The Crystal Structures of the Eukaryotic Chaperonin CCT Reveal Its Functional Partitioning

Nir Kalisman, Gunnar F. Schröder, and Michael Levitt

INTRODUCTION

Group II chaperonins are large nanomachines that are central to protein folding in both eukaryotes and archaea. The overall structure is 16 nm in diameter and comprises two stacked rings of eight subunits each (Yeşenes et al., 2011). Through a cycle that is powered by ATP hydrolysis, these subunits rotate to open and close a central folding chamber. While the archaeal systems are often homo-oligomeric (Kapatai et al., 2006), the eukaryotic chaperonin has evolved (Archibald et al., 2001) to a hetero-oligomeric complex, where each ring is made of eight paralogous subunits (CCT1 to CCT8) with a mutual sequence identity of about 30%. This divergence to eight genes occurred very early in eukaryotic evolution and is highly conserved in all eukaryotic species. The hetero-oligomeric nature of CCT plays a key role in CCT function by allowing for intricate substrate-binding modes in the open form (Llorca et al., 2001; Spiess et al., 2006) and differential ATP hydrolysis in the closed form (Rivenzon-Segal et al., 2005; Amit et al., 2010). This functionality is tightly coupled to a precise arrangement of the eight subunits within each ring and to a fixed relative registration of the two rings (Martín-Benito et al., 2007).

Much of the structural knowledge on group II chaperonins comes from high-resolution crystallography of archaeal systems (Ditzel et al., 1998; Shomura et al., 2004). In contrast, structural studies of CCT either by crystallography (Dekker et al., 2011; Muñoz et al., 2011) or by cryo-electron microscopy (EM) (Martín-Benito et al., 2007; Cong et al., 2010) could not achieve the sufficient resolution to unambiguously identify the different subunits within the electron density. Specifically, the high sequence similarity between the subunits dictates a highly similar backbone trace for all subunits. This limitation can only be overcome by resolving subunit-specific side-chains. Unfortunately, none of these studies had side-chain densities of sufficient quality to make such direct inference. As a consequence, questions pertaining to the hetero-oligomeric nature of CCT have remained unresolved, the most basic of which was that of the subunit arrangement in the native complex. This question is difficult to answer, as there are eight factorial (8! = 40,320) possible arrangements of the subunits within a particle.

To resolve the subunit ambiguity, we present a different approach to low-resolution crystallography, where we abandon the conventional search for specific side-chain densities with the aid of a crystallographic browser. Instead, we collectively and automatically assess the fit of many side-chains of a specific model to the crystallographic data. This automation allows us to test millions of models and to determine objectively if one of them fits the data better than others. When applied to the two crystallographic data sets of CCT that have been published to date, this approach successfully singles out with very high confidence one model for each set. These models describe how to locate the different subunits within the crystal asymmetric units and allow us to fully solve the CCT structures. The resulting structures not only give crystallographic proof of the native subunit arrangement of CCT but also map various asymmetric features to specific subunits. We find that subunit CCT6-Z shows significant structural deviations from the other subunits. We also find that actin and tubulin, the main substrates of CCT, bind mainly around this subunit. The overall picture that emerges is that of a partitioned particle, where specific functions, such as substrate binding or ATP hydrolysis, cluster to specific sides.
RESULTS

Identifying Subunits in the Crystal Asymmetric Units

In this study, we reanalyze X-ray diffraction data of the closed form of yeast-CCT at 3.8 Å resolution (Dekker et al., 2011) and the open form of bovine-CCT at 5.5 Å resolution (Muñoz et al., 2011). Our first aim was to assign the correct subunit type to each position in the two crystal asymmetric units. This problem is harder than determining the subunit arrangement of the particle alone, because each arrangement can fit into the electron density in eight different orientations (Figures 1A and 1D). There are, therefore, 2,580,480 ( = 8! × 8 × 8) and 322,560 ( = 8! × 8) possible models for the asymmetric units of the closed and open data sets, respectively. Because of the large number of possible models and the low resolution of the data, we strictly adhered to an unbiased approach that did not use calculated phases. Thus, we exhaustively built all the possible all-atom models for both asymmetric units and then measured their fit to the crystallographic data by calculating their R values (Vaguine et al., 1999) without doing any refinement. While our fit measure is straightforward, much consideration went into the model building step. First, all our models for a particular form had exactly the same backbone without regard for the specific subunit arrangement or orientation; models differed only in the side-chain atoms added to this fixed backbone. Second, the side-chain rotamers were modeled by SCWRL4 (Krivov et al., 2009) on a backbone taken from an archaeal chaperonin (Shomura et al., 2004).

Without any prior information, the histogram of R values obtained for the closed form (Figure 1B) singles out one model of the asymmetric unit as having the best fit to the data (Rcryst = 49.7%; Z-score = –10.1). This model leads by a large margin over the nearly identical second-best model (Rcryst = 49.86%; Z-score = –8.0), in which the order of two subunits in the ring is swapped. In fact, the entire left tail of the distribution corresponds to models that are very similar to one another (Figure 1C) and gives the best R model as a consensus sequence. Independent support for this model comes from observing that its subunit arrangement is identical to that determined recently by studies that used cross-linking and mass spectrometry (Kalisman et al., 2012; Leitner et al., 2012; Herzog et al., 2012).

For comparison, we also mark in Figure 1B the R values of two other subunit arrangements that were previously reported. The PW model (Dekker et al., 2011) was reported for the crystallographic data used here, and the LFC arrangement (Cong et al., 2010) was reported based on a cryo-EM reconstruction. These arrangements have a worse fit to the data with Z-scores of –1.6 and –0.8 for PW and LFC, respectively. Interestingly, we find that our model of the asymmetric unit reveals a perfect noncrystallographic dyad axis (perpendicular to the plane of Figure 1A) that relates the two particles.

The data set of the open form is much more challenging, not only because of the lower resolution (5.5 Å), but also because a third of the residues in one of the rings were previously reported as unstructured (Muñoz et al., 2011). Indeed, the R value histogram (Figure 1E) is not as clear-cut as for the closed form, and there is no one model that is decisively best. We note, however, that the subunit consensus (Figure 1F) of the top R value models gives a particle order of BDAGZEHA-BDAGZEHA, which is identical in 12 out of the 16 subunits to the arrangement found in the closed form or by cross-linking (BDAGZQHE-BDAGZQHE). This leaves no doubt as to how to orient the previously determined subunit arrangement in the asymmetric unit. In fact, we find the corresponding model to rank among the very top models in the histogram (Rcryst = 46.0%; rank #48). Observing so much side-chain signal in this very difficult crystallographic scenario is a remarkable finding. The resulting model of the asymmetric unit assigns the full subunit in the less structured ring as CCT6-Z and the protruding subunit in the more structured ring as CCT1-A. We note that this CCT6-Z subunit causes some steric interference in an adjacent molecule in the crystal and could be the reason for the protruding CCT1-A subunit in the latter.

Model Refinement

With the subunit arrangement and orientations in the asymmetric units well confirmed, we fully refined the initial models in both forms to reveal the atomic detail needed to explain CCT function. Application of the deformable elastic network (DEN) refinement (Brunger, 2007; Schröder et al., 2010) led to models with Rfree of 41.5% and 39.8% for the closed and open forms, respectively. At this point, we continued to refine only the closed-form structure. Further iterations of manual model building and DEN refinement led to a final structure with Rfree value of 28.37%. Complete refinement statistics for the closed and open form structures are reported in Table 1. CCT subunits contain several insertions of five to ten residues that are subunit-specific. These insertions coincided with additional electron densities (Figure 2A; Figure S2 available online) when they were structured or as clear breaks in the electron density when they were unstructured. We emphasize that these consistencies are independent support for the correctness of our subunit assignment and were not used in the assignment process itself.

Our closed-form structure compares favorably to two other structures that were previously reported for the same data set (Figure 2B). The structure of Dekker et al. (2011) accompanied their initial release of the crystallographic data and was based on an incorrect subunit arrangement (PW). It is therefore not surprising that its Rfree value is six percentage points higher than ours and that its fit to the density is inferior (Figure 2C). Application of DEN refinement to this structure decreased its Rfree value but not substantially. More recently, Leitner et al. reported a structure that was based on the correct subunit arrangement (OMS) and has Rfree value that is two percentage points higher than ours. This difference in Rfree value is likely not caused by the different core backbones, which superimpose to 0.9 Å root-mean-square deviation (rmsd). Rather, the structures differ considerably in the trace of the chain termini and internal loops and in the rotamers of some side-chains. We note that Leitner et al. (2012) based their structure on a particle arrangement that was derived from cross-linking data and did not systematically explore the fit of other possible arrangements or orientations to the data. We also note that their interpretation for the open form data set is wrong, incorrectly identifying the one protruding subunit as CCT7-H instead of CCT1-A.

The Weak CCT6-Z/CCT8-Q Interface

The interfaces between the eight subunits in the ring are significantly different, in spite of their high sequence similarity. In the
closed form, the 6-Z/8-Q interface stands out as a particularly weak interface (Table 2; Figure 3A), with a total buried surface area that is only 63% of the average for a ring interface (2.595 Å²). Moreover, this weakness is apparent throughout the entire interface and is not limited to specific domains of the subunits. Interestingly, the second-weakest interface (8-Q/7-H)
and strongest interface (3-G/6-Z) are flanking 6-Z/8-Q on both sides. We note, however, that their deviations from the average values are not nearly as drastic as that of 6-Z/8-Q and thus might not have functional relevance.

