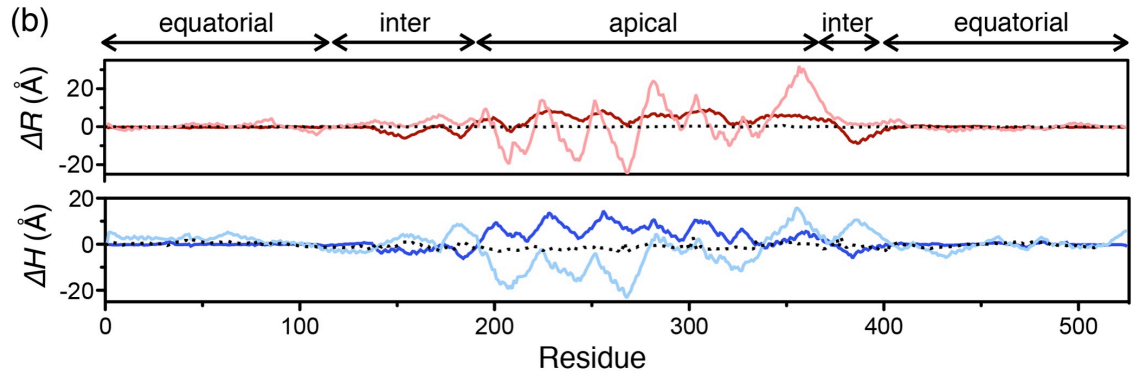


Fig. 3-8 Conformation differences between the R-ADP and T or R'' state, analyzed using difference “signature” plots.

Conformational difference in each residue is quantified by ΔR (upper panel) and ΔH (lower panel). $\Delta R = \langle R \rangle - \langle R' \rangle$, where R is the distance to the 7-fold axis in the **R-ADP**, R' is the corresponding distance in the **T** (red, PDB ID: 1XCK) or **R''** (pink, PDB ID: 1AON) state. $\langle \rangle$ denotes average over seven subunits. Similarly, $\Delta H = \langle H \rangle - \langle H' \rangle$, where H is the height in **R-ADP** state, and H' the corresponding height in a known state (**T** in blue and **R''** in cyan). ΔR and ΔH between two **T** state structures (black dash, PDB ID: 1XCK and 2NWC) serve as the negative control and showing two crystal structures of the same conformational state of GroEL only have little differences.

The difference between the **R-ADP** and the crystal structures of **T** or **R**” (solid lines in Fig. 3-8) is much larger than the difference between two structural models of the same state (black dotted lines in Fig. 3-8). Thus, the **R-ADP** state is distinct from previous known states. Further, the difference is not solely due to the flexibility of ADP bound GroEL, as the difference does not vanish upon averaging over seven subunits within a GroEL ring.

The **R-ADP** is also significantly different from the crystal structure of GroEL-ATP γ S₁₄ complex previously considered to be in the **R** state [64]. We quantified the difference between the **R-ADP** and GroEL-ATP γ S₁₄, using ΔR and ΔH as described in the main text. Although the contacts between nucleotides and nucleotide binding pocket are almost identical in these two structures (Fig. 3-9), the intermediate and apical domains of the two have distinctive conformations (Fig. 3-10 dashed pink and cyan). The difference between the **R-ADP** and GroEL-ATP γ S₁₄ may arise from a difference in crystallization methods. While we co-crystallize ADP and GroEL, ATP γ S was soaked in after the crystal had formed and lattice forces restrained any nucleotide-induced large conformational rearrangements. Therefore, GroEL-ATP γ S₁₄ is almost identical to the nucleotide free **T** state with small difference in the height of apical domain (~ 2 Å) (Fig. 3-10 solid red and blue).

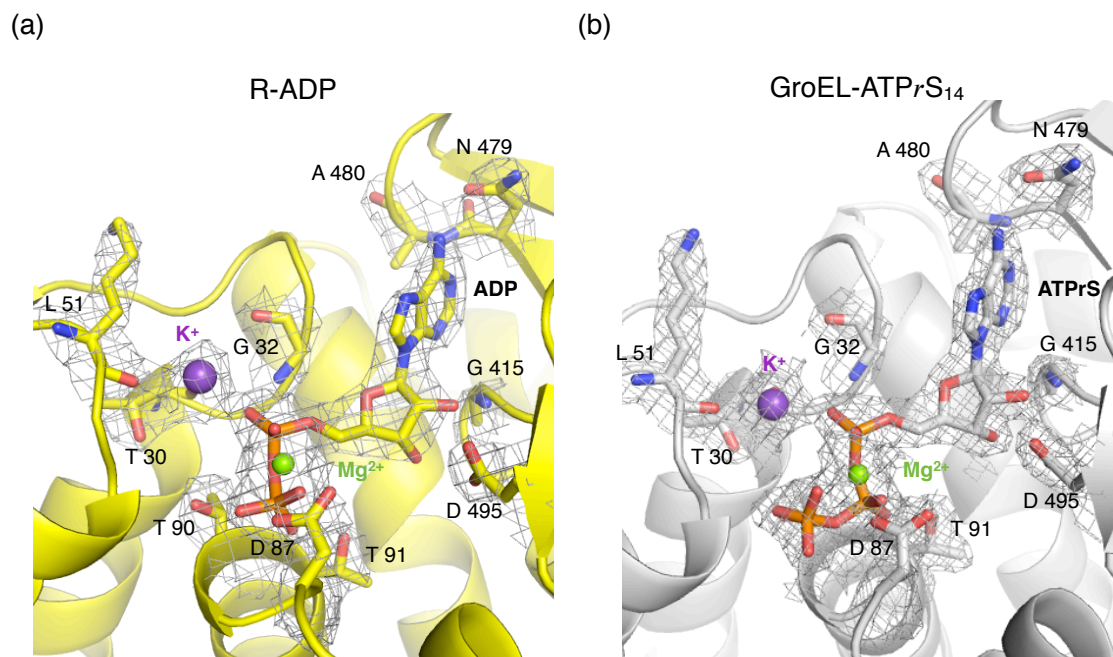


Fig. 3-9 Contact between nucleotides and the nucleotide binding residues in the R-ADP (a, yellow) and GroEL-ATP γ S₁₄ (b, grey PDB ID: 1KP8).

Despite the difference in the nucleotide (KMgADP or KMgATP γ S), the nucleotide binding residues (shown as sticks) in two structures adopt very similar conformations. The electron density maps (2Fo-Fc) are contoured at 1.5 σ .

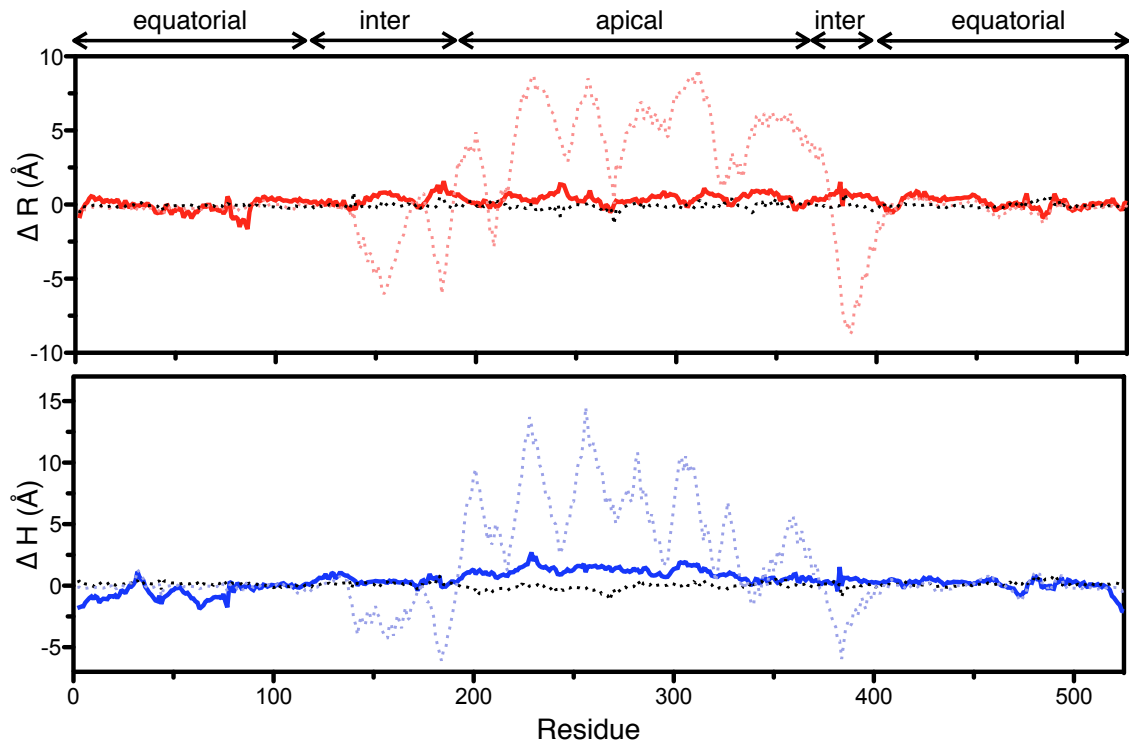


Fig. 3-10 Conformation differences between GroEL-ATP γ S₁₄ and the T or R-ADP state.

Conformational difference in each residue is quantified by ΔR (upper panel) and ΔH (lower panel). $\Delta R = \langle R \rangle - \langle R' \rangle$, where R is the distance to the 7-fold axis in GroEL-ATP γ S₁₄, R' is the corresponding distance in the **T** (red solid trace, PDB ID: 1XCK) or **R-ADP** (pink dashed trace) state. $\langle \rangle$ denotes average over seven subunits. Similarly, $\Delta H = \langle H \rangle - \langle H' \rangle$, where H is the height in GroEL-ATP γ S₁₄, and H' the corresponding height in a known state (the **T** in blue and **R-ADP** in cyan). ΔR and ΔH between two **T** state structures (black dash PDB ID: 1XCK and 2NWC) serve as the negative control and showing two crystal structures of the same conformational state of GroEL have only minor differences.

We also compared the **R-ADP** state GroEL to GroEL structures determined by cryo-EM at lower resolution. The **R-ADP** is similar to a series of cryo-EM **R-ATP** structures [27]. For seven asymmetrical subunits in the **R-ADP** GroEL, chain A has the conformation most close to **R-ATP** state Rd2 and chain D is most similar to **R-ATP** state Rd-open (red, cyan and two dashed black traces in Fig. 3-4).

The conformational differences between the **R-ADP** and other structures are mainly caused by rigid body rotation of the intermediate and apical domains around hinge1 (P137 and G410) and hinge2 (G192 and G375) [38]. For example, compared to the crystal structure of the **T** or **R''** states, although the averaged conformational differences between the intermediate and the apical domains can be greater than 20 Å (Fig. 3-8), the secondary structure within these domains remain largely unchanged. The RMSDs of isolated intermediate domains are only 0.7 ± 0.3 Å (between the **T** and **R-ADP**) and 1.0 ± 0.3 Å (between the **R''** and **R-ADP**). The corresponding deviations of isolated apical domains are 0.5 ± 0.1 Å and 1.1 ± 0.1 Å.

3.3.5 Rigid body rotation of the intermediate domain controls nucleotide release.

A structural comparison of the **R''** and **R-ADP** states indicates that the rotation of the intermediate domain controls access to the nucleotide-binding pocket. In progressing from the **R''** to the **R-ADP** state, the intermediate domains undergo a

10° rotation away from the 2-fold symmetry plane, with respect to hinge-1 (Fig. 3-10a). Consequently the nucleotide-binding pocket, which contains an intermediate domain lid (helices F, G and M) and equatorial domain nucleotide-binding loops, changes from a “fully closed” state to a “half-closed” state. This can be measured from the change in distance between the lid helices and the nucleotide binding loops. The average distance from helices F/G (N153) to the loops (P33) increases from 5.5 ± 0.3 Å to 7.1 ± 0.4 Å (Fig. 3-10b) while the average distance from helix M (R395) to the loops (D52) increases from 7.6 ± 0.1 Å to 10 ± 1 Å (Fig. 3-10c).

