Supporting Information

Cong et al. 10.1073/pnas.0913774107

SI Text

SI Materials and Methods.

Cryo-EM sample preparation. TRiC was purified according to established procedures (1, 2). Purified TRiC was incubated with 1-mM ATP-AlFx for 1 h at 30 °C (3) prior to vitrification on a glow-discharged 200-mesh Quantifoil holey grid with 1.2 × 1.3-μm hole size (Quantifoil Micro Tools) using a Vitrobot (FEI Company).

Map segmentation and similarity analysis. The individual subunits from both the asymmetric and the 2-fold enforced cryo-EM maps were segmented using the EMAN program *segment3d* (4). Rigid body fitting of the individual model into subunit densities was carried out using *Fit in Map* module in Chimera [http://www.cgl.ucsf.edu/chimera/; (5)].

We calculated the pairwise correlation score between each individual subunit map in the *a* ring with each of the eight subunit maps in the *b* ring of the asymmetric map. This allowed us to identify the corresponding subunit pairs between the rings and thus reveal whether the subunits in the two rings were arranged with the same ordering. The location of the 2-fold axis between the rings of TRiC in the asymmetric map was then confirmed by calculating the rotational correlation coefficient between the two rings of this map using EMAN2 (6). To enhance the significance of the similarity score difference, we subtracted the 8-fold symmetrized density from the asymmetric TRiC map. We then calculated the rotational correlation score from this difference map between the *a* ring and the *b* ring with *a* ring being fixed and *b* ring rotated every 1° clockwise up to 360° (Fig. 1*C*). The origin of this curve is arbitrary; we positioned the highest peak at the origin. But no matter where we place the origin of the curve, the location of the 2-fold axis in the map deduced from this curve is the same.

Homology model of bovine TCP1-ring complex or chaperonin containing TCP1 (TRiC/CCT) closed state. We built the initial homology models of the eight bovine TRiC subunits in the both-ring-closed conformation using MODELLER (7). We utilized the available sequences of the eight bovine TRiC subunits, and the crystal structures of homologous group II chaperonin, including thermosome (1A6E) (8) and KS-1 (1Q3Q) (9), served as template structures. Details were described in our previous study (3). Briefly, the eight bovine TRiC subunits have 35 ∼ 41% sequence identity with the template structures; with this level of sequence identity, the models are expected to be quite reliable.

Model optimization. After the subunits were identified, each homology model was adjusted to the corresponding subunit density using the *Regularize Zone* module in COOT (10) to optimize the agreement of the model with cryo-EM density map, while maintaining correct stereochemistry. During this model optimization step, priority was given primarily to placement of Cα atoms to the density. For three subunits [CCT2(β), CCT1(α), and CCT7 (η)], further model improvement was carried out for the entire subunit by optimizing the side-chain atom placement to putative side-chain density, as well as the best possible main chain dihedral angles (4, 11). In the remaining subunits, this additional optimization was carried out only on helix H8 and the adjacent loop, and the apical domain of CCT3(γ).

The electrostatic potential was calculated with CHIMERA's *Coulombic Surface Coloring* module (5).

TRiC cross-linking rationale and approach. Defining the overall subunit arrangement of TRiC using only cross-linking or only MS is extremely challenging, because all of the eight subunits are close in size and isoelectric point, and many of them are modified posttranslationally. Our approach was to identify the most probable neighbors for the subunits using a short, ideally zero-length, cross-linker. We chose formaldehyde as a cross-linker because it has one of the shortest cross-linking span (∼2–3 Å) of any cross-linking reagent (12, 13), thus making it an ideal tool for our approach. The cross-links obtained with formaldehyde are nucleophilic attack reactions, including side-chains containing sulfhydryls, amines, phenols, and imidazoles (12). The cross-linking reaction yields an ensemble of variants that adds to the complexity of TRiC itself. Therefore, we separated the individual cross-linking reaction products using a 2D gel electrophoresis analysis to maximally exploit the slight differences in size and charge. We excised from the 2D gel the few well-separated spots corresponding in size to dimers. As this region is complex, approximately 20 selected spots were excised (Table S1). We next used MS to identify the subunit content of the excised spots. All TRiC subunits could be identified by this analysis. We considered as positive near neighbors only to those arising from well-separated spots that had multiple peptides corresponding to only two subunits. Such decision is based on the assumption that the separation of the two subunits is <3 Å apart as expected from the formaldehyde cross-linker length (12, 13). Importantly, these results provide an unbiased and independent confirmation in the arrangement of the eight subunits obtained separately by cryo-EM modeling (Fig. 3).