Between their apical domains, the 6-Z and 8-Q subunits are pushed apart because of the strain exerted by the strong kink in the helix of the 6-Z subunit that caps the folding chamber (Figure 3B). An analysis of the evolutionary conservation at the interface (Figure S3) implicates residue Arg259 at the tip of that helix as a probable cause of the kink. This arginine is fully conserved in the CCT6-Z subunits of all eukaryotes, but its paralogous positions in the other seven subunits are always hydrophobic and

Table 1. Crystallographic Data and Refinement Statistics

<table>
<thead>
<tr>
<th>Crystallographic Data</th>
<th>Closed Form</th>
<th>Open Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>reference</td>
<td>Dekker et al., 2011</td>
<td>Muñoz et al., 2011</td>
</tr>
<tr>
<td>space group</td>
<td>P1</td>
<td>P2:2:2</td>
</tr>
<tr>
<td>unit cell parameters</td>
<td>a = 159.10 Å</td>
<td>a = 272.7 Å</td>
</tr>
<tr>
<td></td>
<td>b = 162.54 Å</td>
<td>b = 313.5 Å</td>
</tr>
<tr>
<td></td>
<td>c = 268.10 Å</td>
<td>c = 158.3 Å</td>
</tr>
<tr>
<td></td>
<td>α = 85.23°</td>
<td>β = 81.15°</td>
</tr>
<tr>
<td></td>
<td>γ = 61.17°</td>
<td></td>
</tr>
<tr>
<td>wavelength (Å)</td>
<td>1.00</td>
<td>1.114</td>
</tr>
<tr>
<td>resolution range (Å)</td>
<td>90.0–3.8</td>
<td>100.0–5.5</td>
</tr>
<tr>
<td>completeness (%)</td>
<td>91.6 (93.2)</td>
<td>92.5 (95.3)</td>
</tr>
<tr>
<td>mean f(σ)</td>
<td>8.4 (1.92)</td>
<td>9.54 (1.2)</td>
</tr>
<tr>
<td>highest resolution shell (Å)</td>
<td>4.0–3.8</td>
<td>6.0–5.5</td>
</tr>
</tbody>
</table>

Refinement

| resolution range (Å) | 90.0–3.8 | 100.0–5.5 |
| no. of reflections   | 209,671/10,483 | 44,462/2,238 |
| Rwork/Rfree          | 0.2479/0.2837 | 0.340/0.398 |
| average isotropic B-factor (Å²) | 131.4 | 277.3 |

Number of Residues

| protein               | 32 chainsb | 16 chains |
|                       | 16,716 structured residues | 6,842 structured residues |
|                       | (95.5% of all residues) | (78.8% of all residues) |
| ADP                   | 32b        |           |
| BeF₃                  | 32b        |           |
| Mg²⁺                  | 32b        |           |
| rmsd bond angle       | 1.62       | 0.65      |
| rmsd bond length      | 0.0128     | 0.002     |

Ramachandran Statistics (%)

| favored              | 80.3       | 87.1      |
| allowed              | 15.5       | 10.7      |
| outliers             | 4.2        | 2.1       |

aIn the highest resolution shell.

bNCS constraints enforced between the four octameric rings in the asymmetric units.

Figure 2. Refined X-Ray Structures

(A) An insertion (red) unique to the sequence of the CCT1-A subunit is well resolved as an additional density (delineated by arrows) in both the closed and open data sets. Such density is not observed in the adjacent CCT4-D subunit. Overlaid for scale on the CCT traces (green) is the backbone from the archaeal 1Q3R chaperonin template (black). Density is contoured at 2σ and 1σ for the closed and open forms, respectively. See also Figure S2.

(B) Refinement summaries for this work and for previous models of the same data sets. The PW subunit arrangement of the particle is that of Dekker et al. (2011), while the OMS arrangement (Kalisman et al., 2012) is as determined here.

(C) Snapshots at identical positions in the unit cell show that our model (blue) fits the m/z,DF₃ electron density map better than the model of Dekker et al. (2011) (orange). The quality of the density is also improved, showing side-chains more clearly. Maps for both models are averaged over the four rings in the asymmetric unit and contoured at 2σ. Phases are calculated from the published coordinates.
part of the hydrophobic core of the cap. The burial of the arginine side-chain in the hydrophobic core displaces the tip of the 6-Z capping helix in a direction consistent with the kink. In contrast to that example, at other positions along the 6-Z/8-Q interface, the structural weakness is coupled to loss of evolutionary conservation. For example, between the equatorial domains, we find Leu43 in CCT6-Z and Ser540 in CCT8-Q, which are not conserved. Yet the paralogous positions at the other seven interfaces are highly conserved. Overall, these data strongly suggest that the weak 6-Z/8-Q interface is not a crystallographic artifact but an intrinsic property of the closed form with functional importance. The most plausible functional explanation is that the mechanical opening of the ring starts at this interface during the transition to the open form.

The Chain Termini

The subunits of CCT have long N and C termini that are in most cases 10–20 residues long. All termini protrude from the lower part of the equatorial domains into the inner side of the particle, where they were long conjectured to form a flexible “septum” that separates the inner cavities of the two rings. We observed strong residual electron density at the midsection of the closed form that certainly supports the existence of a septum. A structural analysis of the termini in the closed form yielded similar conclusions to those reported by Dekker et al. (2011). We find that the C termini are disordered, but the N termini could be traced for an additional five to ten residues outside the equatorial domain (about half their total lengths). The N termini of all the subunits are traced inside the particle, with the exception of CCT4-D, whose N terminus is instead traced to the outside of the particle through an opening between the rings. The N termini inside the particle mainly interact with the previous subunit in the ring and in some cases with the abutting subunit on the other ring. We could not observe any common structural motif in the interactions of the N termini with the other subunits.

In sequence, the N and C termini are very different. The C termini are conserved across species both in their lengths and residue compositions, which are highly flexible and charged. The N termini, on the other hand, are much less conserved and contain more hydrophobic residues. The lax conservation and the lack of a common structural motif lead us to believe that the structured parts of the N termini do not carry a significant general function. Yet, a consequence of the unspecific “stickiness” of the N termini toward the inside surface of the particle is that they are not a significant part of the septum between the rings. It also means that it is mainly the C termini that interact with the substrate while it is in the folding chamber. Supporting evidence for this comes from a recent cross-linking study by Herzog et al. (2012), in which the C termini cross-linked to a substrate protein (2ABG) at about twice the rate of the

Table 2. Contact Areas at the Subunit Interfaces within a Ring

<table>
<thead>
<tr>
<th>Subunit Pair (Sum of Voronoi Face Area; Å²)</th>
<th>Left</th>
<th>Right</th>
<th>AG</th>
<th>GZ</th>
<th>ZQ</th>
<th>QH</th>
<th>HE</th>
<th>EB</th>
<th>BD</th>
<th>DA</th>
</tr>
</thead>
<tbody>
<tr>
<td>equatorial</td>
<td>844</td>
<td>890</td>
<td>460</td>
<td>770</td>
<td>572</td>
<td>717</td>
<td>761</td>
<td>710</td>
<td></td>
<td></td>
</tr>
<tr>
<td>middle</td>
<td>642</td>
<td>642</td>
<td>369</td>
<td>504</td>
<td>388</td>
<td>571</td>
<td>608</td>
<td>392</td>
<td></td>
<td></td>
</tr>
<tr>
<td>apical</td>
<td>1,065</td>
<td>1,239</td>
<td>796</td>
<td>911</td>
<td>1,242</td>
<td>1,139</td>
<td>1,192</td>
<td>985</td>
<td></td>
<td></td>
</tr>
<tr>
<td>apical</td>
<td>303</td>
<td>283</td>
<td>47</td>
<td>220</td>
<td>620</td>
<td>271</td>
<td>189</td>
<td>472</td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>2,854</td>
<td>3,076</td>
<td>1,627</td>
<td>2,405</td>
<td>2,792</td>
<td>2,698</td>
<td>2,750</td>
<td>2,559</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Surface area values are calculated from a Voronoi decomposition using the program by Gerstein and Chothia (1996). Left and right domains refer to the subunits as written in the letter pairs. The residue ranges of the domains are listed in Table S1. Low values are bold and high values are underlined.

Figure 3. The Weak CCT6-Z/CCT8-Q Interface in the Closed Form

(A) A “crack” that occurs between these two subunits is not observed elsewhere. Only the top ring is shown. CCT6 is brown; CCT8 is gray.
(B) Top view of the closed form overlaid with a perfect 8-fold iris shows deformation of the Z capping helix. The region around Arg259 of CCT6-Z is marked.
(C) Arg259 is buried in the hydrophobic core of the cap with the guanidinium end group of the side chain exposed on the inner side of the particle. This exerts strain on the Arg259 position that is consistent with the helical deformation. In CCT6, this position is completely conserved across all eukaryotes. The corresponding positions in other subunits are always hydrophobic. The electron density map is averaged over the four rings in the asymmetric unit and contoured at 3σ. See also Figure S3.
ATP Binding and Hydrolysis

Every subunit contains a putative ATP-binding site. In the closed form, the ATP sites of all the 16 subunits contained strong densities that fitted well to a model of ADP, BeF$_3$, and Mg$^{2+}$, which was the nucleotide analog used to maintain the closed form. Specifically, the electron density corresponding to the adenosine and ribose moieties was well resolved in all the subunits (Figures 4A and 4B). We also did not observe any significant difference in the ADP binding modes between the subunits as far as permitted by the resolution. An analysis of the evolutionary conservation of residues lining the ATP pockets (Figure 4C) shows that they are highly conserved both across the thirteen tested species and across seven of the subunits. The exception is the CCT8-Q subunit that shows the largest variation from this consensus. For example, the GDGTT motif (positions 133–137) is also ambiguous in its ability to classify this subunit. Interestingly, this duality is reflected in a recent study by Nadler-Holly et al. (2012) that linked CCT3-G to the binding of Q/N-rich protein substrates, a function we attribute more to the nonhydrolyzing substrate: actin in the closed form (Dekker et al., 2011) and tubulin in the open form (Munoz et al., 2011). These substrates suggesting that ATP hydrolysis occurs.