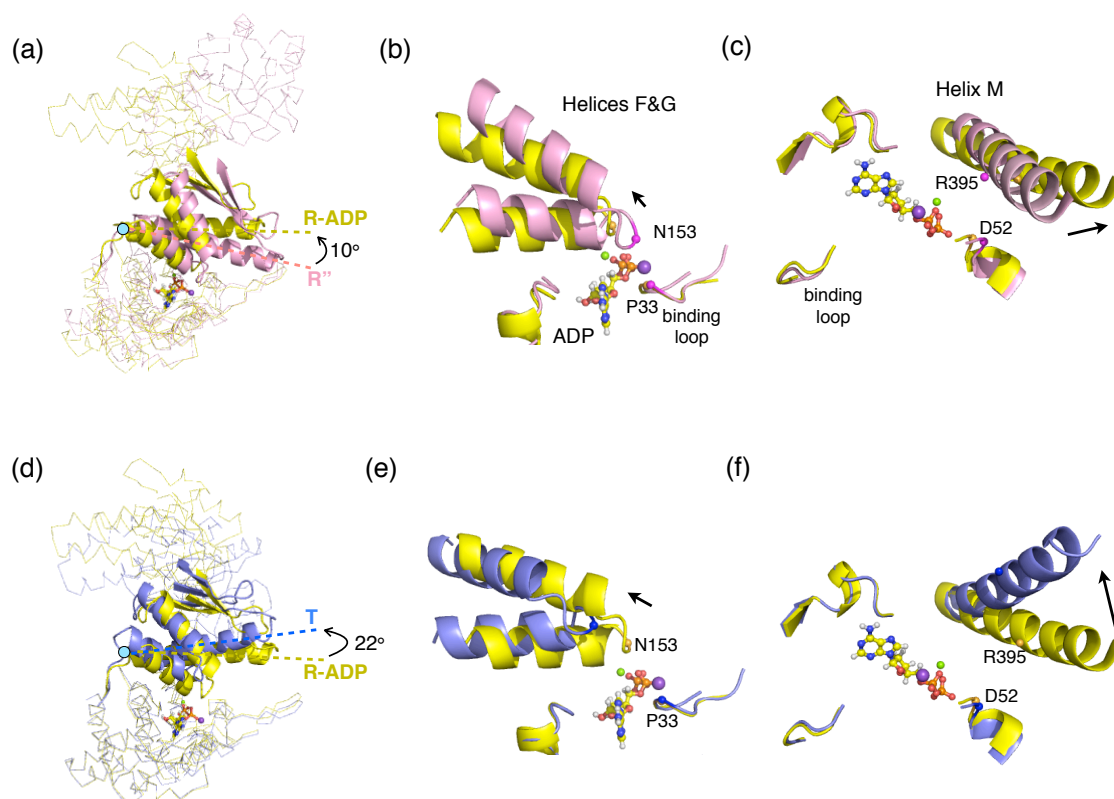


Fig. 3-11 Conformational changes in the intermediate domains, between the T (PDB ID: 1XCK), R-ADP and R' (PDB ID: 1AON) states.

(a) A subunit from the **R**'' (pink) aligned with a subunit from the **R-ADP** (yellow) by superposition of the two equatorial domains. The conformation difference between the intermediate domains (in cartoon) in two structures is the result of domain rotation around hinge-1 (cyan dot). The axes of helix M are shown as dashed lines. Black arrows indicate directions of hinged rotations of intermediate domains. (b-c) Detailed view of helices F&G and helix M in the superimposed structures in (a) showing the distance from these helices to the nucleotide binding loops changes as the result of hinged rotation. Two pairs of residues (P33&N153, D52&R395) are shown to highlight this distance change (C α s spheres). ADP is shown in ball-and stick, Mg²⁺ and K⁺ are shown as green and purple spheres. Black arrows indicate relative position change of these helices from **R**'' state to **R-ADP** state. (d-f) are same as (a-c), except showing the rotation of the intermediate domain and the position of helices F&G and helix M as GroEL switches from the **R-ADP** state (yellow) to the **T** state (blue).

The structural transition of the nucleotide-binding site from “fully closed” in the **R**'' state to the “half-closed” **R-ADP** state is accompanied by an increase in the rate of ADP release. ADP, sequestered in the *cis* ring, in the **R**'' state does not exchange with free ligand in solution at all. On the other hand, ADP dissociates for the “half-closed” **R-ADP** state at a rate of $\sim 0.1\text{s}^{-1}$ compared with the pre-steady state rate of ATP hydrolysis of $\sim 0.6\text{ s}^{-1}$ [17, 39]. In other words, ADP release from the “half-closed” **R-ADP** state, while faster than release from the **R**'' state, is still slower than ATP hydrolysis and thus remains the rate-

determining step in the chaperonin cycle in the absence of SP [17, 19]. By introducing the mutations that stabilized the “half-closed” **R** state, we expect ADP release to be even slower. This is reflected in the three-fold increase in the mean residence time of GroES dissociation, induced upon adding ATP to the resting state complex of GroEL^{D83A/R197A} and wild-type [17].

During the **R-ADP** to **T** transition the intermediate domain rotates away from cavity center and further away from the 2-fold plane of symmetry by total of 22°(Fig. 3-11d). This brings about a change in the nucleotide-binding site from the “half-closed” state to the “open” state. The average distance between the lid helices F/G (N153) and nucleotide binding loop (P33) increases from 7.1±0.4 Å to 12.2±0.2 Å (Fig. 3-10e). Similarly the average separation between lid helix M (R395) and nucleotide binding loop (D52) expanded from 10±1 Å to 14.2±0.1 Å (Fig. 3-11f). The opening of the nucleotide-binding pocket associated with the reformation of the **T** state permits rapid ADP/ATP exchange that is necessary for continued cycling of the chaperonin machine.

While the average distance between lid and loop is related to the rate of ADP release, the variance of distance in the **R-ADP** state explains SP accelerated ADP release. The variance of the distance is significantly larger in the **R-ADP** than in the other two, as the result of asymmetry (Fig. 3-4). For example, the distance from helix M lid to loop varies from 8.5 Å to 11.9 Å in the **R-ADP**, approaching the average lid-to-loop distance in the “open” **T** state (14.2±0.1 Å) on the one

hand, and the “fully-closed” **R”** (7.6 ± 0.1 Å) state on the other. As described above, this variance is the result of flexibility of the **R-ADP**, and very likely exists in solution as well. Thus, it is likely that the nucleotide binding pockets in the **R-ADP** switch stochastically between the “open” and the “fully closed” conformation in solution, due to its structural flexibility.

We suggest that SP accelerates ADP/ATP exchange by selectively binding to GroEL with “open” nucleotide-binding pockets. SP would consequently increase the population and life-time of the “open” conformation permitting rapid nucleotide exchange. This mechanism is similar to the out-of-equilibrium conformational cycling that others have proposed [65].

3.3.6 Rigid body rotation of apical domain enables selective binding to GroES or substrate protein

The apical domains rotate with respect to hinge-2 when GroEL cycles through the **R”** to **R-ADP** to **T** states (cyan dots in Fig. 3-12c). The apical domains bind both SP and GroES that share the same binding site; the groove between helices H and I [13, 66]. Hinged rigid body rotation changes the position of helices H and I and the binding grooves between two helices. Thus, GroEL binds to SP or GroES selectively at different stages of its cycle: the **R”** state binds GroES and the **T** binds SP while the **R-ADP** state has a conformation intermediate.

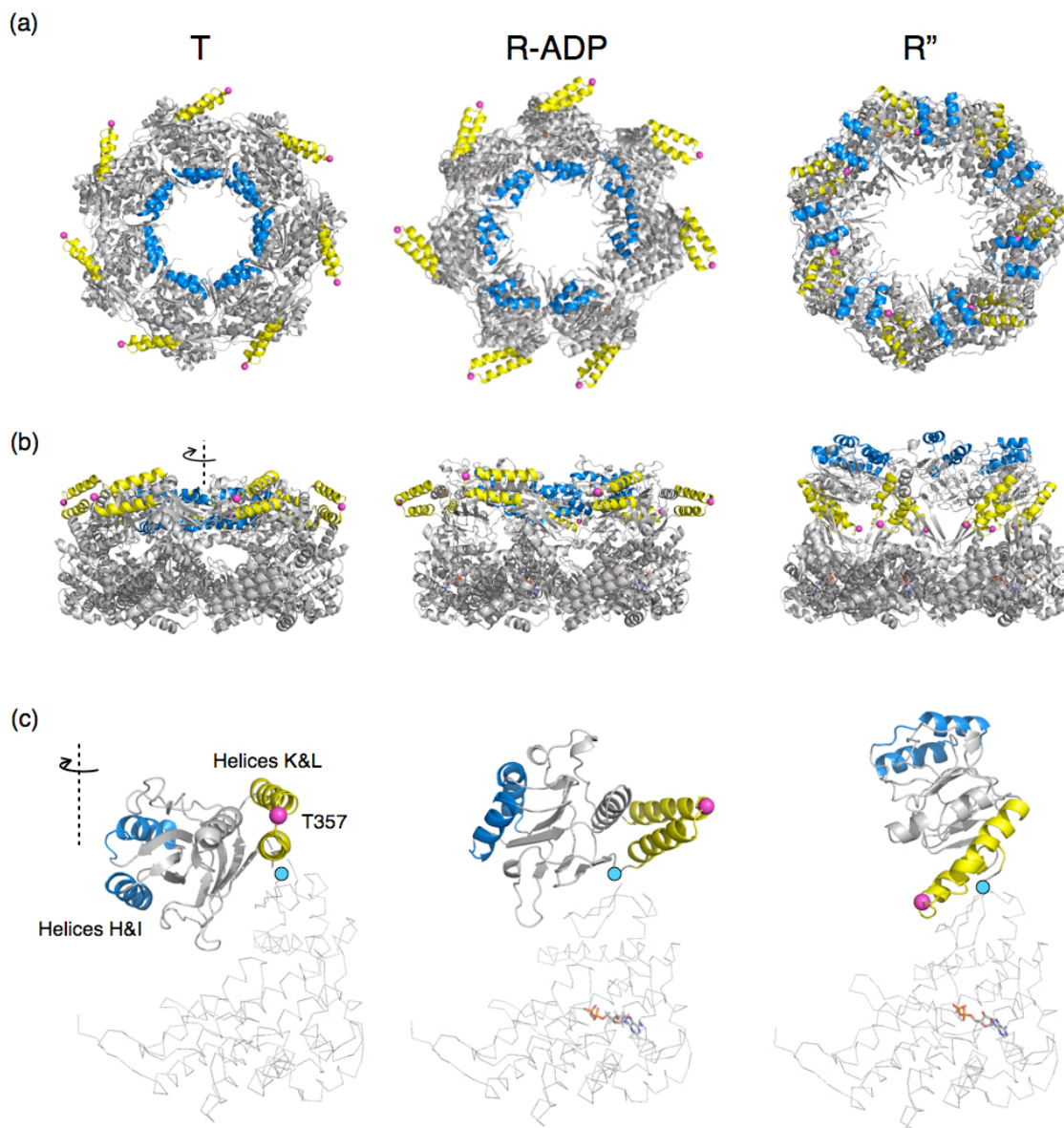


Fig. 3-12 Conformational changes in apical domains, between the T (PDB ID: 1XCK), R-ADP and R'' (PDB ID: 1AON) states.

(a) The top views and (b) sides views of a single GroEL ring in the T (left), R-ADP (middle) and R'' state (right). Positions of helices H and I (blue), helices K and L (yellow) and T357 (pink spheres) are highlighted to trace apical domain rotations. (c) Detailed views of one subunit from each state. The conformation

differences between the intermediate domains (in cartoon) are the result of domain rotation around hinge-2 (cyan dots). ADP is shown in sticks. Note in particular that the major motion of helices K and L, a flip that brings T357 from position on the external surface of GroEL to one on the internal surface, occurs during the **R** to **R''** transition.

In progressing from the **R''** state to the **R-ADP** state, the apical domains in the **R''** state swing about 80° clock-wise and move toward the 2-fold plane (Fig. 3-12, right and middle). As the apical domains move and the overall shape of GroEL changes, the position of the SP/GroES binding helices H and I changes dramatically. In **R''** state, helices H and I present themselves at the upper surface of GroEL ring, with the binding grooves between two helices fully exposed on the GroEL surface and well separated to each other (Fig.3-12a, 3-12b, 3-12c right). In **R-ADP** state, however, these two helices tilt towards the center of the cavity and partially exposing the binding grooves between helices H and I to surface (Fig. 3-12a, middle). At the same time in the **R-ADP** state these helices move closer to each other and line the collar of GroEL cavity (Fig. 3-12a, 3-12b, middle).

The structure of the **R-ADP** suggests that it has lower affinity for GroES, compared to the **R''** state. The binding grooves in **R''** favor GroES binding structurally by adopting a separated and exposed conformation, which adapt to the seven well-separated “mobile” loops of GroES dome [38]. However, in the

R-ADP state the binding grooves rotate towards the center of the ring and become partially buried, so that GroES binds less readily to **R-ADP** state than it does to the **R''** state. Reduced affinity for GroES would prevent it binding to **R-ADP** state GroEL before SP capture.