TRiC cross-linking and nearest-neighbor analysis. TRiC cross-linking was performed in 50 mM KHepes (pH 7.4), 50 mM KCl, 10 mM $MgCl_2$, and 10% glycerol. To initiate cross-linking, formaldehyde (0.1% vol/vol final) was added to 2 μM TRiC and incubated for 20 min at room temperature. The low concentration of TRiC ensured that TRiC was fully monomeric and that no intercomplex cross-links could be obtained. The formaldehyde concentration was empirically determined to maximize the formation of dimers and not higher-order cross-links. The low yield of cross-links ensured that few nonspecific cross-links were produced, and these would not result in discernible spots in the 2D gel. The reaction was quenched by adding 70 mM TrisCl (pH 8.0) (Fig. S5*A*). Crosslinked TRiC (300 μg) was separated by 2D electrophoresis according to the carrier ampholine method of isoelectric focusing (14) by Kendrick Labs (Fig. S5*B*). Isoelectric focusing was performed using pH 5–8 ampholines (GE Healthcare), and proteins were then separated by SDS-PAGE using 7% polyacrylamide and stained with Coomassie blue. In-gel digestion was performed essentially as in ref. 15. Few well-separated spots corresponding in size to dimers were excised from the 2D gel and analyzed by MS on an ESI-QTOF Mass Analyzer (QSTAR® Elite Hybrid Quadrupole TOF, Applied Biosystems). TRiC subunits in each spot were identified from the MS data using ProteinPilot 2.0.1 (Applied Biosystems). Only proteins identified with ≥2 high confidence peptides were considered. For nearest-neighbor analysis, spots consistent with a dimer size and containing predominantly two or fewer subunits were considered (Fig. S5*C*).

A summary of data and complete methods is provided in Table S1. An overlapping or mixed set of neighboring cross-links, which would overlap in the gel, would yield peptides from multiple subunits (e.g., spot G in Table S1). These data demonstrate that we can detect all subunits of TRiC excised from a 2D gel by

MS. Western blot analysis of the 2D gels using subunit-specific antibodies, similar to a previous approach (16), also confirmed specific cross-links [e.g., CCT2(β)–CCT5(ϵ)]; however, the MS approach is more sensitive and specific and does not suffer from the lack of specificity and affinity of some of the commercially available TRiC subunit-specific antibodies.

Methods for in-gel digest, LC-MS data acquisition, and data analysis. In-gel digestion for excised gel slices followed a standard protocol (15) with modifications as follows: DTT was replaced with 5 mM tris(2-carboxyethyl)phosphine in 100 mM ammonium bicarbonate; 200 mM iodoacetic acid in 100 mM sodium hydroxide was diluted to 55 mM using 100 mM ammonium bicarbonate; a sufficient volume of trypsin buffer was added to completely cover each gel slice and incubated overnight. Resulting samples were vacuum concentrated for 1–2 h but generally not taken to dryness. For dried samples, 50 μL of 10% acetonitrile (ACN) in 100 mM ammonium bicarbonate solution was added and the sample was resuspended by vortexing for 2 min.

For chromatographic separation, 10 μL of the extracted peptide solution from each gel spot was autoinjected using an Eksigent Nano-LC (Applied Biosystems) onto a PepMap100 trapping column (0.3 × 5 mm) and washed for 20 min at 15 μL/min. Reverse-phase separation was completed on a PepMap100 column (75 μm × 15 cm) at a flow rate of 300 nL/min using buffers 2% ACN, 0.1% FA (A), and 98% ACN, 0.1% FA (B). The gradient was run as follows: 2–30% B in 15 min, 30–80% B in 2 min, 80% B for 3 min, 80–2% B in 2 min, and 2% B for 30 min. The column flow was directly connected to an ESI-QTOF Mass Analyzer (QSTAR® Elite Hybrid Quadrupole TOF, Applied Biosystems) using electrospray ionization at 2350 V. Precursor ions were scanned from 350 to 1,800 *m/z*. Three product ion scans were set to be collected for each cycle and the mass window for transmission was set to low. Only ions exceeding 35 counts were selected as parent ions for fragmentation, and fragments were scanned from 70 to 1,800 *m/z*. Parent ions and their isotopes were excluded from further selection for 90 s. A mass tolerance was set to 100 ppm. The fragment intensity multiplier was set to 4.0, with a maximum accumulation time of 2.0 s. Collected mass spectra were analyzed using ProteinPilot 2.0.1 (Applied Biosystems). A FASTA file containing all the ORF protein sequences of the TRiC complex was added to the UniProtKB/Swiss-Prot database [downloaded Feb. 2008 (17)]. ProteinPilot was run in rapid identification mode, with trypsin as the cleavage enzyme.