Substrate Binding and Functional Partitioning

The initial studies cocryrstallized CCT with two of its main substrates: actin in the closed form (Dekker et al., 2011) and tubulin in the open form (Muñoz et al., 2011). These substrates were also disordered for any backbone tracing and were reported as residual electron densities inside the rings. Our structures now enable us to locate the specific subunits that interact with these densities (Figure 5). In the open form, the residual tubulin density observed by Muñoz et al. (2011) is next to subunits CCT6-Z and CCT8-Q. In the closed form, we observed strong residual density inside the folding chamber next to subunits CCT3-G and CCT6-Z, which colocalizes with the N termini. These functions of the C termini are also consistent with their evolutionary conservation.

Figure 4. The ATP Binding Sites in the Closed Form
(A and B) The ATP sites of the CCT8-Q and CCT4-D subunits show the ADP-beryllium fluoride analogs (Be, yellow; F, magenta; Mg, green). The density corresponding to the adenosine and ribose moieties is well resolved in all subunits. Side-chains are labeled according to the position indices in the conservation analysis. The density is averaged over the four rings and contoured at 3σ.

(C) The evolutionary conservation of paralogous positions lining the ATP pockets in the eight subunits. The CCT8-Q subunit shows the most variation around the ATP site compared to the other seven subunits. Positions that are identical for at least 10 of the 13 tested eukaryotic species are marked by a capital letter. Positions that are identical for at least 7 of the 13 species are marked by a lowercase letter, and less conserved positions are marked by a star. High entropy ratio (see Experimental Procedures) marks signature positions that are highly conserved across species but differ between subunit types.
density observed by Dekker et al. (2011) in their initial study. Although the closed form was cocrystallized with Plp2 in addition to actin, we follow Dekker et al. (2011) in attributing this internal density to actin alone based on three arguments. First, Plp2 was shown by cryo-electron microscopy reconstruction to bind across and above the folding chamber (Martín-Benito et al., 2004), while this density is internal. Second, photo cross-linking located Plp2 close to CCT4-D (Dekker et al., 2011), which is far from the residual density. Finally, actin is much larger than Plp2. We did not observe any other substantial density that could locate Plp2.

Overall, the picture that emerges is that of a clearly partitioned particle. Subunits CCT3-G, 6-Z, and 8-Q bind the major substrates of CCT but have apparently lost their ATP-hydrolysis capabilities. They are located opposite in the ring to the five “power-subunits”, those that strongly hydrolyze ATP: CCT7-H, 5-E, 2-B, 4-D, and 1-A. Of special interest is subunit CCT6-Z, which appears to play a central role in substrate handling. First, at the center of the substrate-binding section, it likely participates in the binding of both actin and tubulin. Second, it is involved in both the weakest and strongest intra-ring interfaces. Lastly, unlike its neighbor, CCT8-Q, its sequence is highly conserved across eukaryote species and shows sensitivity to mutations in putative substrate-binding residues (Lin et al., 1997). The central role of CCT6-Z may explain why this is the only subunit for which there is an expressed isoform in mammals. It is suggestive that such polymorphism achieves the highest flexibility in substrate handling for the minimal amount of gene duplication, since the ATP-utilizing section is unaltered.

Interestingly, the functional partitioning is mirrored in the mutual sequence similarity between subunits (Figures 5B and 6). The sequences of the five power-subunits are highly similar to each other, while the sequences of the substrate-binding subunits have diverged both from them and from each other. It is beyond the scope of this work to answer whether we are observing conservation pressure due to ATP-related functions or evolutionary diversification due to substrate-related functions (or both). Nonetheless, this partitioning adds new complexity to the intriguing issue of how this hetero-oligomeric complex evolved from its homo-oligomeric archaeal ancestor.

**DISCUSSION**

The unbiased approach presented here confidently assigns the correct subunit type to specific electron densities in crystallographic data sets of CCT. This allows us to link features in the structures to specific subunits, to match these features between structures, and to correlate other published data of CCT with the structures. The merging of these data on the two available structures of CCT shows that functions are clearly partitioned in the particle. We focus here on the implications of this partitioning on the opening of the particle. The particle opening is associated with the release of the ATP-hydrolysis products (Meyer et al., 2003). If indeed the particle first opens at the CCT6-Z/CCT8-Q interface as we suggest, and since very little of these products is occurring at these subunits, then the opening signal must propagate through nearly half the ring before achieving a mechanistic effect. While such a long route is in accord with
For the open form, CCT was purified from bovine testes and was verified by mass-spectrometry analysis to contain tubulin with the stoichiometry of one tubulin molecule per one CCT particle. CCT was incubated with ATP-γ-S, crystallized in hanging drop, and a 5.5 Å data set was collected (Table 1).

**Homology Modeling of the Crystal Asymmetric Unit**

Homology models of the individual yeast-CCT and bovine-CCT subunits were based on the template of the single repeating subunit from the archaeal thermosome of Thermococcus strain KS-1 (determined at 2.9 Å; PDB code 1GQR) and on alignments adapted from our earlier work (Tables S2 and S3; Kalisman et al., 2012). We modeled all subunits with an identical backbone of 503 core residues by removing residues at alignment positions that had an insertion/deletion in any of the sequences. Side-chain positions were modeled by the SCWRL4 software (Krivov et al., 2009). To place a subunit model into the asymmetric unit, the model was broken into its three main domains (Table S1), and each domain was rigidly superimposed on the backbone coordinates deposited by the earlier studies (Dekker et al., 2011; Muñoz et al., 2011). The typical root mean square deviation was 0.6 Å per domain. The domain-wise model building increased the accuracy of our starting models by accounting for large variations in particle shape between thermosome and CCT. This model building protocol resulted in all-atom models for the asymmetric unit that had identical backbone coordinates, regardless of their subunit order, and only differed in the coordinates of the side-chain atoms.

**Reefinement of the Closed Form Model**

Following the identification of the arrangement with the best crystallographic fit, we focused on its further refinement. The starting homology model had 4,024 ( = 8 × 503) residues per ring, which constituted 92.0% of its total sequence lengths (not counting the calcium-binding protein insert). This model was initially refined with DEN (Brunger, 2007; Schröder et al., 2010) without imposed symmetry to Rfree = 41.48%. For the next round of refinement, we introduced strong noncrystallographic symmetry (NCS) restraints designed to keep the structures of the four rings similar and used torsion-angle simulated annealing. We ran 24 parallel refinements that scanned the DEN parameter space by testing all combinations of the γ-parameter (0.0, 0.2, 0.4, 0.6, 0.8, and 1.0) and the DEN restraint weight (1.0, 10.0, 100.0, and 1,000.0). The largest drop in the Rfree value to 36.9% was obtained for γ = 0.2 and wDEN = 1.0.

Subsequent refinement attempts were hindered by over-fitting, as indicated by the tendency of Rfree to increase. We therefore switched to NCS constraints, which reduced the number of degrees of freedom 4-fold and allowed further refinement to Rfree values of 27.06%/30.34%, respectively. Several model building and refinement cycles were performed, with all manual modeling done in COOT (Emsley et al., 2010). A total of 4,179 residues with all their constituent atoms were finally included in the ring model (95.5% of all residues). Rather than performing a costly DEN parameter search at each modeling/refinement cycle, we kept the γ value at 1.0 and the weight wDEN value at 100.0.

The density in the nucleotide-binding sites was not clear enough to unambiguously place the ADP molecule together with BeF3 and the Mg2+ ion. The ADP-BeF3-Mg complex was therefore restrained to maintain the coordination of BeF3 and Mg with respect to the ADP molecule, as observed in the homologous site of the GroEL structure (PDB code 1SX3; determined at 2.0 Å).

A restrained ENCAD (Levitt et al., 1995) run was used to improve stereochemistry (Figure S4). Following the ENCAD run, we replaced some of the side-chain rotamers in the subunit interface regions (see Supplemental Information). These optimizations resulted in a drop of 0.3% in Rfree.

In the last refinement round, the model quality was good enough to allow us to switch from torsion angle dynamics to Cartesian dynamics with DEN restraints. This further decreased Rfree to 24.79%/28.37% to give our final structure. The effect of the DEN restraints is minimal at this stage of refinement: minimization of this final model in Cartesian coordinates without DEN restraints yields only slightly higher Rfree values of 25.0%/28.5%, which indicates that the structure has been refined to a stable optimum.