During the **R-ADP** to **T** transition the apical domains swing about 30° counter-clock-wise and move closer to the 2-fold plate (Fig. 3-12, middle and left). The partially exposed binding groove between helices H and I in the **R-ADP** state now re-orient such that they point towards the center of cavity in the **T** state (Fig. 3-12c, middle and left). The circumference of the binding collar that contains seven pairs of helices H and I decreases (Fig. 3-12a, middle and left), as a consequence of which, the SP binding sites on adjacent subunits move 7 Å closer on average. Such conformational difference between the **R-ADP** and **T** state suggest that the **T** state binds SP tighter than the **R-ADP**. Because SP need to interact with multiple binding sites in order to bind GroEL with nanomolar-affinity [22], the **T** state with smaller SP binding collar and shorter distance between adjacent binding sites is likely to have higher affinity to SP than the **R-ADP** state.

In *E. coli* GroEL interacts with half of the soluble, misfolded proteins, diverse in sequence and structure [10]. In order to capture so many heterogeneous proteins, the binding sites on GroEL have to be adaptive. The flexibility of apical domains would allow the **R-ADP** to adjust its binding collar according to a specific

misfolded protein. While it enables high promiscuity, the flexibility of the apical domain may lead to weaker affinity for SP. Nevertheless, not until SP binds could GroEL release bound ADP and proceed to the **T** state.

3.4 Summary

The crystal structure at 2.7 Å resolution of the GroEL in its relaxed **R** allosteric state that we report here differs in important respects from the ensemble of cryo-EM structures solved at lower resolution [27]. Due to symmetry that was artificially imposed on the cryo-EM images, an ensemble of GroEL rings was created that essentially consists of the sub-unit conformations that we observe within a single ring. In contrast, our data show that the asymmetry within a ring results from the intrinsic flexibility of the apical domain. Inter-subunit, apical domain contacts stabilize the **T** state. When these contacts are broken upon the addition of ATP/ADP, the apical domains become free to express the intrinsic flexibility that is characteristic of the **R** state.

Hitherto, conventional wisdom dictated that the oligomeric class I chaperonin rings were fundamentally symmetrical. Previously determined crystal structures of GroEL in the nucleotide free **T** state and GroES bound **R'** state were both almost perfectly symmetric (RMSD between chains <0.3) [11, 38]. As the result, it was further speculated that this symmetry was maintained in the ADP/ATP bound **R** state and during the various allosteric transitions [27]. However, asymmetry in the related class II chaperonins has been known for some time

[67]. Here, we present an example of how this symmetry is broken in the class I chaperonin, GroEL, merely by removing two salt bridges that are normally broken during the **T** to **R** transition. Intriguingly, one example has very recently emerged in classes I chaperonins [32], raising the possibility that transient departures from structural symmetry may be an important and intrinsic part of the mechanism(s).

Chapter 4: The crystal structure of GroEL^{D83A/R197A}

4.1 Introduction

The first crystal structure of GroEL in the **R** state, highlighting its unprecedented asymmetry was described in chapter 3. Although biochemical studies of this mutant have shown that these two mutations only destabilize the **T** state but do not eliminate it [39, 43], one might still argue these mutations are the source of the asymmetry.

To confirm that by removing two salt bridges (D83-K327 and E386-R197) alone, which ordinarily break during **T** to **R** transition, did not cause significant conformational changes and asymmetry, we determined the crystal structure of nucleotide free (apo) GroEL^{D83A/R197A} at the resolution of 3.11 Å.

4.2 Methods specific to chapter 4

4.2.1 Crystallization

The GroEL^{D83A/R197A} was screened for crystallization by using the sitting-drop vapor diffusion method (drop: 0.5 µL well solution: 100 µL), using the ARI Crystal PHOENIX robot. Diffraction quality crystals were grown at the screen condition with well solution consisting of 25% PEG 3000 (w/v), 0.1M acetic acid pH 4.5. Crystals were optimized by the hanging-drop vapor diffusion method (drop: 3 µL well solution: 1 mL). Long cubic shaped crystals appeared after 3 days in drops containing 2:1 mixture of protein sample and well solution.

Three days after the crystallization trays were set up, single crystals were removed from the mother liquor and Cryo-protected in well solution containing 20% glycerol before being mounted and flash-frozen by liquid nitrogen.

4.2.2 Data collection and structural determination

Diffraction data were collected at the NE-CAT beamline 24-ID-E located at the Advanced Photon Source, Argonne National Laboratory (Argonne, IL, USA). 300 frames with 0.5° oscillation were collected at 100K. The data were indexed and integrated using iMosflm [68]. The structure was solved by molecular replacement using apo wild-type GroEL (PDB ID: 1XCK) as the search model. Molecular replacement and structure refinement were carried out using AutoMR, Refine and Coot in Phenix suites [60].

4.3 Results and Discussion

4.3.1 apo-GroEL^{D83A/R197A} is in the T state

The overall architecture of GroEL is preserved in this mutant. The 14 subunits are arranged into two heptameric rings, stacked back-to-back. Each subunit consists of three domains: equatorial, intermediate and apical, as previously described for other high-resolution structures (Fig. 4-1). Data and refinement statistics are listed in Table 4.1. The asymmetric unit contains a full functional unit, i.e. a 14mer. One of the 14 subunits, subunit K, has a weaker apical domain density than the other subunits.

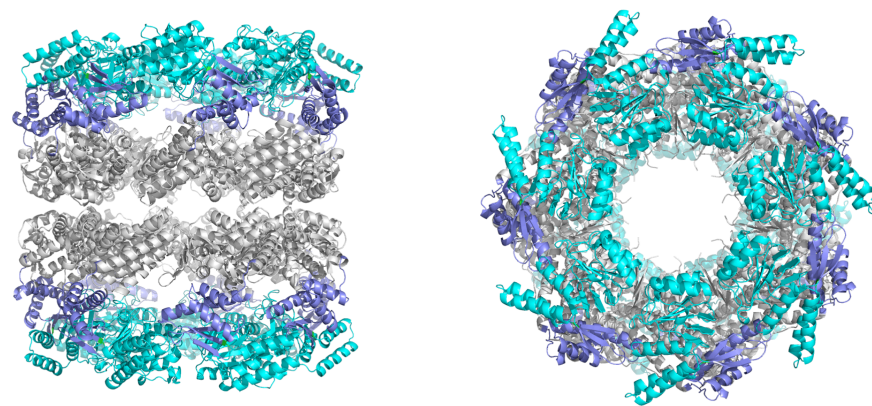


Fig. 4-1 Overall structure of GroEL^{D83A/R197A}.

Equatorial, intermediate and apical domains are colored in grey, blue and cyan respectively.

Space group	P 21 21 21	
Unit cell a, b, c (Å)	135.62/259.71/280.85	
Resolution range (Å)	123.5-3.13(3.232-3.13)	
Rmerge (%)	13.8(68.8)	
Total number of reflections	174049(17260)	
Redundancy	5.3(5.6)	
Completeness (%)	99.62(99.84)	
I/σ(I)	10.2(3.0)	
Number of residues/asymmetry unit		3668
Number of protein atoms		27012
Number of ligand atoms		368
Number of water atoms		597
R (%)	16.65 (26.31)	
Rfree (%)	22.09 (33.63)	
RMSD		
bond lengths (Å)		0.01
bond angles (degree)		1.36
Ramachandran plot		
Most favored (%)		98
allowed(%)		1.52
outlier(%)		0.48
Average B factors (Å²)		
Protein		81.7

Table. 4-1 Data collection and refinement statistics of GroEL^{D83A/R197A}.

Compared with other known conformation states, including the **T**, **R**, **R'** and **R''** state. GroEL^{D83A/R197A} is very similar to the wild-type **T** state and significantly different from **R-ADP** structure in the same mutation background (Fig.4-2). For example the apical domain helices H and I in GroEL^{D83A/R197A} form a small and compact collar just like in the wild-type **T** state, while in GroEL^{D83A/R197A}-ADP₁₄ a ring of seven helices H and I move away from each other (Fig. 4-2a).

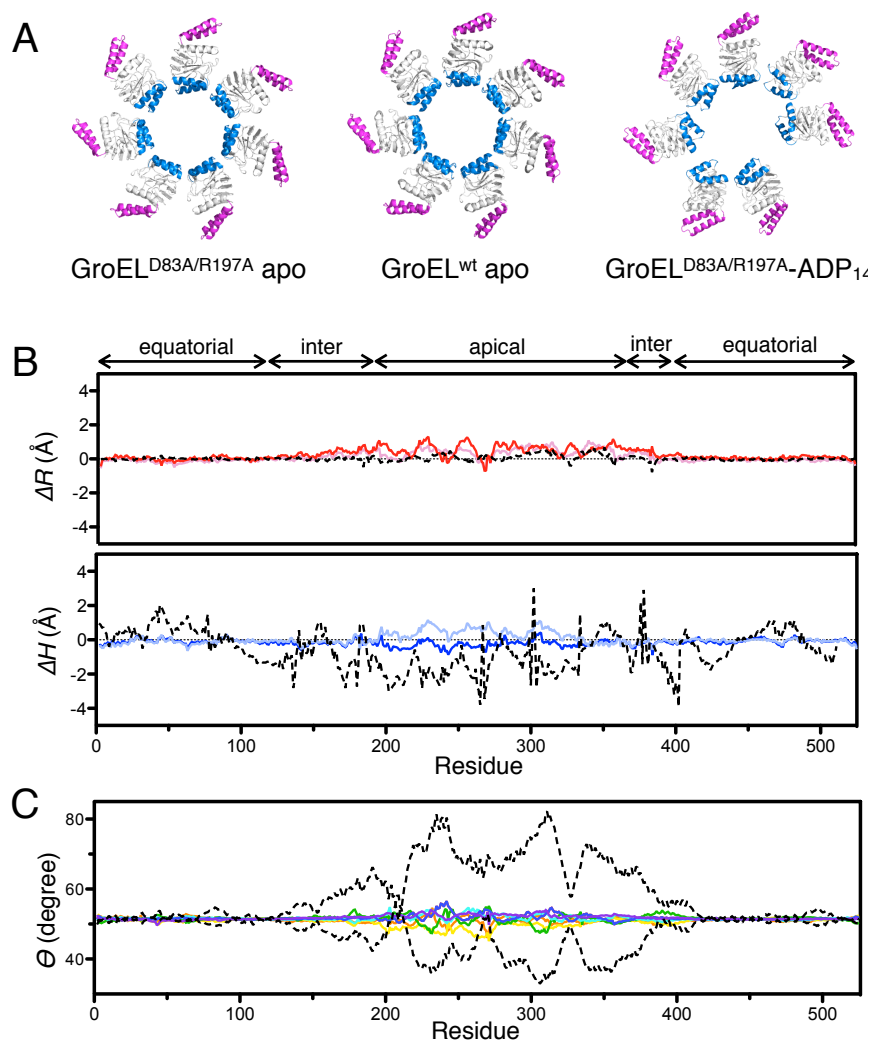


Fig. 4-2 Nucleotide-free GroEL^{D83A/R197A} is in the T state.

(A) Comparison between apo-GroEL^{D83A/R197A} and (i) the apo-GroEL^{wt} in the T state and (ii) GroEL^{D83A/R197A}-ADP₁₄ in the R state. Helices H and I are colored in blue, helices K and L are colored in pink and the rest of apical domains are colored in grey. (B) Quantitative comparison between apo-GroEL^{D83A/R197A} and apo-GroEL^{wt}, using difference “signature” plots. $\Delta R = \langle R \rangle - \langle R' \rangle$, where $\langle R \rangle$ is the average distance from the C α of residue i to the seven-fold symmetry axis (red: apo-GroEL^{D83A/R197A} subunit A-G; pink: apo-GroEL^{D83A/R197A} subunit H-N)

and $\langle R' \rangle$ is the corresponding distance in apo-GroEL^{wt}. Similarly $\Delta H = \langle H \rangle - \langle H' \rangle$, where $\langle H \rangle$ is the average distance from the C α of residue i to the two-fold symmetry axis between rings (blue: apo-GroEL^{D83A/R197A} subunit A-G; light blue: apo apo-GroEL^{D83A/R197A} subunit H-N) and $\langle H' \rangle$ is the corresponding distance in apo-GroEL^{wt}. (C) θ plots show the degree of asymmetry of apo-GroEL^{D83A/R197A}. Seven colored traces (red, orange, yellow, green, cyan, blue and purple) correspond to seven subunits in a ring of apo-GroEL^{D83A/R197A}, two dotted traces showing the maximum degree of asymmetry observed in the R state. ΔR and ΔH between two T state structures (black dashed lines; PDB ID: 1XCK and 2NWC) serve as negative controls.

The quantitative “signature plot” analysis agrees with visual inspection. The difference “signature plots”, ΔR and ΔH show the differences between both rings of GroEL^{D83A/R197A} and GroEL^{WT} are within experimental error (Fig. 4-2b). The “signature plot” of theta shows seven subunits in one ring of GroEL^{D83A/R197A} only slightly deviate from seven-fold symmetry, by up to 5 degrees. But for GroEL^{D83A/R197A}-ADP₁₄, in which the seven subunits in one ring all have different conformations, the asymmetry in theta exceeds 30 degrees for some residues (Fig. 4-2c). In summary apo GroEL^{D83A/R197A} is in the T state, just like apo-GroEL^{WT}.