1. Ferreyra RG, Frydman J (2000) Purification of the cytosolic chaperonin TRiC from bovine testis. *Method Mol Biol*, 140:153–160.
2. Meyer AS, et al. (2003) Closing the folding chamber of the eukaryotic chaperonin requires the transition state of ATP hydrolysis. *Cell*, 113:369–381.
3. Booth CR, et al. (2008) Mechanism of lid closure in the eukaryotic chaperonin TRiC/CCT. *Nat Struct Mol Biol*, 15:746–753.
4. Ludtke SJ, et al. (2008) De novo backbone trace of GroEL from single particle electron cryomicroscopy. *Structure*, 16:441–448.
5. Pettersen EF, et al. (2004) UCSF Chimera—A visualization system for exploratory research and analysis. *J Comput Chem*, 25:1605–1612.
6. Tang G, et al. (2007) EMAN2: An extensible image processing suite for electron microscopy. *J Struct Biol*, 157:38–46.
7. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol*, 234:779–815.
8. Ditzel L, et al. (1998) Crystal structure of the thermosome, the archaeal chaperonin and homolog of CCT. *Cell*, 93:125–138.
9. Iizuka R, et al. (2004) Role of the helical protrusion in the conformational change and molecular chaperone activity of the archaeal group II chaperonin. *J Biol Chem*, 279:18834–18839.
10. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D* 60(Pt 1, Pt 12):2126–2132.
11. Jiang W, et al. (2008) Backbone structure of the infectious epsilon15 virus capsid revealed by electron cryomicroscopy. *Nature*, 451:1130–1134.
12. Nadeau OW, Carlson GM (2005) Protein interactions captured by chemical cross-linking. *Protein–Protein Interactions: A Molecular Cloning Manual*, eds Golemis E, Adams PD (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York), 2nd Ed, pp 105–128.
13. Rice NA, Nadeau OW, Yang Q, Carlson GM (2002) The calmodulin-binding domain of the catalytic gamma subunit of phosphorylase kinase interacts with its inhibitory alpha subunit: Evidence for a Ca^{2+} sensitive network of quaternary interactions. *J Biol Chem*, 277:14681–14687.
14. O'Farrell PH (1975) High resolution two-dimensional electrophoresis of proteins. *J Biol Chem*, 250:4007–4021.
15. Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M (2007) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc*, 1:2856–2860.
16. Liou AK, Willison KR (1997) Elucidation of the subunit orientation in CCT (chaperonin containing TCP1) from the subunit composition of CCT micro-complexes. *EMBO J*, 16:4311–4316.
17. UniProt Consortium (2009) The Universal Protein Resource (UniProt) 2009. *Nucleic Acids Res*, 37(Suppl 1):D169–174.

Fig. S1. (*A*) Sequence alignment of eight bovine TRiC subunits. For each subunit, a unique insertion or sequence stretch is highlighted by yellow and red frames, which is further illustrated in Figs. S3 and S4. (*B*) Ribbon diagram showing the general architecture of a TRiC subunit, with α-helices longer than two turns and β-sheets being labeled.

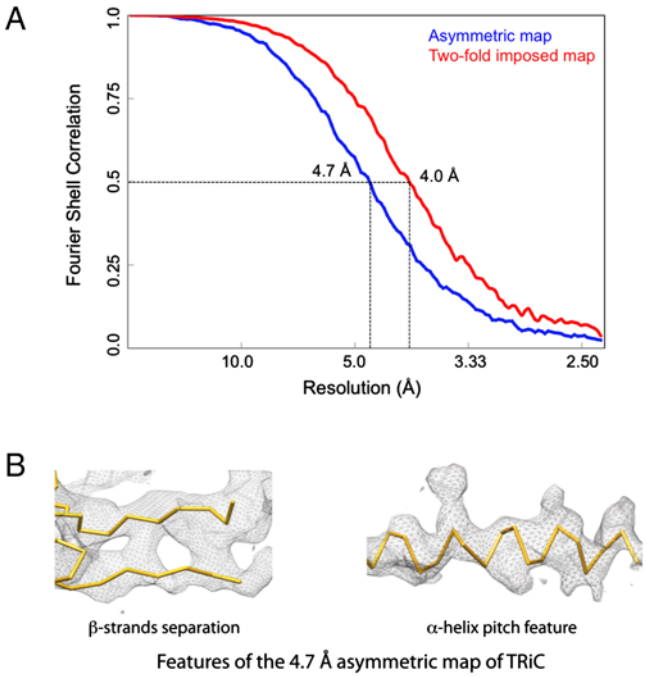

Fig. S2. (*A*) Resolution assessment of the asymmetric map (*Blue Curve*) and the 2-fold imposed map (*Red Curve*) based on Fourier shell correlation of two independent reconstructions. Using the 0.5 criteria (dashed line), a resolution of 4.7 Å was obtained for the asymmetric map, and 4.0 Å for the 2-fold imposed map. (*B*) Representative features of the 4.7-Å resolution density map with no symmetry imposed.