**Reefinement of the Open Form Model**

The models for the exhaustive R value analysis included only domains that were traced in the coordinates deposited by Muñoz et al. (2011). This meant that the subunits in the top ring comprised of all three domains (equatorial,
intermediate, and apical), but that the subunits in the bottom ring were mostly incomplete (only one subunit had all three domains, one subunit had two domains, and the other six subunits comprised only of the equatorial domain for a total of 1,142 missing residues compared to the top ring). Following the identification of the native arrangement of the asymmetric unit, we examined its electron density and found that the six missing intermediate domains in the bottom ring can all be docked into clear unassigned densities. Our starting model for refinement was therefore only missing seven apical domains in the bottom ring (total of 854 missing residues compared to the top ring). This model was refined by DEN using the default parameters. The resulting final structure had $R_{cryst}/R_{free}$ of 34.0%/39.8%.

Quantification of the Phenotypes in Amit et al. (2010)
That study mutated a conserved aspartic residue at the ATP pocket of each subunit in turn and observed the resulting phenotypic effects on the yeast cells (mutation Del6 in CCT1-A or its equivalents in the other seven genes). Despite the relatively benign nature of this mutation, the resulting phenotypic effects were drastic. Since this aspartic residue coordinates the magnesium ion at the ATP site, it is reasonable to assume that the severity of the phenotype is reporting on the extent of the ATP hydrolysis that occurs at each subunit. We quantified the severity of the phenotypes as described ahead and plotted these numbers for each subunit in Figure 5A. There were five quantifiable phenotypes measured by Amit et al. (2010): (1) growth rate at 15°C; (2) growth rate at 30°C; (3) growth rate at 37°C; (4) susceptibility to a toxic actin polymerization inhibitor, and (5) susceptibility to a toxic tubulin polymerization inhibitor. For each phenotype, we ranked the subunits according to the phenotype’s severity going from one (least affected compared to wild-type) to eight (most affected). Since the mutation to CCT4 was lethal, CCT4 was always given the rank of eight. The “average phenotypic severity” of a subunit was then simply its average rank. The values calculated for subunits A, G, Z, O, H, E, B, and D were 4, 2.8, 1.6, 1.6, 5.4, 5.4, 6.4, and 8, respectively.

Evolutionary Conservation and Sequence Identity between Subunits
CCT sequences for 13 species that span the eukaryotic evolutionary tree were retrieved and aligned as previously described (Kalisman and Levitt, 2010). Entropy ratio is defined as

$$S_{ent} = \frac{\sum_i n_i \ln(n_i)}{\sum_i \sum_k m_{ik} \ln(m_{ik})}$$

where $n_i$ and $m_k$ are counts of amino acid type $i$ in all subunits or a specific subunit, respectively. Residues that are the same in all subunits are marked with "C" in the entropy ratio column.

ACCESSION NUMBERS

The coordinates and structure factors of the refined structures have been deposited in the Protein Data Bank under accession codes 4GOL and 4APK for particles 1 and 2 of the closed form, respectively, and 4B2T for the open form.

SUPPLEMENTAL INFORMATION

Supplemental information includes four figures, three tables, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.str.2013.01.017.

ACKNOWLEDGMENTS

This work was supported by NIH grant GM063817. Computations were done on the Bio-X2 cluster (National Science Foundation award CNS-0619926). M.L. is the Robert W. and Vivian K. Cahill Professor of Cancer Research.

Received: December 5, 2012
Revised: January 20, 2013
Accepted: January 28, 2013
Published: March 7, 2013

REFERENCES


Supplemental Information

The Crystal Structures of the Eukaryotic Chaperonin CCT Reveal Its Functional Partitioning
Nir Kalisman, Gunnar F. Schröder, and Michael Levitt

Inventory of Supplemental Information
Fig S1 related to Fig 1.
Fig S2 related to Fig 2.
Fig S3 related to Fig 3.
Fig S4 related to Experimental Procedures.
Table S1 related to Table 2.
Table S2 related to Experimental Procedures.
Table S3 related to Experimental Procedures.
Supplemental Experimental Procedures and References.
### Supplementary Figures

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>PW arrangement</td>
<td>LFC arrangement</td>
<td>OMS arrangement</td>
</tr>
<tr>
<td>CNS DEN refined</td>
<td>CNS DEN refined</td>
<td>CNS DEN Refined</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>OMS arrangement</td>
<td>Side-chains truncated at Gamma atom</td>
<td>Side-chains truncated at Delta atom</td>
</tr>
<tr>
<td>CNS unrefined</td>
<td>CNS DEN refined</td>
<td>CNS DEN Refined</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>G</th>
<th>H</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>OMS arrangement</td>
<td>SFCheck unrefined</td>
<td>Number of Occurrences</td>
</tr>
<tr>
<td>CNS DEN refined</td>
<td>Reversed OMS</td>
<td>R-Value</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>J</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Occurrences</td>
</tr>
</tbody>
</table>
Supplementary Figure 1. Controls confirming the role of side-chains in the R-value signal. Related to Fig. 1.

For a given subunit arrangement of the CCT particle, there are 64 (=8 x 8) possible models of the asymmetric unit of the closed form, since both its particles have eight unique orientations each (Fig. 1A). We plot the R-values as an 8-by-8 array, where each dimension denotes the rotation of one of the particles in the asymmetric unit.

(A, B and C) show the R-values for all orientations with subunit arrangements PW (Dekker et al., 2011), LFC (Cong et al., 2010), and OMS (Kalisman et al., 2012), respectively. These subunit arrangements are very different from each other, and were all reported in previous studies based on different structure determination methods. For the OMS arrangement, the lowest fifteen R-values are arranged in a “+” pattern that is not observed for the other two arrangements. This pattern is expected from the correct arrangement irrespective of what that arrangement is: at the central point both particles are correctly oriented, along the arms the orientation of just one of the particles is correct, whereas everywhere else both particles are incorrect. The values shown are (R_{ij}-R_{min})*10, where R_{min}, the minimum percentage R-value in each plot, is 42.68%, 42.70%, and 41.48% for PW, LFC, and OMS, respectively. The orientations in the asymmetric unit are specified by a pair of integers from 0 to 7, so OMS has the best (lowest) R-value at orientation (4,4). The OMS arrangement at position (4,4) is identical to the model with best R-value identified in Figure 1B (model N). The R-values shown are obtained using the DEN (Schröder et al. 2010) refinement option in CNS (Brunger, 2007). Each value is the final free R-value for refinement against all the X-ray structure factors of a model with the particular orientation. It takes 40 core-hours of CPU time to calculate each individual R-value.

(D, E and F) The CNS R-value before refinement is calculated in just 2 core-hours and shows the same ‘+’ pattern. The dependency of the R-value signal on the side-chains is further demonstrated by repeating this analysis with truncated models of the side-chains. If the models comprise only side-chain atoms up to and including the δ atoms, the ‘+’ signal is reduced. The ‘+’ signal is further reduced when the side-chains are truncated at the γ atoms. As would be expected, truncating sidechains at the β atoms eliminates the ‘+’ signal completely (data not shown). Panels D through F used the OMS arrangement.

(G) The R-values of the OMS arrangement can be calculated much more quickly using the program SFCHECK (Vaguine et al. 1999; 13 core-seconds), which does not do the solvent
flattening done by CNS. Although the values change, the ‘+’ is still clear. The orientations examined in panels I and J are circled in their respective colors.

(H) The ‘+’ signal disappears when the CCT subunits are modeled with the OMS arrangement but with their sequences in reverse. Much smaller perturbations of the OMS arrangement including a shift of the alignment of the sequence relative to the backbone by plus or minus one also abolish the ‘+’ signal (data not shown).

(I and J) Surprised by the apparent significance of small changes in the initial R-value we used bootstrapped data sets to calculate the distribution of R-values of models at specific orientations. Each bootstrapped data set is a unique selection of approximately 63.2% (fraction 1-1/e) of the 209,733 reflections measured by Dekker et al. 2011 with additional copies taken from those selected to make a set of 209,733 structure factors; this set is degenerate and includes many duplicates, triplicates, etc, and it can be chosen in many different ways. For each set, we calculate the R-value using SFCHECK. This was repeated for 256 different bootstrap selections giving the distribution of R-values for each orientation. For clarity, only the values of the four circled orientations in Panel G are shown. In this work, the relative R-value is used in place of the absolute R-value so we also look at the R-value relative to that of orientation (4,4). The distributions of the relative R-value (panel J) are much narrower with an average standard deviation of 0.0106±0.0024% that is a factor of 7 less than that of the absolute R-value (panel I) for same four orientations. Thus the differences shown in the ‘+’ plots are actually accurate to 0.1 units of \((R_{ij}-R_{\text{min}})*10\).
Supplementary Figure 2. Insertion in CCT1-A. Related to Fig. 2.

A large insertion (red) that is unique to the equatorial domain of CCT1 occurs both in Yeast (A) and Bovine (B). The extra electron densities that corresponds to these insertions are shown in Fig. 2a.
Supplementary Figure 3. Residues at subunit interfaces. Related to Fig. 3.

Paralogous positions in the interface between subunits in the ring. The notation is as in Figure 4C. The right-hand column shows pairwise interactions across the subunit interface. The line thickness measures the extent of interaction of the two residues and the stronger interactions are listed first. The table only lists residues that make the strongest interaction ($\geq 11 \text{Å}^2$ of contact area per residue) and also have Entropy Ratios greater than 3.5. Positions 340, 81 and 711 correspond to residues Arg259 (CCT6-Z), Leu43 (CCT6-Z) and Ser540 (CCT8-Q) that are discussed in the main text.
Supplementary Figure 4.  Rotamer search and ENCAD minimization. Related to Experimental Procedures.