The structure of apo-GroEL^{D83A/R197A} suggests that removal of two salt bridges stabilizing the T state is insufficient to eliminate the T state. This result is

consistent with biochemical studies, which shows this mutant populates the **T** state at very low ATP concentrations. Moreover, the similarity between apo-GroEL^{D83A/R197A} and apo-GroEL^{WT} support our previous claim. The unique conformation of GroEL^{D83A/R197A}-ADP₁₄ including its dramatic deviation from seven-fold symmetry, is not caused by the mutations, but instead caused by ADP. The cryo-EM structures of the wild-type GroEL-ATP complex at ~ 9 Å resolution displayed similar asymmetry [27]. Thus it can be concluded that this asymmetric state is populated at least once and perhaps twice during the course of the chaperonin cycle.

4.3.2 A systematic analysis of all inter-subunit salt bridges.

Salt bridges are commonly observed non-covalent interactions, which stabilize a protein's tertiary and quaternary structures. For GroEL, conformational changes are correlated with the breaking and formation of salt bridges [63]. To understand why breaking two salt bridges did not make the **T** state of apo-GroEL^{D83A/R197A} asymmetrical, we analyzed the inter-subunit, intra-ring salt bridges in three different conformational states, **T** (PDB ID: 1XCK), **R-ADP** (PDB ID: 4KI8) and **R'** (PDB ID: 1AON).

Salt bridges are a combination of two non-covalent interactions, hydrogen bond and electrostatic interactions [69, 70, 71]. The strength of a salt bridge depends on the distance between two charged side chains. The distance between two

centroids of charges has to be less than 4 Å for a strong salt bridge while weaker salt bridges extend out to ~ 6Å [69, 70, 71].

However, some measurements of the distances between two centroids have larger uncertainties than the others. That is because some of the charged side chains are not crystallographically well defined. With the average resolution of approximately 3 Å [11, 38, 39], side chains of the flexible residues of the apical domain including Arg or Lys, are not visible in the electron density map contoured at 1.0 σ .

To account for all possible salt bridges including those involving flexible side chains, we estimate the uncertainty of distance between two centroids of charges using B factors. The B factor of an atom is proportional to the mean square displacement of this atom caused by thermal motion:

$$B = 8\pi^2 \langle \mu^2 \rangle$$

where $\langle \mu^2 \rangle$ is the mean square displacement of an atom

Assuming that the B factor is isotropic, meaning that the atom has the same amount of thermo flexibility in x, y, and z directions, then we can deduced the uncertainty of the distance between two centroids of charges:

$$\delta D = 0.5 \times (\sqrt{B1} + \sqrt{B2}) \times (|X2 - X1| + |Y2 - Y1| + |Z2 - Z1|) \div \pi \times \sqrt{6} \div D$$

Where $B1$ and $B2$ are B factors of two centroids of charges, $(X1, Y1, Z1)$ and $(X2, Y2, Z2)$ are the coordinates of two centroids.

The interaction between two charged side chains is a salt bridge when:

$$D - \delta D \leq 4 \text{ \AA}$$

Where D is the distance between two centroids of charges and δD is the uncertainty of the distance.

Using the method described above, we found the T state has the largest number of inter-subunit salt bridges, more than twice as many as the **R-ADP** and **R''** state (Fig. 4-3). Although a salt bridge between E386 and R197 will be disrupted by the R197A mutation, the negatively charged E386 interacts with other positively charged side chains, including R285 and K277. Therefore the interactions between neighboring intermediate and apical domains are maintained, with a total of 16 salt bridges still present. However, if an E386A mutant were created instead of R197A, the **T** state and inter-subunit interactions would be destabilized even more.

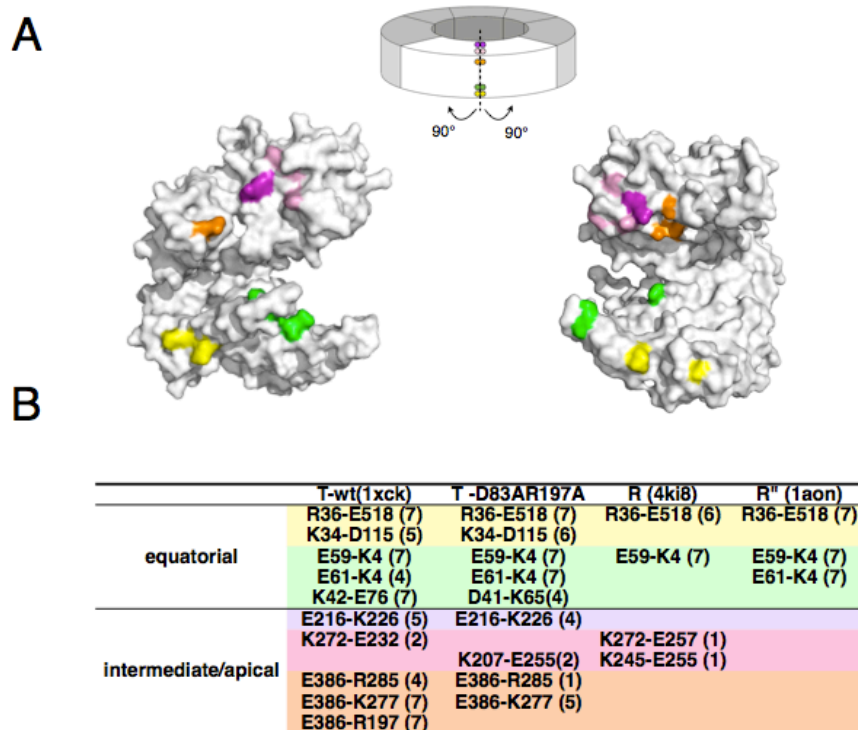


Fig. 4-3 Analysis of all inter-subunit salt bridges in the T, R, and R'' state.

(A) All inter-subunit salt bridges can be categorized into five patches, based on their locations. The inter-subunit interface is shown in white surface, residues involved in inter-subunit salt bridges are colored as in (B). **(B)** Residues involved in inter-subunit salt bridges. Two residues of opposite charges are considered forming a salt bridge when the $D - \delta D < 4.0 \text{ \AA}$. d is the distance between the centroids of two charged side chains, δD is the uncertainty of distance calculate from the B factors.

4.3 Summary

The crystal structure of nucleotide free GroEL^{D83A/R197A} is shown in this chapter. This mutant has very similar conformation to the apo-GroEL^{wt} in the **T** state but different from GroEL^{D83A/R197A}-ADP₁₄ (Fig. 4-2). This indicates that the asymmetry we observed in GroEL^{D83A/R197A}-ADP₁₄ is not caused by mutations.

We also performed a systematic analysis on inter-subunit salt bridges of the **T**, **R** and **R''** states. With the exception of one inter-subunit salt bridge removed in this mutant (E386-R197), the apical and intermediate domains of the **T** state are connected with four other salt bridges (Fig.4-3). Therefore the seven-fold symmetry in GroEL^{D83A/R197A} is maintained.

Chapter 5 Crystal structures of symmetric “football”

complexes: $[\text{GroEL}-(\text{ADPBeF}_3)_7\text{-GroES}]_2$

5.1 Introduction

For many years, the GroEL/ES machine has been described as a two-stroke motor [16, 50]. But the two-stroke motor model is only accurate in the absence of SP. Using multiple spectroscopic probes, our lab recently showed that in the presence of misfolded proteins, GroEL/ES works as a parallel processor [18, 19]. When the two GroEL rings process misfolded proteins simultaneously under turnover conditions, “football” complexes $[\text{GroEL-SP-GroES}]_2$ are predominate functional forms [18, 19, 33].

In normal circumstances the football complexes formed in the presence of SP are dynamic, with a lifetime of ~ 1 s, after which time only a tiny fraction ($<1\%\sim 2\%$) of the transiently encapsulated SP will have folded to the native state [18, 72].

Under these cycling conditions it is unclear exactly where the SP folds.

Nevertheless, the chaperonins function as parallel processing devices and not as alternating machines. Regardless of SP, the football complexes become indefinitely stable in the presence of $\text{ADP}+\text{BeF}_3$, enabling the formation of diffraction-quality crystals.

Two crystal structures of the football complex were determined; one is empty, the other contains encapsulated SP in both chambers. However, the encapsulated

SP is not visible in the electron density map. Compared with the bullet-shaped complex (PDB ID: 1AON), the football complexes differ at the interface between the rings, suggesting a structural basis for negative inter-ring cooperativity.

5.2 Methods specific to chapter 5

5.2.1 Crystallization

The “MT-football” complex: GroEL:GroES₂ was made by mixing 150 μ M GroEL^{wt}, 900 μ M GroES^{wt}, 3.6 mM ATP, 4.8 mM BeCl₂ and 48 mM KF in Buffer A (50mM Tris-Acetate pH 7.5, 20 mM MgCl₂, 200 mM KCl). Crystals containing both GroEL and GroES were first grown in 25% PEG 550 monomethyl ether (v/v) and 0.1 M Tris-HCl buffer, pH 8.5. In crystal optimization followed, diffraction-quality crystals were grown in 8.5% PEG 550 monomethyl ether (v/v) and 0.1 M acetic acid-KCl buffer, pH 5.0. These crystals took about 10 days to fully-grow and can reach their maximum size, approximately 1 mm \times 0.5 mm \times 0.5 mm. Before crystals were mounted, they were removed from mother liquor and soaked in dehydration buffer (12% PEG 550 monomethyl ether (v/v), 20% ethylene glycol, 0.1 M acetic acid pH 5.25, 20 mM MgCl₂, 200 mM KCl, 1 mM BeCl₂ and 10 mM KF) for 3 to 7 minutes.

The Rubisco containing “football” complex: GroEL-GroES₂-Rubisco₂ (“SP-Football”) was prepared by first mixing 150 μ M GroEL^{wt}, 900 μ M GroES^{wt} and

42 μ M acid-denatured Rubisco-His6 in Buffer A. The protein mixture was placed at room temperature for 15 minutes. After that, ATP (3.6 mM), BeCl₂ (4.8 mM) and KF (48 mM) were added to the protein mixture. Crystals of “SP-Football” were obtained at conditions similar to the “MT-Football”. Single cubic shaped crystals were mounted after a similar dehydration procedure.

5.2.2 Data collection and structure determination

Diffraction data for the SP free “football” complex were collected at the NE-CAT beamline 24-ID-E located at the Advanced Photon Source, Argonne National Laboratory (Argonne, IL, USA). 300 frames with 0.5° oscillation were collected at 100K. The data were indexed and integrated using iMOSFLM [68]. The structure was solved by segmented molecular replacement. Seven search models were fitted one at a time. Each model consisted of one GroEL subunit and one GroES subunit extracted from the cis-ring of GroEL-GroES₁-ADP₇ (1AON).

Diffraction data of the “SP-football” complex were collected at the NE-CAT beamline 24-ID-C located at the Advanced Photon Source, Argonne National Laboratory (Argonne, IL, USA). 900 frames with 0.2° oscillation were collected at 100K. The data were indexed and integrated using iMOSFLM [68]. The structure was solved by molecular replacement using “MT-football” complex as the search model.

Structural refinement of the “MT-football” and “SP-football” complexes were carried using standard procedures. Refinements include rigid body, NCS (non crystallographic symmetry) and TLS refinement. Molecular replacement and structure refinement were carried out using AutoMR, Refine and Coot in Phenix suites [60]. Over 96% of the residues in the model of “MT-football” and “SP-football” are within the favored regions of the Ramachandran plot. The asymmetric unit of both “MT-” and “SP-football” contain a full functional unit: GroEL₁₄-GroES₁₄.

5.2.3 Interface analysis.

Hydrogen bonds, salt bridges and interface surface areas are analyzed using PISA [73].

5.3 Results and Discussion

5.3.1 Overall structures of the “football” complexes, with or without SP.

Two crystal structures of two [GroEL-(ADPBeF₃)₇-GroES]₂ “football” complexes were determined; one devoid of SP (MT football) (Fig. 5-1) and the SP football containing encapsulated Rubisco (Fig. 5-1 and Table. 5-1) The crystal packing of both football complexes is almost identical. One layer of footballs pack with their seven-fold axis parallel to one another and the footballs in the next layer pack with their seven-fold axis orthogonal to the footballs in the first layer (Fig. 5-2). Both football complexes consist of two heptameric GroEL rings, capped by

two heptameric GroES “lids.” All 14 nucleotide-binding sites on GroEL are saturated with the ATP analog ADP-BeF₃.