An overlay of the CCT1-A/CCT3-G interface before (light gray side chains) and after (magenta side chains) a refinement step that included rotamer optimization and a round of ENCAD minimization. The equatorial and middle domains of the CCT1-A subunit are colored red and green, respectively. The equatorial domain of the CCT3-G subunit is in dark gray, but only helix 2 is shown. Overall, the ENCAD minimization resulted in very small changes to the backbone and side-chain conformations. Rotamer optimization improved the geometry of poorly formed hydrogen bonds and occasionally resulted in more significant changes seen here for the glutamate side chain in the center. The new conformation of the glutamate side chain is a more commonly observed rotamer and makes hydrogen bonds with both the arginine side chain and the N-cap of the green helix. This particular glutamate geometry is also seen in some thermosome X-ray structures with higher resolution.
### Supplementary Table 1. Subunit parsing into domains. Related to Table 2.

<table>
<thead>
<tr>
<th>Subunit type</th>
<th>N-terminal</th>
<th>Equatorial</th>
<th>Middle</th>
<th>Apical</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Yeast</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A (CCT1)</td>
<td>1-18</td>
<td>19-151; 415-558</td>
<td>152-220; 380-414</td>
<td>221-379</td>
</tr>
<tr>
<td>B (CCT2)</td>
<td>1-11</td>
<td>12-148; 397-527</td>
<td>149-210; 362-396</td>
<td>211-361</td>
</tr>
<tr>
<td>G (CCT3)</td>
<td>1-13</td>
<td>14-146; 409-534</td>
<td>147-221; 374-408</td>
<td>222-373</td>
</tr>
<tr>
<td>D (CCT4)</td>
<td>1-13</td>
<td>14-145; 407-528</td>
<td>146-211; 372-406</td>
<td>212-368</td>
</tr>
<tr>
<td>E (CCT5)</td>
<td>1-39</td>
<td>40-175; 436-562</td>
<td>176-241; 401-435</td>
<td>242-400</td>
</tr>
<tr>
<td>Z (CCT6)</td>
<td>1-11</td>
<td>12-145; 408-546</td>
<td>146-211; 373-407</td>
<td>212-372</td>
</tr>
<tr>
<td>H (CCT7)</td>
<td>1-18</td>
<td>19-153; 406-537</td>
<td>154-216; 371-405</td>
<td>217-370</td>
</tr>
<tr>
<td>Q (CCT8)</td>
<td>1-20</td>
<td>21-154; 414-558</td>
<td>155-224; 379-413</td>
<td>225-376</td>
</tr>
<tr>
<td><strong>Bovine</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A (CCT1)</td>
<td>1-10</td>
<td>11-143; 405-544</td>
<td>144-212; 370-404</td>
<td>213-369</td>
</tr>
<tr>
<td>B (CCT2)</td>
<td>1-17</td>
<td>18-154; 403-534</td>
<td>155-216; 368-402</td>
<td>217-367</td>
</tr>
<tr>
<td>G (CCT3)</td>
<td>1-15</td>
<td>16-147; 404-536</td>
<td>148-216; 369-403</td>
<td>217-368</td>
</tr>
<tr>
<td>D (CCT4)</td>
<td>1-29</td>
<td>30-161; 421-542</td>
<td>162-228; 386-420</td>
<td>229-385</td>
</tr>
<tr>
<td>E (CCT5)</td>
<td>1-26</td>
<td>27-160; 415-541</td>
<td>161-226; 380-414</td>
<td>227-379</td>
</tr>
<tr>
<td>Z (CCT6)</td>
<td>1-12</td>
<td>13-143; 404-531</td>
<td>144-208; 369-403</td>
<td>209-368</td>
</tr>
<tr>
<td>H (CCT7)</td>
<td>1-14</td>
<td>15-150; 402-533</td>
<td>151-212; 367-401</td>
<td>213-366</td>
</tr>
<tr>
<td>Q (CCT8)</td>
<td>1-21</td>
<td>22-155; 405-538</td>
<td>156-219; 370-404</td>
<td>220-369</td>
</tr>
</tbody>
</table>
Supplementary Table 2. Alignment of yeast-CCT sequences to the thermosome sequence from Thermococcus strain KS-1 (PDB code 1Q3R). Related to Experimental Procedures. The '^' symbols mark the 503 positions that made the core backbone models of the CCT subunits. These positions appear in all eight sequences, and allowed us to model an identical backbone to all the eight subunits. The '*' symbols mark fully conserved positions.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1Q3R</td>
<td>GPKGDKMLVDLS-IGDVNTGATILKIDLQHPAAMKMEVEAKTVQDEKEAGDGTTAVIAGE 106</td>
</tr>
<tr>
<td>YeastA</td>
<td>GPVGLDKMLVDG-IDGTFTNGATILSSPDVHPQAGKILVELAQQQDREIGGDGTTVVIASE 107</td>
</tr>
<tr>
<td>YeastB</td>
<td>GPVGLDKMLVDG-IDGTFTNGATILSSPDVHPQAGKILVELAQQQDREIGGDGTTVVIASE 107</td>
</tr>
<tr>
<td>YeastG</td>
<td>GPVGLDKMLVDG-IDGTFTNGATILSSPDVHPQAGKILVELAQQQDREIGGDGTTVVIASE 107</td>
</tr>
<tr>
<td>YeastD</td>
<td>GPVGLDKMLVDG-IDGTFTNGATILSSPDVHPQAGKILVELAQQQDREIGGDGTTVVIASE 107</td>
</tr>
<tr>
<td>YeastE</td>
<td>GPVGLDKMLVDG-IDGTFTNGATILSSPDVHPQAGKILVELAQQQDREIGGDGTTVVIASE 107</td>
</tr>
<tr>
<td>YeastZ</td>
<td>GPVGLDKMLVDG-IDGTFTNGATILSSPDVHPQAGKILVELAQQQDREIGGDGTTVVIASE 107</td>
</tr>
<tr>
<td>YeastH</td>
<td>GPVGLDKMLVDG-IDGTFTNGATILSSPDVHPQAGKILVELAQQQDREIGGDGTTVVIASE 107</td>
</tr>
<tr>
<td>YeastQ</td>
<td>GPVGLDKMLVDG-IDGTFTNGATILSSPDVHPQAGKILVELAQQQDREIGGDGTTVVIASE 107</td>
</tr>
<tr>
<td></td>
<td>^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^</td>
</tr>
<tr>
<td></td>
<td>**                   *    *                   *    ****</td>
</tr>
<tr>
<td>1Q3R</td>
<td>LLRKAEELL-DQNIHPSIITKGYALAEOAQEKIDEAIARVD-----PDDEETLKIATAITGKN 166</td>
</tr>
<tr>
<td>YeastA</td>
<td>LLRKAEELL-DQNIHPSIITKGYALAEOAQEKIDEAIARVD-----PDDEETLKIATAITGKN 166</td>
</tr>
<tr>
<td>YeastB</td>
<td>LLRKAEELL-DQNIHPSIITKGYALAEOAQEKIDEAIARVD-----PDDEETLKIATAITGKN 166</td>
</tr>
<tr>
<td>YeastG</td>
<td>LLRKAEELL-DQNIHPSIITKGYALAEOAQEKIDEAIARVD-----PDDEETLKIATAITGKN 166</td>
</tr>
<tr>
<td>YeastD</td>
<td>LLRKAEELL-DQNIHPSIITKGYALAEOAQEKIDEAIARVD-----PDDEETLKIATAITGKN 166</td>
</tr>
<tr>
<td>YeastE</td>
<td>LLRKAEELL-DQNIHPSIITKGYALAEOAQEKIDEAIARVD-----PDDEETLKIATAITGKN 166</td>
</tr>
<tr>
<td>YeastZ</td>
<td>LLRKAEELL-DQNIHPSIITKGYALAEOAQEKIDEAIARVD-----PDDEETLKIATAITGKN 166</td>
</tr>
<tr>
<td>YeastH</td>
<td>LLRKAEELL-DQNIHPSIITKGYALAEOAQEKIDEAIARVD-----PDDEETLKIATAITGKN 166</td>
</tr>
<tr>
<td>YeastQ</td>
<td>LLRKAEELL-DQNIHPSIITKGYALAEOAQEKIDEAIARVD-----PDDEETLKIATAITGKN 166</td>
</tr>
<tr>
<td></td>
<td>^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^</td>
</tr>
<tr>
<td></td>
<td>*      *                                     *    ***</td>
</tr>
<tr>
<td>1Q3R</td>
<td>AESHKELLAKLAVEAVKQVAEKKD-------GKYVVDLDNIKFEKKAGEGVEESELVGRVVD-- 222</td>
</tr>
<tr>
<td>YeastA</td>
<td>AESHKELLAKLAVEAVKQVAEKKD-------GKYVVDLDNIKFEKKAGEGVEESELVGRVVD-- 222</td>
</tr>
<tr>
<td>YeastB</td>
<td>AESHKELLAKLAVEAVKQVAEKKD-------GKYVVDLDNIKFEKKAGEGVEESELVGRVVD-- 222</td>
</tr>
<tr>
<td>YeastG</td>
<td>AESHKELLAKLAVEAVKQVAEKKD-------GKYVVDLDNIKFEKKAGEGVEESELVGRVVD-- 222</td>
</tr>
<tr>
<td>YeastD</td>
<td>AESHKELLAKLAVEAVKQVAEKKD-------GKYVVDLDNIKFEKKAGEGVEESELVGRVVD-- 222</td>
</tr>
<tr>
<td>YeastE</td>
<td>AESHKELLAKLAVEAVKQVAEKKD-------GKYVVDLDNIKFEKKAGEGVEESELVGRVVD-- 222</td>
</tr>
<tr>
<td>YeastZ</td>
<td>AESHKELLAKLAVEAVKQVAEKKD-------GKYVVDLDNIKFEKKAGEGVEESELVGRVVD-- 222</td>
</tr>
<tr>
<td>YeastH</td>
<td>AESHKELLAKLAVEAVKQVAEKKD-------GKYVVDLDNIKFEKKAGEGVEESELVGRVVD-- 222</td>
</tr>
<tr>
<td>YeastQ</td>
<td>AESHKELLAKLAVEAVKQVAEKKD-------GKYVVDLDNIKFEKKAGEGVEESELVGRVVD-- 222</td>
</tr>
<tr>
<td></td>
<td>^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^</td>
</tr>
<tr>
<td></td>
<td>*      *                                     *    ***</td>
</tr>
<tr>
<td>1Q3R</td>
<td>-KEVVHPR--MPKRVE------NAKIALAKELEVKTE-TDAKINITSPOQLMSFLEEOEMK-- 277</td>
</tr>
<tr>
<td>YeastA</td>
<td>-KEVVHPR--MPKRVE------NAKIALAKELEVKTE-TDAKINITSPOQLMSFLEEOEMK-- 277</td>
</tr>
<tr>
<td>YeastB</td>
<td>-KEVVHPR--MPKRVE------NAKIALAKELEVKTE-TDAKINITSPOQLMSFLEEOEMK-- 277</td>
</tr>
<tr>
<td>YeastG</td>
<td>-KEVVHPR--MPKRVE------NAKIALAKELEVKTE-TDAKINITSPOQLMSFLEEOEMK-- 277</td>
</tr>
<tr>
<td>YeastD</td>
<td>-KEVVHPR--MPKRVE------NAKIALAKELEVKTE-TDAKINITSPOQLMSFLEEOEMK-- 277</td>
</tr>
<tr>
<td>YeastE</td>
<td>-KEVVHPR--MPKRVE------NAKIALAKELEVKTE-TDAKINITSPOQLMSFLEEOEMK-- 277</td>
</tr>
<tr>
<td>YeastZ</td>
<td>-KEVVHPR--MPKRVE------NAKIALAKELEVKTE-TDAKINITSPOQLMSFLEEOEMK-- 277</td>
</tr>
<tr>
<td>YeastH</td>
<td>-KEVVHPR--MPKRVE------NAKIALAKELEVKTE-TDAKINITSPOQLMSFLEEOEMK-- 277</td>
</tr>
<tr>
<td>YeastQ</td>
<td>-KEVVHPR--MPKRVE------NAKIALAKELEVKTE-TDAKINITSPOQLMSFLEEOEMK-- 277</td>
</tr>
<tr>
<td></td>
<td>^^^^^^^^^^^^^^^^^^^^^^^^^^^^^</td>
</tr>
<tr>
<td></td>
<td>*      *                                     *    ***</td>
</tr>
<tr>
<td>1Q3R</td>
<td>-KEVVHPR--MPKRVE------NAKIALAKELEVKTE-TDAKINITSPOQLMSFLEEOEMK-- 277</td>
</tr>
<tr>
<td>YeastA</td>
<td>-KEVVHPR--MPKRVE------NAKIALAKELEVKTE-TDAKINITSPOQLMSFLEEOEMK-- 277</td>
</tr>
<tr>
<td>YeastB</td>
<td>-KEVVHPR--MPKRVE------NAKIALAKELEVKTE-TDAKINITSPOQLMSFLEEOEMK-- 277</td>
</tr>
<tr>
<td>YeastG</td>
<td>-KEVVHPR--MPKRVE------NAKIALAKELEVKTE-TDAKINITSPOQLMSFLEEOEMK-- 277</td>
</tr>
<tr>
<td>YeastD</td>
<td>-KEVVHPR--MPKRVE------NAKIALAKELEVKTE-TDAKINITSPOQLMSFLEEOEMK-- 277</td>
</tr>
<tr>
<td>YeastE</td>
<td>-KEVVHPR--MPKRVE------NAKIALAKELEVKTE-TDAKINITSPOQLMSFLEEOEMK-- 277</td>
</tr>
<tr>
<td>YeastZ</td>
<td>-KEVVHPR--MPKRVE------NAKIALAKELEVKTE-TDAKINITSPOQLMSFLEEOEMK-- 277</td>
</tr>
<tr>
<td>YeastH</td>
<td>-KEVVHPR--MPKRVE------NAKIALAKELEVKTE-TDAKINITSPOQLMSFLEEOEMK-- 277</td>
</tr>
<tr>
<td>YeastQ</td>
<td>-KEVVHPR--MPKRVE------NAKIALAKELEVKTE-TDAKINITSPOQLMSFLEEOEMK-- 277</td>
</tr>
<tr>
<td></td>
<td>^^^^^^^^^^^^^^^^^^^^^^^^^^^^^</td>
</tr>
<tr>
<td></td>
<td>*      *                                     *    ***</td>
</tr>
</tbody>
</table>