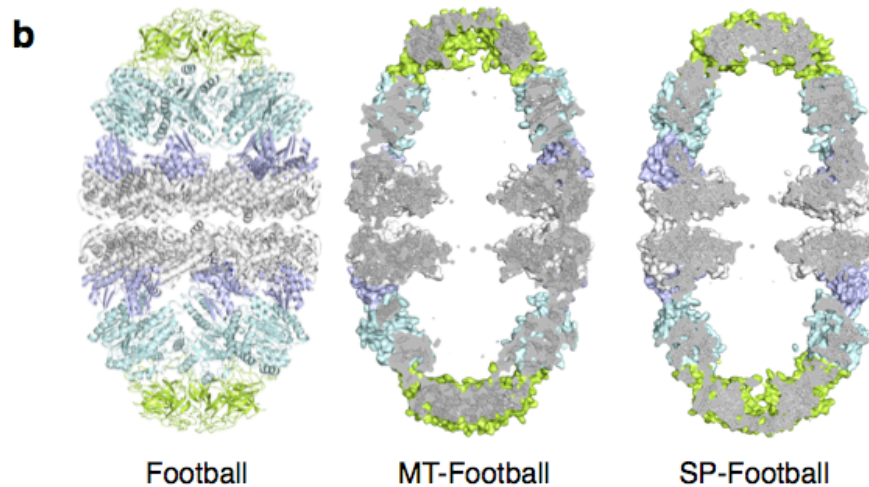


Fig. 5-1 Crystal structures of the “MT-football” complex and the “SP-football” complex with encapsulated Rubisco.

Three domains of GroEL are colored in grey (equatorial domain), blue (intermediate domain), and cyan (apical domain), respectively. GroES is colored in lime. The left panel shows the overall view of the “football” complex as a ribbon diagram. The middle panel shows the cross-section of the “MT-football” complex in surface representation (along a plane containing the axis of 7-fold symmetry). The right panel shows the cross-section of the “SP-football” complex containing encapsulated Rubisco. Note that no electron density attributable to Rubisco was apparent in the central cavity even though biochemical evidence indicates the presence of Rubisco [33]. Electron densities are shown as grey mesh. All 2Fo-Fc electron density maps in this manuscript are contoured at 1σ .

	MT-Football (PDB ID: 4PKO)	SP-Football (PDB ID: 4PKN)
Space group	P 21 21 21	P 21 21 21
Unit cell a, b, c (Å)	169.79/174.49/410.16	171.95/173.65/411.27
Resolution range (Å)	90.9-3.84 (3.91-3.84)	122.19-3.66 (3.86-3.66)
Rmerge (%)	2.9 (70)	18.6 (65.5)
Total number of reflections	842530 (41896)	826508 (125226)
Number of unique reflections	116895 (5711)	135754 (19644)
Redundancy	7.2 (7.3)	6.1 (6.4)
Completeness (%)	100 (100)	98.9 (99.2)
I/ σ (I)	9.2 (2.8)	6.9 (2.34)
R (%)	18.30 (25.18)	18.80 (26.38)
Rfree (%)	24.80 (30.85)	24.10 (31.69)
RMSD		
bond lengths (Å)	0.008	0.006
bond angles (degree)	1.1	1.207
Ramachandran plot		
Most favored (%)	96	95
outlier(%)	0.3	0.3
Average B factors (Å²)	160	106

Table. 5-1 Data collection and refinement statistics of MT-football and SP-football.

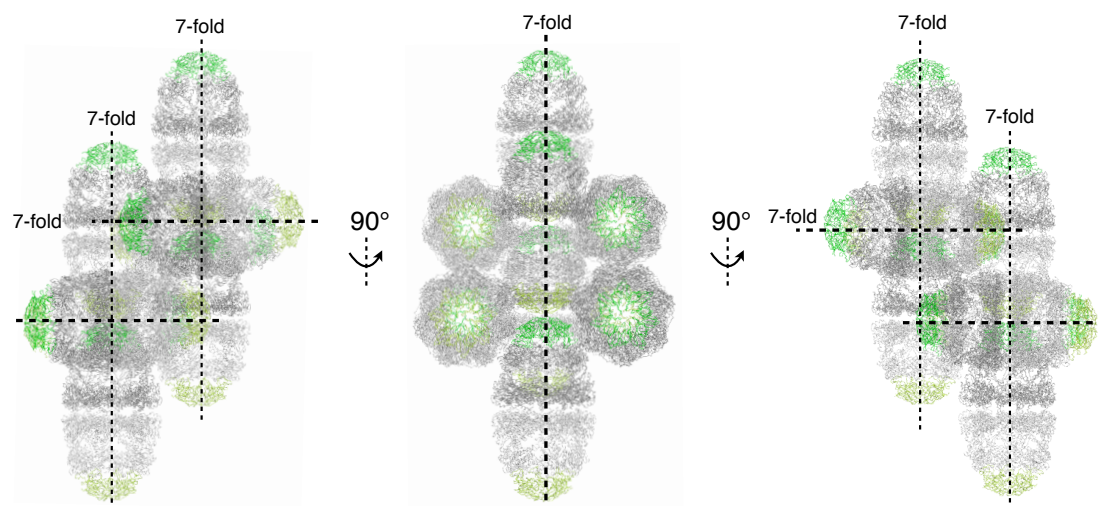


Fig. 5-2 Packing of the “football” complex in the crystal.

Both of the “MT-football” and “SP-football” are packed in the same way in the crystals. GroEL is shown in grey and GroES in green/lime.

Our biochemical data indicate that the SP football contains Rubisco in both cavities [33]. Given the time for crystal formation it is likely that the Rubisco monomer in the cavity will have assumed a native-like, folded state. However, the encapsulated Rubisco is not visible on the electron density map and does not significantly distort the GroEL/GroES structure overall (rmsd = 0.74; Fig. 5-3). Here we are dealing with an asymmetric object (Rubisco) in a container of near seven-fold symmetry (GroEL). So the conformation and orientation of encapsulated Rubisco varies between different unit cells in the crystal. For the same reason, an SP-containing bullet complex is also devoid of electron density attributable to encapsulated SP [74].

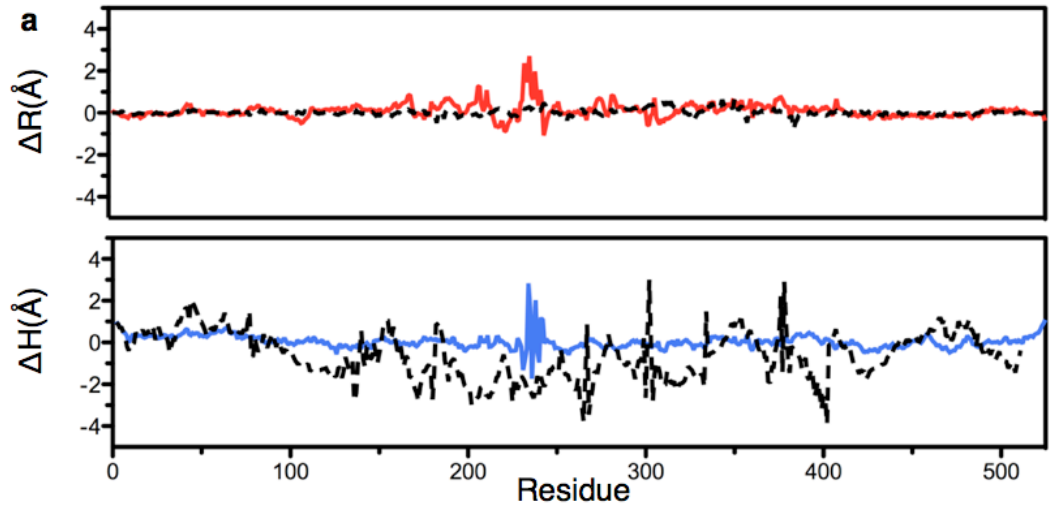


Fig. 5-3 Quantitative analysis of the conformational differences between “MT-football” and “SP-football”.

$\Delta R = \langle R \rangle - \langle R' \rangle$, where R is the distance from $C\alpha$ to the seven-fold axis in the “MT-football” complex, R' is the corresponding distance in the “SP-football”

complex. $\langle \rangle$ denotes average over fourteen subunits. Similarly, $\Delta H = \langle H \rangle - \langle H' \rangle$, where H is the distance from $C\alpha$ to the twofold axis between rings in the “MT-football” complex, H' is the corresponding distance in the “SP-football” complex. $\langle \rangle$ denotes average over fourteen subunits. ΔR and ΔH between two T state structures (PDB ID: 1XCK and 2NWC) serves as negative control, showing the level of structural “noise”.

5.3.2 Structural Plasticity of the Football Complexes.

Although both MT footballs and SP footballs seem symmetric, closer inspection reveals that the apical domains of GroEL and the GroES are not truly seven-fold symmetric, but rather pseudosymmetric. The θ plots (Fig. 5-4) show that the GroEL apical domains and GroES in both footballs deviate from perfect seven-fold symmetry by up to 10° . We attribute this asymmetry in these football complexes to their intrinsic flexibility, similar to but of a smaller magnitude than the asymmetry in the **R** state GroEL-ADP₁₄ [39].

However, the asymmetry in the football complexes is not identical to the asymmetry we observed in the **R** state. First, the overall degree of asymmetry decreases by 70% in the football complexes compared with the **R** state (Fig. 5-4B, C). Second, rather than being distributed throughout the entire apical domain, the asymmetry in the football complexes is restricted to several regions of the apical domain. This suggests that when GroES binds to the **R** state GroEL it gathers GroEL’s flexible apical domains together, making GroEL more rigid.

Finally, greater asymmetry occurs at the solvent-exposed residues facing the inside of the cavity, indicating that the encapsulated SP is surrounded with a plastic chamber rather than a rigid cage (Fig. 5-4C and Fig. 5-5). Such plasticity might allow GroEL to closely interact with SPs of different sizes and shapes during encapsulation. In this regard we note that the average B-factor of apical domains in the SP football is lower than that of the MT football (147 \AA^2 vs. 225 \AA^2), which could be caused by SP-GroEL interactions (Fig. 5-6).

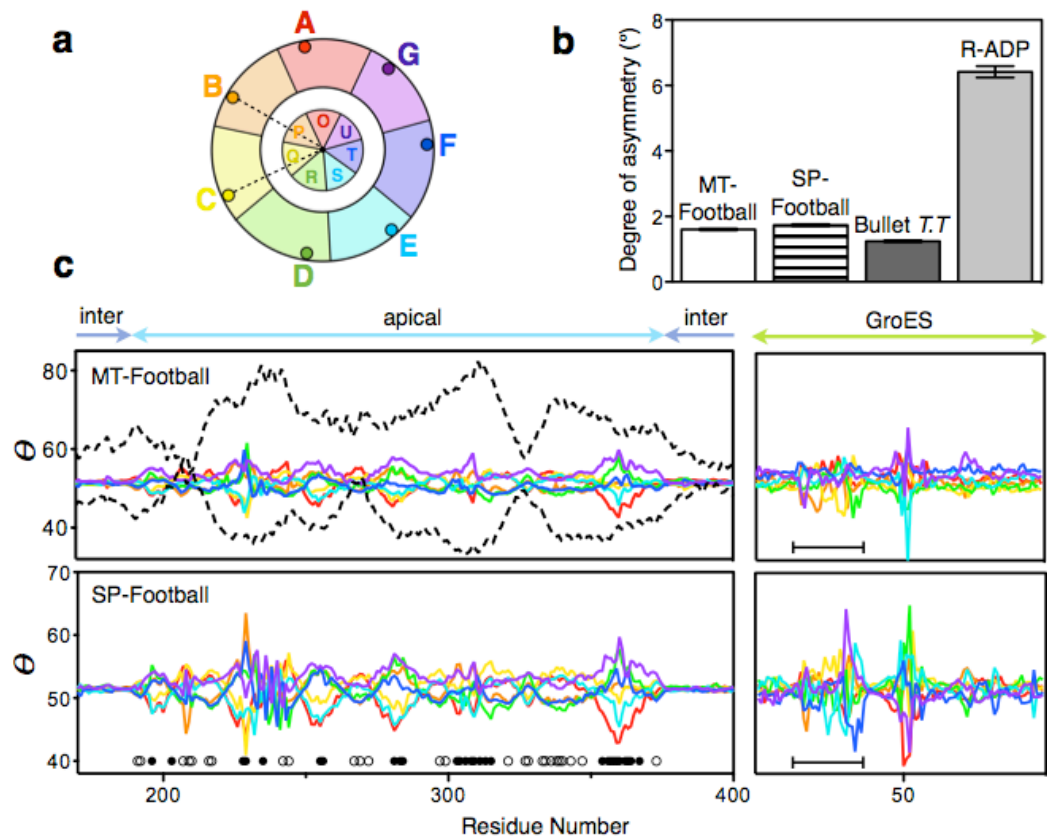


Fig. 5-4 Asymmetry in the “football” complexes.