The '^^^' symbols mark the 503 positions that made the core backbone models of the CCT subunits. These positions appear in all eight sequences, and allowed us to model an identical backbone to all the eight subunits. The '*' symbols mark fully conserved positions.
1Q3R DMVDHIAQTG---------ANVVFQ-----KGIDDLAQHYLAKYGIMAVRRVKKSDMEKLAKA 327
YeastA ERVKKIIDAG---------AQVVLTT-----KGIDDLKEFVEAKMGVRCCREEDLARARA 332
YeastB NKIAISKFG---------INTFINR-----QLIYDYPEQLFTDLGINSIEHADFEGOVERLAV 320
YeastG LMEQILARV--------PILYVCT------KGIDLAHYLLGCGSVLRKVKDSNNRIV 319
YeastD NICKKKKAK---------CWVPLIQSKILARNDVNLALHFLSKNIKMVLDKEERIEIFLSKG 327
YeastE EMMDDVKKKAG---------ADVVICQ-----WGFDDEANHLLLQNDLNPDMLPAVRWGQGQLEHIAIS 357
YeastZ AKLKKIDLNEVGMDDPGKVFVIIQ-----KGIDPMVLDFVNHILARAKRRNMPNERQLV 331
YeastH EKLQVQVEE---------ANIVLSDK------LPIGDLATQFFADRNIFCAGRVSADDNMVRQA 329
YeastQ AMMKEIADMG---------VECIVAG------AGVGEHALHYGINLYGVIKVFDEMRLRCV 335

1Q3R TGAKIVTNVKDLT------PEDLGYAE-VVEERKLA--GENMIFVEGCKN----PKAVTILLRGGT 380
YeastA TGATLVSMSNLGEFTFFESPLLVLGCD-EEVQAKFSS--DDECILIKGTKS----HSSSIIILRGN 391
YeastB TEGVEVSTFDEPS------KKLGLCED--VIEEIMLG--EQFKLFSGCAK------GEACTIVLVRAG 373
YeastG TGAATNVREDKL------ESEVTVNGCFLKVEMIGK--DYEFSFLLNCKE----PKACTIMLRGGS 385
YeastD LGCKFIADEILF------EDRILAD-SADDESD--GSKIVRVTGIRNNAARVTVSVRVRG 383
YeastE TNGRIPFQREGL------RITYESEQFGETTKDRLILIEEQSKE------TKTTCFTVRGNSR 412
YeastZ TGEAEQTVNVEDL------FTQILGFSG-LVYQETIG--EEKFYTVEVTD-----PKSTTILLKGR 382
YeastH VGSSIPSTTDIK------PLHEGCTA-LEFMEOIQ--SERNYLFQGCPQ------ACTKTLRRGGA 384
YeastQ CGATPLFRGQAT------PEELGLVGE--TVKTEIG-----------GDRVTQVFQEQG-EISRTTILLRG 390

1Q3R EHVIDEVERALEDAVKVVDMED-GAVLPAGGAIEELAIRLDEYAKQVG------GKEALAIENF 440
YeastA DSYLDENERSLHDSLVKVRLEES--GNVPPPQGCCVCEALAINYINLDFATTGV---SREQLAIAEF 451
YeastB DQTLDEAERSLHDALSVLSQTTKE--TRTVLGGGCAEMVMSKADVTEAQNI------GKSLAVEAF 433
YeastG KDLNIEDRNDAMAVARNVML---PSLPPSGQGATEMSVCLAEKARLEQG------GIGQQWYQAV 445
YeastD NMIETIERSLHLAVCIRLCLVKE--RGLIAAGGAIEIERSLSKEARSM------GQVAFQIQEF 443
YeastE KIMVDEARLHDSLCVRNLKVD---SRVYGGGAAEVMTSLVSEAEADKQR------GIDQYAFRGR 472
YeastZ HYLALGKTDRAVRGVLANKV------KNIPGQAFFIALSRLYIASMNKLGAKGKTKEDEIEAF 447
YeastH EGVIAEVERSLSDAIMVKRALQN------KLIAGGATEMSVCLRDSKTA-----GKQOMINAF 442
YeastQ QNNLDIERAIDDDGVAVAKLMPGKGLPKQGATPEIEILSIRITKGYERTPF-----GLLQLAIKF 452