(a) Definition of asymmetry probe θ , as previously described in chapter 2. **(b)** A histogram showing the average deviation from perfect 7-fold symmetry (θ)

$=360^\circ/7=51.4^\circ$) for “SP-football”, “MT-football”, the cis-ring of *T. Thermophilus* GroEL:GroES₁ (7) and the **R-ADP** structure of a GroEL^{D83A/R197A} mutant devoid of the salt-bridges which break during the **T** to **R** allosteric transition in the normal chaperonin cycle [37, 41]. **(c)** Quantitative θ plots showing that the apical domains and GroES mobile loops (black bars) of both the “MT-football” and the “SP-football” deviate from 7-fold symmetry. Subunit colors are as the same as in (a). Also shown (dashed lines) are θ plots of the two most asymmetric subunits of the **R-ADP** structure. In both “football” complexes, greater deviations from symmetry occur in solvent-exposed residues of the central cavity (black circles) than in exterior solvent exposed residues (white circles).

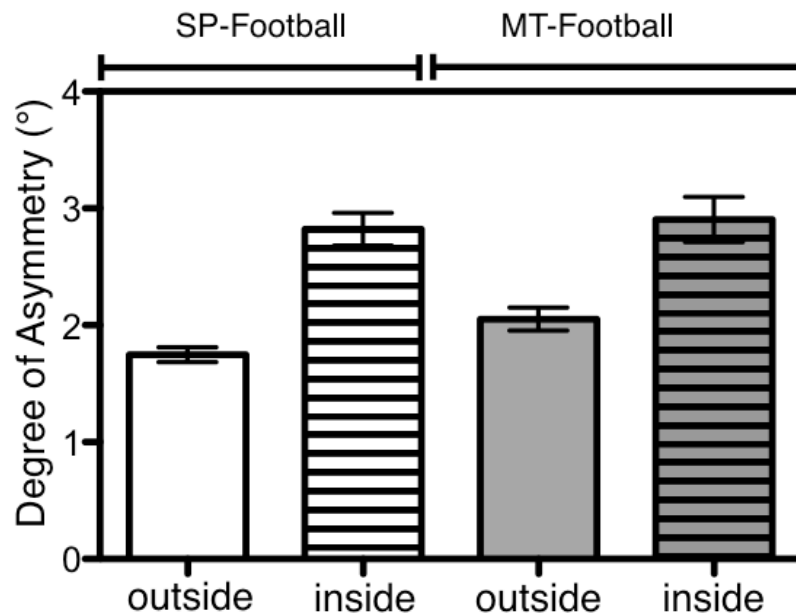


Fig. 5-5 Solvent-exposed residues inside the central chamber are more asymmetric than those outside the chamber.

Histograms showing the average deviations from perfect 7-fold symmetry ($\theta = 360^\circ/7=51.4^\circ$) for residues exposed to the inside and outside the central chambers of “SP-football” and “MT-football”. The error bars show SEM.

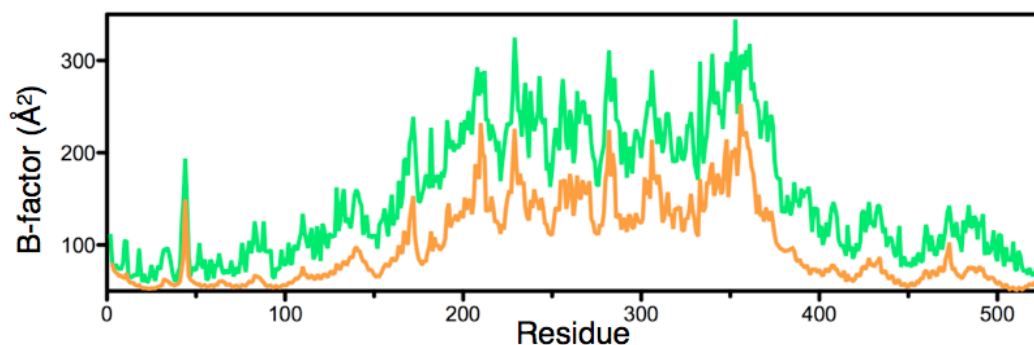


Fig. 5-6 Averaged B factors of the C α of each residue in the “MT-football”(green) and “SP-football”(orange) complexes.

None of the apical domains in the football complexes have identical conformations and the interaction at the GroEL/GroES interfaces was heterogeneous. Indeed, no two GroES “mobile loops” have the same conformation, and each GroEL/GroES interface is maintained by a unique set of hydrogen bonds (Fig. 5-7). This is quite different from the GroEL/GroES interface in the asymmetric bullet complex (PDB ID: 1AON), where, owing to imposed symmetry, all seven GroEL/GroES interfaces consist of the same hydrogen bonds between residues from helix I and the mobile loop of GroES (Fig. 5-7, red).

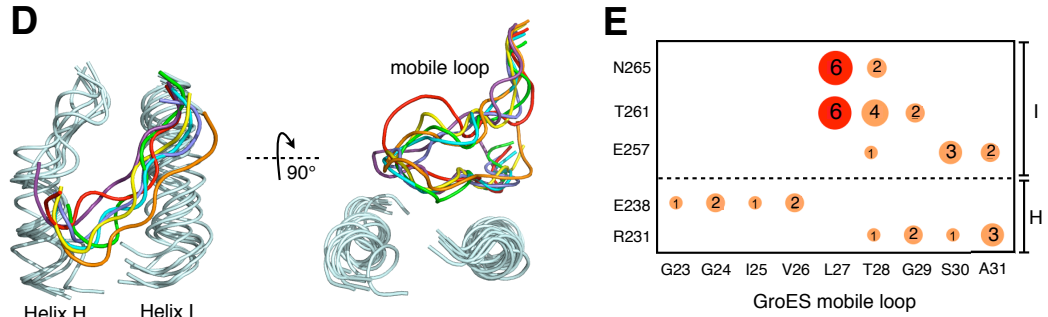


Fig. 5-7 GroEL-GroES interfaces in the “football” complexes.

(d) Interactions between GroEL subunit and GroES subunit are heterogeneous.

(e) Hydrogen bonds stabilize GroEL-GroES interfaces are represented by circles.

Area of circles represents the relative occurrence of hydrogen bonds. Two hydrogen bonds presented in the “bullet” complex are colored in red.

The nonspecific interactions between GroEL and GroES mobile loops could be important for GroES binding. Before the binding of GroES, both the GroEL apical domains, including helices H and I, and the mobile loops of GroES are extremely flexible [14, 27, 39]. Having a large number of nonspecific interactions in the GroEL/GroES interface would permit the engagement of these flexible partners with fewer entropic costs that highly specific interactions would entail. For example, T28 on the mobile loop interacts through hydrogen bonds with up to four residues in helices H and I (Fig. 5-7), which increases the probability of GroES capture.

5.3.3 Conformational Changes at the Inter-Ring Interface. Although the football complex encapsulates two SP molecules simultaneously, the two GroEL/GroES

cavities do not act independently. Biochemical evidence has shown that the communication between two GroEL rings is crucial for GroES release [66]. When such inter-ring communication is disrupted, as it is in the single-ringed version SR1, GroEL fails to release GroES, SP, and ADP, resulting in a “dead-end” complex [75].

The crystal structure of the football complex at ~ 3.7 Å resolution allows us to analyze the inter-ring communication during chaperonin’s natural catalytic cycle. The two GroEL rings in the football complex communicate through the same two inter-ring interfacial sites, L and R, as previously reported (Fig. 5-8, Inset) [38].

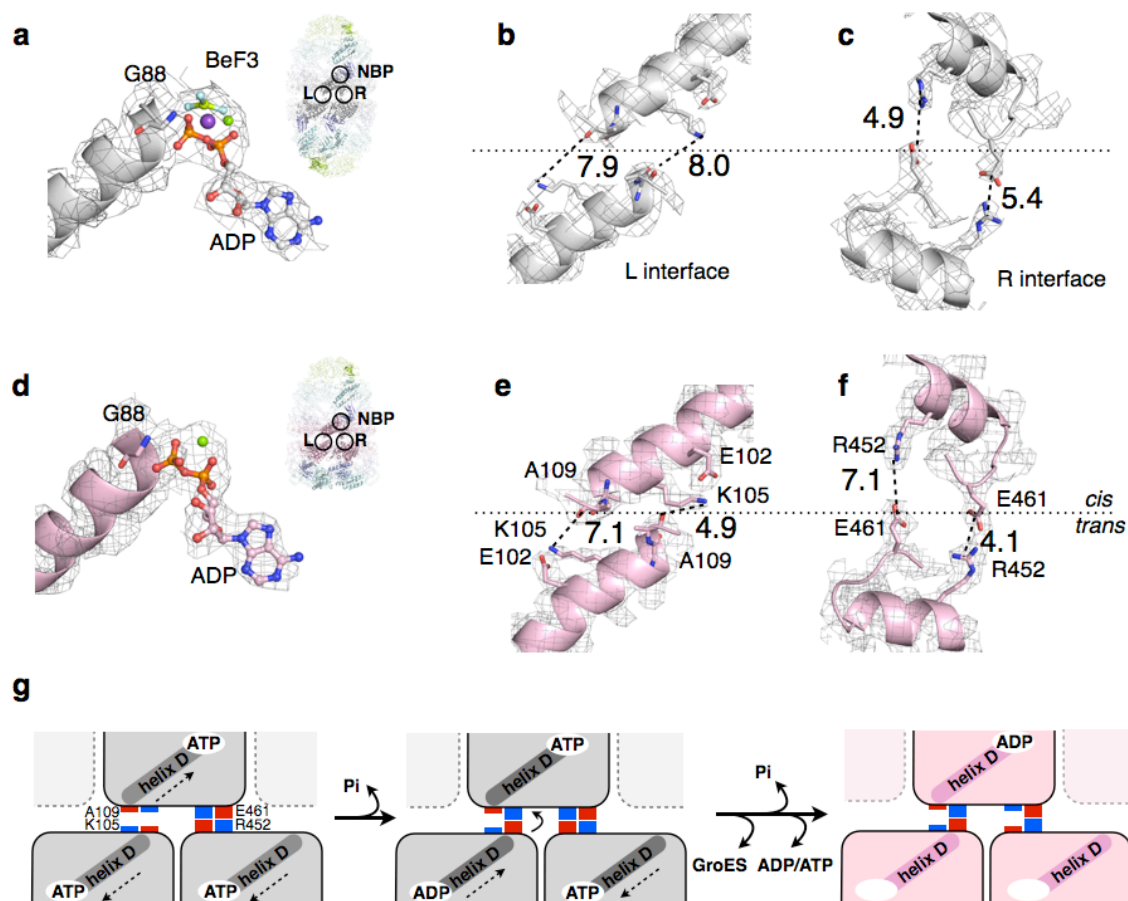


Fig. 5-8 Switching of electrostatic interactions at the inter-ring interface.

(a) Interaction between ADP-BeF₃ and the N terminus of helix D in the “football” complex. The relative position of inter-ring interfaces L, R, and the nucleotide-binding pocket (NBP) is shown in the inset. **(b)** and **(c)** are close views of the inter-ring interfaces L and R in the “football” complex. The 2-fold axis of symmetry is shown as a long dashed line. **(d)** Interaction between ADP and the N termini of helix D in the “bullet” complex. The relative position of inter-ring interfaces L, R, and the nucleotide-binding pocket (NBP) is shown in the inset. **(e)** and **(f)** are the same as (b) and (c), except showing the inter-ring interface L and R in the “bullet” complex. **(g)** A structure based mechanism for

sensing ATP hydrolysis. The dashed arrows indicate movement of helix D in response to ATP binding and hydrolysis.

The L interface involves interactions between helix D of one subunit with the same helix in the opposite ring. The axes of the two helices D are anti-parallel to one another. The C termini of each helix (A109) are slightly offset from one another, across the twofold axis of symmetry (Fig. 5-8 B, E). The R interface involves interactions between helix P of one subunit and the same helix in the opposite ring. The axes of the two helices P are nearly anti-parallel to one another and also to the twofold axis of symmetry (Fig. 5-8 C, F).

When GroEL/GroES switches from the football complex to the bullet complex, the total contact surface area of L and R interfaces increases only slightly (from 2,232 Å² to 2,464 Å²). However, close inspection shows the relative size of the two interfaces has changed. The L interface expands by ~50% (from 764 Å² to 1,132 Å²), whereas the R interface shrinks by ~25% (from 1,472 Å² to 1,132 Å²). This change in the inter-ring interface is caused by a reduction in the radius (up to 6 Å) of the equatorial plate of the trans ring (Fig. 5-9A) plus a slight rotation (up to 8°) of the trans GroEL ring relative to the cis GroEL ring (Fig. 5-9B).