1Q3R ADALKIIPKTLAQEGLDTVEMLVKVISEHKNR------------GLGIGIDEF------EGKPADML 490
YeastA AAAALLIIKPKTLAVNAAKDSSELVAKLRSYHAASSMQMADKVRRSRNYGDL---TGKIVDEI 501
YeastB AARLQVPTLADNAAGDSSSELVSKLRSYIYNG-----------ISTSGDLN------NGTIAIMGD 483
YeastG KDILNEIDRNDAMAVARNVML---PSLPPSGQGATEMSVCLAEKARLEQG------GIGQQWYQAV 445
YeastD ASAMECFRTILQNNAGGGPRLLLISQRLRAHQG----------NFTTGIDGQ---RGKTVMDV 495
YeastE ADSEVILEPITLASEAVLIRSIDVIAAKATKPEGQGGMPGGMGGMGGMG 548
YeastA HAGVLEPTISKVKSLKSALEACVAILRIDTMITVDPEPPKEDPH------IERTGIDML 559
YeastB QLGIVESYKLKRAVVSSASEELVLRVVNIARIRTANQHM------EIKWYVFED---TENIGDNF 492
YeastG SYGIWEPEVIKQQSVKTAIESACLLLRVDDIVSGVRKQE------------------- 534
YeastD EEHILQPVLTPGITLSTLSKQQLKE--------------KISNIGDCL------GYSNDMDK 523
YeastE AEALLIPVTLLNGFIDPPL overwhelms the primary language. It looks like the full document contains only text, so there is no need for further processing.
**Supplementary Table 3. Alignment of bovine-CCT sequences to the thermosome sequence from Thermococcus strain KS-1 (PDB code 1Q3R). Related to Experimental Procedures.**

The '^' symbols mark the 502 positions that made the core backbone models of the CCT subunits. These positions appear in all eight sequences, and allowed us to model an identical backbone to all the eight subunits. The '*' symbols mark fully conserved positions.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1Q3R</td>
<td>------------MAQLSGQPVYLPEGQRYGVRDAQLRNLARIAEEVTRTTLGFKGMKMLV 53</td>
</tr>
<tr>
<td>BovA</td>
<td>-------------------MEGPLSGQPVYLPEGQRYGVRDAQLRNLARIAEEVTRTTLGFKGMKMLV 46</td>
</tr>
<tr>
<td>BovB</td>
<td>------------MASLSLVIFNIGAEEAERTALKLSIGAIGDLVKSTLGFKGMKMLL 53</td>
</tr>
<tr>
<td>BovG</td>
<td>--------------MMGHRPVLVLSQNTKRESGRKVQNSINAAAKAVADTSGLGFKGMKMLQ 65</td>
</tr>
<tr>
<td>BovD</td>
<td>MPENVAPRTGPAAAGAAGGKSAQDLRDPKAPQRISNFNSAASKAVADTSGLGFKGMKMLQ 65</td>
</tr>
<tr>
<td>BovE</td>
<td>---MASVGTLAFDEYRPFILLIQDQDRKSRLMGLEALKSHMAAKAVANTMTKTSGLFGNLMDMV 62</td>
</tr>
<tr>
<td>BovZ</td>
<td>------------MAAVKTLNPAVARRAQALAVNSISRARLGQVLRNLPGKFGKTMKML 48</td>
</tr>
<tr>
<td>BovH</td>
<td>---MMPTVILKETDSSGQPIQLVSINACQVIAEAVRPTLPGFKGMLKMLV 50</td>
</tr>
<tr>
<td>BovQ</td>
<td>------------MALHPVKAPFQAGMLKEGAKSHFGLGEEVRYNIAQCKELAQTRTAYIPGMNKMMV 57</td>
</tr>
</tbody>
</table>

**1Q3R**
```
1Q3R  ------------MAQLSGQPVYLPEGQRYGVRDAQLRNLARIAEEVTRTTLGFKGMKMLV 53
BovA  -------------------MEGPLSGQPVYLPEGQRYGVRDAQLRNLARIAEEVTRTTLGFKGMKMLV 46
BovB  ------------MASLSLVIFNIGAEEAERTALKLSIGAIGDLVKSTLGFKGMKMLL 53
BovG  --------------MMGHRPVLVLSQNTKRESGRKVQNSINAAAKAVADTSGLGFKGMKMLQ 65
BovD  MPENVAPRTGPAAAGAAGGKSAQDLRDPKAPQRISNFNSAASKAVADTSGLGFKGMKMLQ 65
BovE  ---MASVGTLAFDEYRPFILLIQDQDRKSRLMGLEALKSHMAAKAVANTMTKTSGLFGNLMDMV 62
BovZ  ------------MAAVKTLNPAVARRAQALAVNSISRARLGQVLRNLPGKFGKTMKML 48
BovH  ---MMPTVILKETDSSGQPIQLVSINACQVIAEAVRPTLPGFKGMLKMLV 50
BovQ  ------------MALHPVKAPFQAGMLKEGAKSHFGLGEEVRYNIAQCKELAQTRTAYIPGMNKMMV 57
```

**1Q3R**
```
1Q3R  DS--LGDIVVTNDGATILDKIDLQHPAAKMMVEVAKTDKEAGDGTTSVVIAELLENKNADELVK 116
BovA  DD--IGDVTITNDGATILKLLEVEHPAAKVLCDMGLQDKEVGDGTTSVVIIAAELLKNADELVK 109
BovB  SSGRDASLMVTNDGATILKNIGVDNPAAKVLVDMSKQDIEAGDGTTSVTVLGALLEEAEQLLD 118
BovG  DP--MGGIVMTNDGATILREIQVQHPAAKMIEISRTQDDEEGDTTSVTVAGELMSLVAEHFL 114
BovD  DG--KGDIITNDGATILRSSKDLQHPAAKVLCDMGLQDIEAGDGTTSVTVLGALLEEAEQLLD 118
BovE  DK--DDGVTITNDGATILSMSKDQRPAAKVLVDMSKQDIEAGDGTTSVTVLGALLEEAEQLLD 118
BovZ  SG--AGDIITNDGATILREIQVQHPAAKMIEISRTQDDEEGDTTSVTVLGALLEEAEQLLD 118
BovH  DG--RGKATITNDGATILRSSKDLQHPAAKVLCDMGLQDIEAGDGTTSVTVLGALLEEAEQLLD 118
BovQ  NH--LEKFLVTNDATITLEQYVPAKMMVASMRQEQVTDIETTLVFAGAGLALLEELL 120
```

**1Q3R**
```
1Q3R  QNIHPSIITKGYALAAEKAQEILDEIAIRVD----PDDEETLLKIAATSITGKNAESHKELLAKL 177
BovA  QKIHPSISVYSLCAKEVRYISENLIINT----DELGRAYCLAAKTSMSMKVINGDFPAFLN 171
BovB  KKHPIQTVIGWREATKAARGAILNSAVDHGS-DEVKFRQDMNATGTTLSKLTHHDKHTKL 182
BovG  QQMHPTVISAYKALDISSLKISIPVD----TSNDRTMNINSITSTKSVISWSLACNI 175
BovD  KGIIHPISTIRFQKALEIGIRKLTMSRVPVE--LSDERTLNSAATLNSKKSQVSSLSIPM 175
BovE  ERHPIRIAGYEAARIAEHLKDIELSDVVE--DMKNTPELQIAKTSTLGKVSNCHRQMAEI 188
BovZ  EHLHPRIITEFGEAAMEKLQFLQVKVSKVE----MDRETLDVATSIRTKVHAELVDAVLETA 171
BovH  EHLHPQIIARFATQAELNKEIATVKKEVQRLKELCAMTALSLSQKAFKAK 178
BovQ  LGSLSEVIEVYBIACKAHILFPLDVCSSA--NRDVEVSLLHTSHMSKQYGHNFVFKLAKL 183
```

**1Q3R**
```
1Q3R  AVEAKVQVAEKKD--GKYVVDLNIKFKEKKAEGVEEVESELVRGVID----KEVHFRMPKRVENAK 238
BovA  VVDAVLAIYDITQIQYRPVFRNVLNFSWVNHHIKGSAASRQEMSMLINGYALV----CVGFQSMMPKRIVNKA 233
BovB  AVEAVLRKQ-------SGNLHVI1KLSGLASDYLDEGFL-----KKMVQPKRVENAK 236
BovG  ALDAVKTVQFENGREIKDIYKARVEKIPGGIIIEDSSCVRGVMN-----KDVTHFMRMRYIKNPR 237
BovD  SVDAMVTVIPIGTDVSDMLKKGVDGTTCDCELVEGLV-----KVQAENVITREVEAK 248
BovE  AVNAVLTVADMQ---RRDVDFELIKVEKGGVGRLEDTKLKGVID----KDHFQPQMPKVQVEDAK 247
BovZ  VVSDLAIKQKD-----EPIDLFMVEMEMHSHKTDTSLRIGLVD----HAGHPRDMKPVREDAY 229
BovH  VVDAVMLDLD--------LQLMKIQIGKQVGGALEESILQGAFVAKFKTSYAGFEMQPKKYHNPM 238
BovQ  AQACVSIFPDS-----GHIENVMDRCVKGSSHVLHMVF------KTEGVDTSVQKDAK 238
```

**1Q3R**
```
1Q3R  IALINEALEVKKTE-TDAKINITSPDQIMSFLPSLEQEEMLKDMVDHIAQTG----------ANVFVF 293
BovA  IACLDFSQKTMKM--GKYVVDLNIKFKEKKAEGVEEVESELVRGVID----KEVHFRMPKRVENAK 238
BovB  ILIANTGDMTDDKIQIFGSRVRVDSTKVAEIAEAEKEMEKVERILKHE-----INCFIN 292
BovG  IVLDDSLEEYKGE-SQTDIBIEEEDDRQIPMLQMEEEYIQELCEDIILQK------PDVVT 292
BovD  IGQLCFLQAPKDT---MDQIVVSDQVMDRVRERAYILNVQIKKTG----------CNVLLI 303
BovE  IALLTCFPEFPFETKHKLTVSVEDFRAKQYEOKEFEMIFIRQETG-------ANLARC 302
BovZ  ILNCVSLEYIKE-VNSHGYKSAEEREVLKAVREKFIERDVDVKKKLCVDGGDSGFKV 293
BovH  IALLVELEKAE-DNAEIKVHVYDAIIVDAEWNLIDKLEIHHSHG--------AKVVL 291
BovQ  IAVYSCEPFGMITE-TGTVLKSAAEMNFSGSHEONLMDAQVKAIDAD----------ANVVT 293
```

[^1]: ^ ^ ^
[^2]: * * *
1Q3R  Q-----KGDIALQYHLAKIGMAMVRKVSMDKLEAKATGAIKTINVDLTT-------PEDLGYA  347
BovA  T-----GGIDDMLCFVGRAGAMRVRKLKSDKRMKASGATVLSTLANLEGGEETFEASMGLQA  348
BovB  R-----QLIYNYEQPFGAOGVAAEMHDFVPGVEgLTVTGEIAESTFDHFE-------LVKLGSC  346
BovG  E-----KGISDIAQYLMARITAIARRVKTDDNRHARACGARIVSRPEELR-------EEDVGTG  346
BovD  Q-----WGDDEAANHLIIQDNLPVPRVRGQDEIEELAIATGGRIVPRFSEELT-------AEKLGPA  356
BovZ  Q-----KGDIPDLSALAEKAARLRRKNMERLALCGSIALNLSLDDL------PDCGLHA  347
BovH  K-----LPIGVATQGYAFDORFCAGRVPEELDKRTMACGSISITVSQVNL------SDVLGRC  345
BovQ  G-----GRVADLHYNKINLYNKLWDLRLCKTVGATLPRNLPPY-------LEEMGHC  347
^  ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^      ^^^^^^^
*                     *
1Q3R  E-VVEERKLA--GENMIFVEGCKN-PKAVTILIRGKTEHVIDEVEERALADAVVVKVDMEGAVL  408
BovA  E-EVQERIC--DDELILIKNTKA-RTSATILRGANDFMCDEMERSLHADCVKRLEKSVV  409
BovB  K-LIEEVMIG--DEYFTFITECKD-PACTILRLGASKEILSEVENLQDAMVCNVLDDLPQLV  408
BovD  E-LAEVSLNL-SSGKLKIKITCGASPQVTIVRVSNGKLVIEAIERSLHADCVKLKLRRAL  409
BovE  G-LVEIFSGTDDKLMLVEQICN-SRAVFTIGGNNMKIIIEAEKRSLHADCVLIRLIRNDRV  419
BovZ  G-LVYETLGL--EEKFIIKCCN-PRSSTLIIKGNHTLQIKDAIRDLGAVKAIDDCGCCV  406
BovH  Q-VFEETQIG--GERNFTTGCPS-AKCTIIILGGAEQFMEEETERSLHADМИRAKINDSVV  406
BovQ  D-SYLVSGEV-DOJQVVFKHEKDASTVILGSTDNLMDIERAVDGVNFTKVLTDRKRLV  409
^  ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^  
*              *
1Q3R  PAGGAEIELAIRLDEYAQVGKAEALAEFADALKIIPKTLAENAGLTDVEMVLKVXISHKR  473
BovA  PGGGAVEAALSILYENATSMSQGRELALAIENFADALKIIPKTLAENAGLTDVEMVLKVXISHKR  474
BovB  YGGGCVMWMLMAHAVTQLASRTPGKEAVAMESYAKALRMLPTIIADNAYSDALVQALRAHSEG  472
BovG  PGGGASEMAVAHALTEKSKAMTGVEQWPYRAVAQALEVIPRTLNGCSTIITTSLRANKTHQ  473
BovD  AGGGAPEIELAIRLARATRPTSLGMECYTARAFAMVESIPSTLAENAGLNPSTTVREHNGA  490
BovE  YGGGAEIASCALAVSEQEADCKPTLEQYAMRFAEADLVEIPMALAENSGMNIFQMTVERAQVKE  484
BovZ  PGAGAVEMAVAEYAPEVKYKPSKVEGFAQAFADDILIPKVLEAPGSDFGDLTVEKQVAEHSEQ  473
BovH  AGGGAEMESKYLDRHYSTQPIGKQALLILGAYAKALEIIPQRLCDNGAFDANIILNKLRHAAOG  471
BovQ  PGGGAEIELAIRLDEYALQVGKAEALAEFADALKIIPKTLAENAGLTDVEMVLKVXISHKR  474
^  ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^  
*              *
1Q3R  --------GLGIGIDVF--EGKPADMEKGIIIEPLRVKKQAIKSASEAAIMLRRDDVIAAKATK  528
BovA  QVNPERKNLWIGDLVV-NGKPRDKQAGVFEPITIVKVKSKFLAEAAITILRIDLLIHLPES  537
BovB  --------KTTAGLDMK-EGTIDMSVLGITESFPOQVQVGLLSSAEAAEVLNDAKPP  527
BovG  --------NCETGQVNGE--TQTLVMDKGLIEPEALVKQYKTATAVELLLLRRIDDVSHKKK  529
BovD  --------EKTTGINVR--KGGISNILEELVQPLLVSVLALTETVRSILKIDDVNR---  542
BovE  --------VNPALGIDC-HKLGMNKRQHVTIETLQQISLADQMVRKLICAKDDKGFSE  540
BovZ  --------QQLVGVDDLM--TGPMVAAEAGWDNYCVKQKLLHCSTVTATNLIVEDMAGMS  528
BovH  --------GMWYGVDIN---TEJDNVEFVWPEFAMVRINALTAASEAAALISVDETINKPRST  526
BovQ  --------NKVGLDIEAEVPAVKKMKALRTGVLTVTQLKWAIAK种VEAVTLRVEQTMKAFAG  531
^  ^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^  
*              *
1Q3R  PEGGQGGGMMPGGMGGMGMG  548
BovA  KDKHGHYEDAVHSALDA-  556
BovB  RVFDPHPDC------  535
BovG  GDDQSQGGAPEAQDE-----  545
BovD  542
BovE  E-------------  541
BovZ  LKG-------------  531
BovH  VDASPAAAGRRGRRGRRLH--  543
BovQ  GPFPSSGGKWWDQND-  548
Supplementary Methods

Describing an asymmetric unit model by a letter string. For the closed form, the nomenclature describing the asymmetric unit by a 32-letter subunit string is as follows. Each successive 8 letters describe the subunit order in one ring with a direction that follows the intrinsic slant of the subunits, i.e., when the particle is viewed from the outside a string for the top ring denotes a subunit progression to the right whereas a string for the bottom ring denotes a progression to the left. The positions in the asymmetric unit that correspond to the first letters of the top rings (letters 1 and 17) are defined arbitrarily. The positions corresponding to the first letters in the bottom rings (letters 9 and 25) are abutting to the first positions of the top rings. The arrangement that best fit the data according to our analysis has a string of: BDAGZQHE-BDAGZQHE-BDAGZQHE-BDAGZQHE. In our deposited coordinates (PDB codes 4AOL and 4APK) these positions correspond to chains BDAGZQHE-bdagzqhe (4AOL) and BDAGZQHE-bdagzqhe (4APK), respectively.

For the open form, the nomenclature describing the asymmetric unit by a 16-letter subunit string is as described above for the closed form: each successive 8 letters describe the subunit order in one ring. The arrangement that best fit the data according to our analysis has a string of: BDAGZQHE-BDAGZQHE. In our deposited coordinates (PDB code 4B2T) these positions correspond to chains BDAGZQHE-bdagzqhe.

Side-chain modeling at subunit interfaces. Without moving the backbone or allowing steric clashes we changed rotamers of polar or charged side-chains that are within 5 Å of another subunit in the ring. The scoring function maximized was $2\times E_{hb} + \log(0.1 + Pr(\text{rot}))$, where $E_{hb}$ is a tally of hydrogen bonds & salt bridges and $Pr(\text{rot})$ is the probability of occurrence of the rotamer as determined by Dunbrack's 2002 backbone dependent rotamer library (Dunbrack, 2002). These steps improved the $R_{cryst}/R_{free}$ values to 25.78% / 29.00%.

Deposited PDB model. The chains in the deposited PDB structures are named after the gene (A, B, G, D, E, Z, H and Q). The chains in the top ring are named with capital letters and those in the bottom ring with lowercase letters. To prevent residue numbering mistakes, the residues numbers in the bottom ring were all increased by 1,000 compared to their counterparts in the top ring. The asymmetric unit of the closed form was separated, due to its size, into two PDB entries (4AOL and 4APK) each comprising of just one particle. In these two entries, the wild-type
residue numbering of the CCT3 subunit was unchanged, since the large CBP insert after P374 was completely unstructured.

**Representation Graphics.** Molecular images were drawn by PyMol ([http://www.pymol.org](http://www.pymol.org)) and Chimera (Peterson *et al.*, 2004).

**Supplementary References**