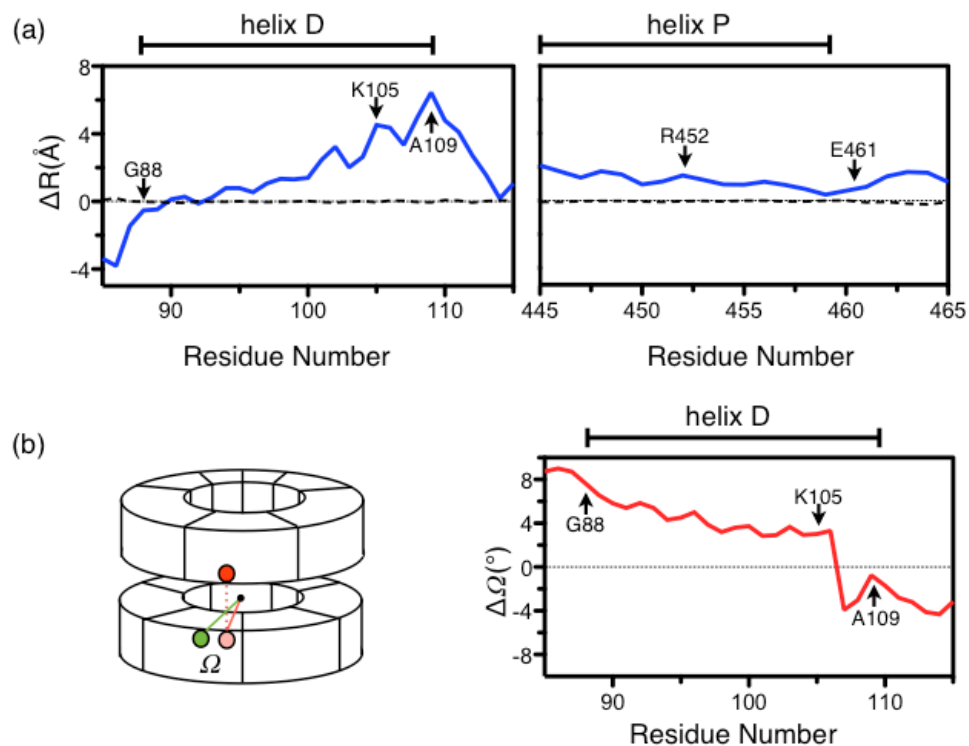


Fig. 5-9 Quantitative analysis of the conformational changes at the inter-ring interface during the “football” to the “bullet” transition.

(a) We use a cylindrical coordinate system to analyze the changes in the inter-ring interface, as previously described in chapter 2. $\Delta R = \langle R \rangle - \langle R' \rangle$, where R is the distance to the seven-fold axis in the “football” complex, R' is the corresponding distance in the “bullet” complex. $\langle \rangle$ denotes average over seven subunits. **(b)** Definition of Ω , the quantity used to measure ring to ring rotation. Ω is the angle between two vectors, one from the $C\alpha$ of residue i in subunit j , to the seven-fold axis of symmetry; another from the $C\alpha$ of residue i to the seven-fold axis of symmetry, in the subunit form L interface with subunit j , from the opposite ring. $\Delta \Omega = \langle \Omega \rangle - \langle \Omega' \rangle$, where Ω is the ring-to-ring angle in the

“football” complex, Ω' is the corresponding angle in the “bullet” complex. $\langle \rangle$ denotes average over seven L interfaces.

We further analyzed the change in electrostatic interactions at both L and R interfaces. Accompanying the release of one GroES, the two interacting D helices at the L interface move closer and twist to established one electrostatic interaction between e-amino group of K105 of the cis ring and helix dipole-induced charge on the carbonyl oxygen of A109 of the trans ring. (Fig. 5-9 B, E). At the R interface, the two helices P move apart and twist so one of the two salt bridges between E461 of the trans ring and R452 of the cis ring breaks (Fig. 5-9 C, F).

5.3.4 Structural Basis for Inter-Ring Communication.

During the chaperonin’s natural catalytic cycle, the dissociation of GroES from the football complex requires the hydrolysis of ATP and the development of nucleotide asymmetry (i.e., the difference in the number of ATPs hydrolyzed between two GroEL rings) [18, 19]. We propose that helix D senses and transmits the signal of ATP hydrolysis and ATP asymmetry by exploiting the helix dipole that is positively charged at the N-terminal G88 and negatively charged at the C-terminal A109 (Fig. 5-8 A, D).

Before the football complex hydrolyzes ATP and releases GroES, the N terminus of helix D (G88) in the two GroEL subunits from opposite rings both form

electrostatic interactions with the ATP γ -phosphate (Fig. 5-8A, G). The interaction between γ -phosphate and the N terminus of the two helices D draws the two helices apart from one another (arrows in Fig. 5-8G). Once ATP in one subunit is hydrolyzed and the γ -phosphate is released, helix D in that subunit moves closer to helix D in the other ring and a cross-ring electrostatic interaction forms between A109 and K105 (Fig. 5-8G). As more ATP is hydrolyzed and a critical number of cross-ring A109-K105 interactions is reached, one or the other GroES departs [76].

The mechanism of inter-ring communication in the football complex is quite different from a previously proposed model that is based on the comparison between the interfaces of apo-GroEL and the bullet complex [77]. However, our current understanding of the chaperonin cycle (Fig. 1-2, Fig. 1-3) assigns no role whatsoever to apo-GroEL. Exactly how the events at the equatorial plate are transmitted to the apical domains leading to the dissociation of GroES is yet to be determined, however.

5.4 Summary

We report structures of two such “football” complexes to ~ 3.7 Å resolution; one is empty whereas the other contains encapsulated SP in both chambers.

Compared with the bullet-shaped GroEL:GroES₁ complex, the GroEL:GroES₂ football complex differs conformationally at the GroEL-GroES interface and also at the interface between the two GroEL rings. We propose that the electrostatic

interactions between the ϵ -NH₃⁺ of K105 of helix D in one ring with the negatively charged carboxyl oxygen of A109 at the carboxyl end of helix D of the other ring provide the structural basis for negative inter-ring cooperativity.

Chapter 6: Summary and final discussion

6.1 What happens during SP capture?

In chapter 3 and chapter 4 of this thesis I tried to answer two questions related to SP capture. First, how does GroEL capture SPs, which are different in sequence and structure? Second, how does SP promote the release of ADP?

Previous experiments have shown that without SP, GroEL idles with ADP bound to the trans-ring. In that case the turnover of ATP is limited by the slow release of ADP so that ATP will not be consumed when there is no “work” for GroEL to do. In the presence of SP, the release of ADP is accelerated by more than 100 fold so GroEL turns over both ATP and SP rapidly [17,19]. However without structural information, it is not clear how SP accelerates the release of ADP from previous cycle. It is also not clear if the ADP bound trans-ring has other roles in addition to limiting ATP turnover.

In order to answer these two questions we determined two crystal structures: GroEL^{D83A/R197A}-ADP₁₄ and GroEL^{D83A/R197A}. We found that upon ADP binding, two out of three domains of GroEL, the intermediate domains and the apical domains deviate from seven-fold symmetry. The deviation from symmetry is caused by the loss of inter-subunit interactions and increased domain flexibility.

Flexible apical domains allow GroEL to capture SPs of different sizes and shapes. Such plasticity might be necessary for GroEL to capture SP in solution.

After SP is captured, it catalyzes ADP release by favoring the **T** state and holding the lid of nucleotide binding pocket of the flexible intermediate domain in the open position.

6.2 What happens during SP encapsulation?

In chapter 5 I tried to answer a few questions related to SP encapsulation. In each chapronin cycle, the “football” complex briefly encapsulates SP. What happens during encapsulation is quite mysterious. Are the encapsulation cavities in the “football” complex different from the single cavity in the “bullet” complex? Does the cavity interact with encapsulated SP and actively promote SP folding? And finally how is encapsulation timed by ATP hydrolysis?

To answer these questions, we determined two crystal structures of the “football” complexes. One contains no SP, another encapsulates one misfolded Rubisco in each cavity. The two cavities of the “football” complex are very similar to the cavity of the “bullet” complex, except the mobile loops of GroES in the “football” have heterogeneous GroEL contacts. Solvent accessible residues in the cavity are flexible and could be stabilized upon SP encapsulation.

Based on the “football” structures and other biochemical evidence, we proposed that the duration of encapsulation is determined by nucleotide asymmetry. The signal of ATP hydrolysis and release of phosphate is sensed by a helix dipole and transmitted through inter-ring electrostatic interactions between K105 and A109.

6.3 Final discussion

GroEL is the first molecular chaperone discovered, which helps misfolded proteins regain their native structures and functions [15]. After years of studies, each step in GroEL/ES assisted protein folding is becoming clear. The flexible apical domains of GroEL capture heterogeneous SPs and SP promotes ADP to ATP exchange in GroEL's nucleotide binding pockets. ATP binding unfolds SP and helps SP to overcome the energy barrier between misfolded state and native state. Partially unfolded SP is encapsulated in a closed GroEL/GroES chamber. At saturating SP concentration, the “football” complexes with two GroEL/GroES chambers and two encapsulated SPs is the predominate species. The timer of encapsulation is ATP hydrolysis. The signal of phosphate release is transmitted through inter-ring electrostatic interactions involving K105 and A109.

What is the most important step for SP folding? The answer is likely to depend on the SP. For stably misfolded SPs, ATP induced unfolding maybe the most important step; and for SPs with complicated native topology, encapsulation maybe crucial because it bring distant residues together. Thus the mechanism of chaperonin assisted protein folding very likely differs among substrates, and even for the same substrate in different misfolded conformations. Therefore the active turnover of ATP is crucial, it enables a misfolded protein to try different folding strategies every 2 seconds, with a different starting structure each time.

GroEL is very robust in helping different SP to reach their respective native states [35, 78].

To prove this model, it is necessary to probe the conformation of different SPs during GroEL/ES assisted folding. Given the conformational heterogeneity of misfolded SP, we have shown that it is impossible to study GroEL-SP interaction using X-ray crystallography [33]. High-resolution cryo-EM will be very helpful in studying such heterogeneous ensembles [32]. Another advantage of cryo-EM is that the same sample containing SP, GroEL, GroES and ATP may be frozen at different times. This allows us to analyze the conformation of SP at different steps of assisted folding such as during capture and encapsulation, at realistic time scale.

Reference

1. Crick F (1970) Central dogma of molecular biology. *Nature* 227 (5258):561-3
2. Anfinsen C (1972) The formation and stabilization of protein structure. *Biochem. J.* 128 (4):737-49
3. Selkoe DJ (2003) Folding proteins in fatal ways. *Nature* 426 (6968): 900–904
4. Chiti F, Dobson C (2006). Protein misfolding, functional amyloid, and human disease. *Annual review of biochemistry* 75:333–366
5. Hashimoto M, Rockenstein E, Crews L and Masliah E (2003) Role of protein aggregation in mitochondrial dysfunction and neurodegeneration in Alzheimer's and Parkinson's diseases. *Neuromolecular Med.* 4 (1–2):21–36.
6. Turner PR, O'Connor K, Tate WP and Abraham WC (2003). Roles of amyloid precursor protein and its fragments in regulating neural activity, plasticity and memory. *Prog. Neurobiol.* 70 (1):1–32.
7. Brookmeyer R, Gray S and Kawas C (1998) Projections of Alzheimer's Disease in the United States and the Public Health Impact of Delaying Disease Onset. *American Journal of Public Health.* 88(9):1337–42
8. Brookmeyer R, Johnson E, Ziegler-Graham K and Arrighi H.M. (2007) Forecasting the global burden of Alzheimer's disease. *Alzheimer's & Dementia.* 3(3):186–91.
9. Goloubinoff P, Gatenby A and Lorimer GH (1989) GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose biphosphate carboxylase oligomers in *Escherichia coli*. *Nature* 337:44-47

10. Viitanen PV, Gatenby AA and Lorimer GH (1992) Purified chaperonin 60 (groEL) interacts with the nonnative states of a multitude of Escherichia coli proteins. *Protein Sci.* 1(3):363-9.
11. Braig K, Otwinowski Z, Hegde R, Boisvert DC, Joachimiak A, Horwich AL and Sigler PB (1994) The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature* 371:578-586
12. Mou J, Sheng S, Ho R and Shao Z (1996) Chaperonins GroEL and GroES: views from atomic force microscopy. *Biophys. J.* 71(4):2213–2221.
13. Chen L and Sigler PB (1999) The crystal structure of a GroEL/peptide complex: plasticity as a basis for substrate diversity. *Cell.* 99(7):757-68.
14. Hunt JF, Weaver AJ, Landry SJ, Gierasch L and Deisenhofer J (1996) The crystal structure of the GroES co-chaperonin at 2.8 Å resolution. *Nature.* 379(6560):37-45.
15. Goloubinoff P, Christeller JT, Gatenby AA and Lorimer GH (1989) Reconstitution of active dimeric ribulose biphosphate carboxylase from an unfolded state depends on two chaperonin proteins and Mg-ATP. *Nature.* 342(6252):884-9.
16. Lorimer GH (1997) Protein folding. Folding with a two-stroke motor. *Nature.* 388(6644):720-1, 723.
17. Grason JP, Gresham JS and Lorimer GH. (2008) Setting the chaperonin timer: a two-stroke, two-speed, protein machine. *Proc. Natl. Acad. Sci. U S A.* 105(45):17339-44.

18. Yang D, Ye X and Lorimer GH. (2013) Symmetric GroEL:GroES2 complexes are the protein-folding functional form of the chaperonin nanomachine. *Proc. Natl. Acad. Sci. U S A.* 110(46):E4298-305
19. Ye X and Lorimer GH. (2013) Substrate protein switches GroE chaperonins from asymmetric to symmetric cycling by catalyzing nucleotide exchange. *Proc. Natl. Acad. Sci. U S A.* 110(46):E4289-97.
20. Stan G, Brooks BR, Lorimer GH and Thirumalai D. (2006) Residues in substrate proteins that interact with GroEL in the capture process are buried in the native state. *Proc. Natl. Acad. Sci. U S A.* 103(12):4433-8.
21. Stan G, Brooks BR, Lorimer GH and Thirumalai D. (2005) Identifying natural substrates for chaperonins using a sequence-based approach. *Protein Sci.* 14(1):193-201.
22. Farr GW, Furtak K, Rowland MB, Ranson NA, Saibil HR, Kirchhausen T and Horwich AL. (2000) Multivalent binding of nonnative substrate proteins by the chaperonin GroEL. *Cell.* 100(5):561-73.
23. Ranson NA, Burston SG and Clarke AR. (1997) Binding, encapsulation and ejection: substrate dynamics during a chaperonin-assisted folding reaction. *J. Mol. Biol.* 266(4):656-64.
24. Walter S, Lorimer GH and Schmid FX. (1996) A thermodynamic coupling mechanism for GroEL-mediated unfolding. *Proc. Natl. Acad. Sci. U S A.* 93(18):9425–9430.

25. Sharma S, Chakraborty K, Müller BK, Astola N, Tang YC, Lamb DC, Hayer-Hartl M and Hartl FU. (2008) Monitoring protein conformation along the pathway of chaperonin-assisted folding. *Cell*. 133(1):142-53.
26. Lin Z, Madan D and Rye HS. (2008) GroEL stimulates protein folding through forced unfolding. *Nat. Struct. Mol. Biol.* 15(3):303-11.
27. Clare DK, Vasishtan D, Stagg S, Quispe J, Farr GW, Topf M, Horwich AL and Saibil HR. (2012) ATP-triggered conformational changes delineate substrate-binding and -folding mechanics of the GroEL chaperonin. *Cell*. 149(1):113-23.
28. Yokokawa M, Wada C, Ando T, Sakai N, Yagi A, Yoshimura SH and Takeyasu K. (2006) Fast-scanning atomic force microscopy reveals the ATP/ADP-dependent conformational changes of GroEL. *EMBO J*. 25(19):4567-76.
29. Corsepius NC and Lorimer GH. (2013) Measuring how much work the chaperone GroEL can do. *Proc. Natl. Acad. Sci. U S A*. 110(27):E2451-9
30. Chakraborty K, Chatila M, Sinha J, Shi Q, Poschner BC, Sikor M, Jiang G, Lamb DC, Hartl FU and Hayer-Hartl M. (2010) Chaperonin-catalyzed rescue of kinetically trapped states in protein folding. *Cell*. 142(1):112-22.
31. Ellis RJ (1996) Revisiting the Anfinsen cage. *Fold. Des.* 1:R9–R15.
32. Chen DH, Madan D, Weaver J, Lin Z, Schröder GF, Chiu W and Rye HS. (2013) Visualizing GroEL/ES in the act of encapsulating a folding protein. *Cell*. 153(6):1354-65.

33. Fei X, Ye X, LaRonde NA and Lorimer GH. (2014) Formation and structures of GroEL:GroES₂ chaperonin footballs, the protein-folding functional form. *Proc. Natl. Acad. Sci. U S A.* pii: 201412922. [Epub ahead of print]
34. Thirumalai D and Lorimer GH. (2001) Chaperonin-mediated protein folding. *Annu. Rev. Biophys. Biomol. Struct.* 30:245-69.
35. Weissman JS, Rye HS, Fenton WA, Beechem JM and Horwich AL (1996) Characterization of the active intermediate of a GroEL-GroES-mediated protein folding reaction. *Cell* 84: 481-490.
36. Horovitz A, Fridmann Y, Kafri G and Yifrach O (2001) Review: allostery in chaperonins. *J. Struct. Biol.* 135(2):104-14.
37. Yifrach O and Horovitz A (1995) Nested cooperativity in the ATPase activity of the oligomeric chaperonin GroEL. *Biochemistry* 34(16):5303-8.
38. Xu Z, Horwich AL and Sigler PB. (1997) The crystal structure of the asymmetric GroEL-GroES-(ADP)₇ chaperonin complex. *Nature.* 388(6644):741-50.
39. Fei X, Yang D, LaRonde-LeBlanc N and Lorimer GH. (2013) Crystal structure of a GroEL-ADP complex in the relaxed allosteric state at 2.7 Å resolution. *Proc. Natl. Acad. Sci. U S A.* 110(32):E2958-66.
40. Murai N, Makino Y and Yoshida M. (1996) GroEL locked in a closed conformation by an interdomain cross-link can bind ATP and polypeptide but cannot process further reaction steps. *J. Biol. Chem.* 271(45):28229-34.

41. Nojima T and Yoshida M. (2009) Probing open conformation of GroEL rings by cross-linking reveals single and double open ring structures of GroEL in ADP and ATP. *J. Biol. Chem.* 284(34):22834-9.
42. Chaudhry C, Farr GW, Todd MJ, Rye HS, Brunger AT, Adams PD, Horwich AL and Sigler PB. (2003) Role of the gamma-phosphate of ATP in triggering protein folding by GroEL-GroES: function, structure and energetics. *Embo J.* 22:4877-4887
43. Yang D (2014) PhD thesis (Univ. of Maryland, College Park, MD)
44. Schmidt M. et al. (1994) Symmetric complexes of GroE chaperonins as part of the functional cycle. *Science* 265:656–659
45. Azem A. et al. (1995) The protein-folding activity of chaperonins correlates with the symmetric GroEL₁₄(GroES₇)₂ heterooligomer. *Proc. Natl. Acad. Sci. U S A* 92:12021–12025.
46. Llorca O, Marco S, Carrascosa JL and Valpuesta JM. (1997) Symmetric GroEL-GroES complexes can contain substrate simultaneously in both GroEL rings. *FEBS Lett.* 405:195–199.
47. Sparrer H, Rutkat K and Buchner J. (1997) Catalysis of protein folding by symmetric chaperone complexes. *Proc. Natl. Acad. Sci. U S A* 94:1096–1100.
48. Taguchi H. et al. (2004) BeF(x) stops the chaperonin cycle of GroEL-GroES and generates a complex with double folding chambers. *J. Biol. Chem.* 279:45737–45743.
49. Sameshima T. et al. (2008) Football- and bullet-shaped GroEL-GroES

- complexes coexist during the reaction cycle. *J. Biol. Chem.* 283:23765-23773.
50. Takei Y, Iizuka R, Ueno T and Funatsu T. (2012) Single-molecule observation of protein folding in symmetric GroEL-(GroES)₂ complexes. *J. Biol. Chem.* 287:41118–41125.
 51. Grason JP, Gresham JS, Widjaja L, Wehri SC and Lorimer GH. (2008) Setting the chaperonin timer: The effects of K⁺ and substrate protein on ATP hydrolysis. *Proc. Natl. Acad. Sci. U S A* 105:17334-8.
 52. Martin ACR (1982). Rapid Comparison of Protein Structures. *Acta. Crystallogr.,Sect.A* 38 (6):871-873.
 53. Bartolucci C, Lamba D, Grazulis S, Manakova E and Heumann H. (2005) Crystal structure of wild-type chaperonin GroEL. *J. Mol. Biol.* 354:940-951
 54. Kiser PD, Lodowski DT and Palczewski K. (2007) Purification, crystallization and structure determination of native GroEL from *Escherichia coli* lacking bound potassium ions. *Acta. Crystallogr.,Sect.F* 63:457-461
 55. Lin Z and Rye HS (2006) GroEL-mediated protein folding: making the impossible, possible. *Crit. Rev. Biochem. Mol. Biol.* 41(4):211-39.
 56. Yifrach O and Horovitz A (1996) Allosteric control by ATP of non-folded protein binding to GroEL. *J. Mol. Biol.* 255(3):356-61.

57. Yifrach O and Horovitz A (1994) Two lines of allosteric communication in the oligomeric chaperonin GroEL are revealed by the single mutation Arg 196->Ala. *J. Mol. Biol.* 243(3):397-401.
58. White HE, Chen S, Roseman AM, Yifrach O, Horovitz A and Saibil HR. (1997) Structural basis of allosteric changes in the GroEL mutant Arg 197->Ala. *Nat. Struct. Biol* 4(9):690-694.
59. Grason JP (2004) PhD thesis Allostery in GroEL: its role in the refolding of protein substrate. (Univ. of Maryland, College Park, MD)
60. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macro- molecular structure solution. *Acta. Crystallogr. D Biol. Crystallogr.* 66:213-21.
61. Shimamura T, Koike-Takeshita A, Yokoyama K, Masui R, Murai N, Yoshida M, Taguchi H and Iwata S (2004) Crystal structure of the native chaperonin complex from *Thermus thermophilus* revealed unexpected asymmetry at the cis-cavity. *Structure* 12(8):1471-1480.
62. Danziger O, Rivenzon-Segal D, Wolf SG and Horovitz A. (2003) Conversion of the allosteric transition of GroEL from concerted to sequential by the single mutation Asp-155→ Ala. *Proc. Natl. Acad. Sci. U S A.* 100(24): 13797-802
63. Hyeon C, Lorimer GH and Thirumalai D (2006) Dynamics of allosteric transitions in GroEL. *Proc. Natl. Acad. Sci. U S A.* 103(50):18939-44.

64. Wang J and Boisvert DC (2003) Structural basis for GroEL-assisted protein folding from the crystal structure of (GroEL-KMgATP)₁₄ at 2.0Å resolution. *J. Mol. Biol.* 327(4):843-55.
65. Frank GA, Goomanovsky M, Davidi A, Ziv G, Horovitz A and Haran G (2010) Out-of-equilibrium conformational cycling of GroEL under saturating ATP concentrations. *Proc. Natl. Acad. Sci. U S A.* 107(14):6270-4.
66. Todd MJ, Viitanen PV and Lorimer GH (1994) Dynamics of the chaperonin ATPase cycle: implications for facilitated protein folding. *Science* 265(5172):659-66 (1994).
67. Kalisman N, Schröder GF and Levitt M. (2013) The crystal structure of the eukaryotic chaperonin CCT reveal its functional partitioning. *Structure* 21(4): 540-9.
68. Battye TGG, Kontogiannis L, Johnson O, Powell HR and Leslie AGW. (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta. Crystallogr. D. Biol. Crystallogr.* 67:271-81.
69. Barlow DJ and Thornton JM. (1983) Ion-pairs in proteins. *J. Mol. Biol.* 168(4):867-85.
70. Kumar S and Nussinov R. (2002) Relationship between Ion Pair Geometries and Electrostatic Strengths in Proteins. *Biophys. J.* 83(3):1595-1612.
71. Kumar S and Nussinov R. (1999) Salt Bridge Stability in monomeric proteins. *J. Mol. Biol.* 293(5):1241-55.

72. Tehver R and Thirumalai D (2008) Kinetic Model for the Coupling between Allosteric Transitions in GroEL and Substrate Protein Folding and Aggregation. *J. Mol. Biol.* 4, 1279-1295.
73. Krissinel E and Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372:774-797.
74. Shimamura T. et al. (2004) Crystal Structure of the Native Chaperonin Complex from *Thermus thermophilus* Revealed Unexpected Asymmetry at the cis-Cavity. *Structure* 12:1471-1480.
75. Horwich AL, Burston SG, Rye HS, Weissman JS and Fenton WA. (1998) Construction of single-ring and two-ring hybrid versions of bacterial chaperonin GroEL. *Methods Enzymol.* 290:141-146.
76. Takei Y, Iizuka R, Ueno T and Funatsu T (2012) Single-molecule observation of protein folding in symmetric GroEL-(GroES)₂ complexes. *J. Biol. Chem.* 287(49):41118-41125.
77. Horwich AL (2011) Protein folding in the cell: An inside story. *Nat. Med.* 17(10):1211-1216.
78. Todd MJ, Lorimer GH and Thirumalai D (1996) Chaperonin-facilitated protein folding: optimization of rate and yield by an iterative annealing mechanism. *Proc. Natl. Acad. Sci. U S A.* 93(9):4030-4035.