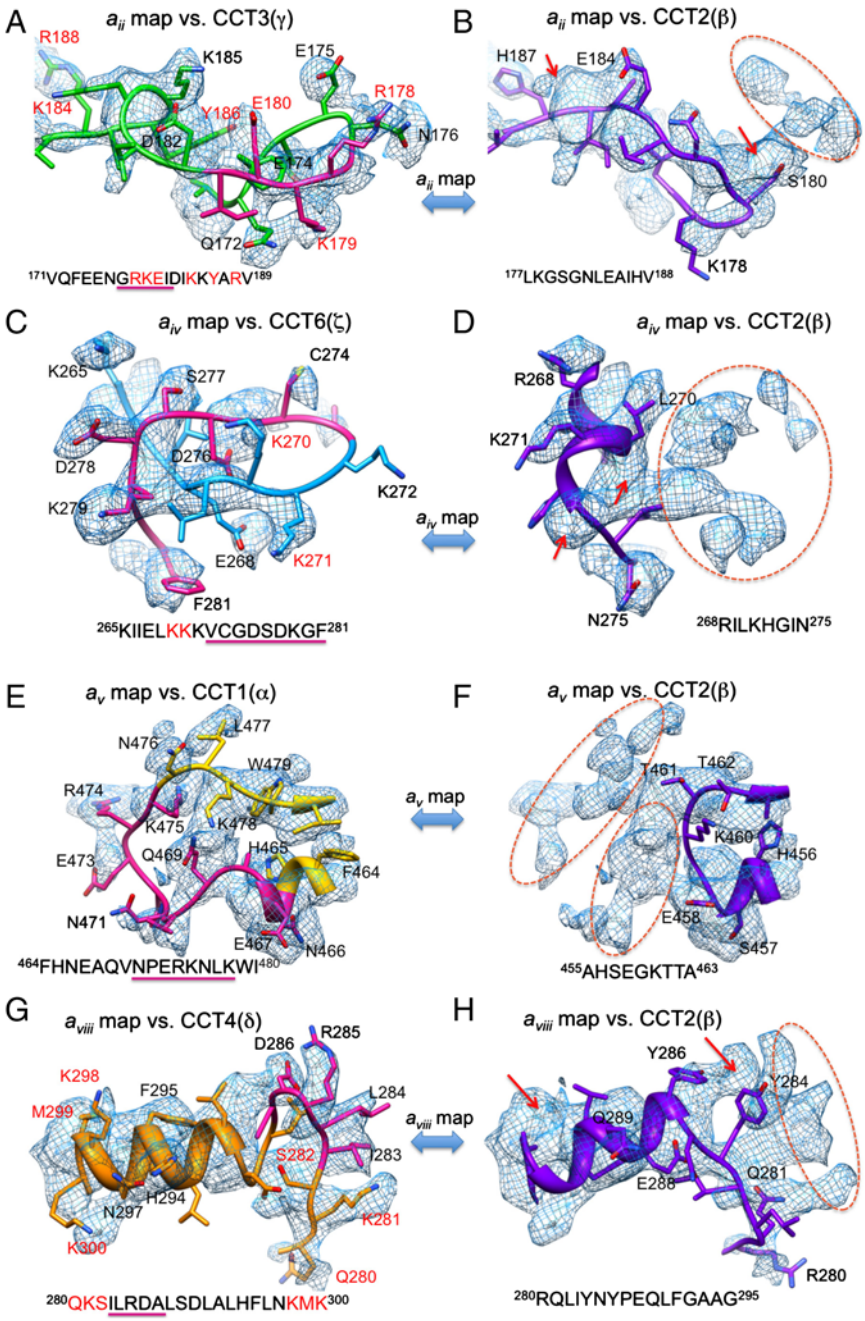

Fig. S3. (*A*, *C*, *E*, and *G*) Examples of a unique fit between the map and the identified model in a characteristic insertion region for four TRiC subunits with such insertions: CCT3(γ) in *A*, CCT6(ζ) in *C*, CCT1(α) in *E*, and CCT4(δ) in *G*. The insertion sequence is highlighted in dark purple in both the model and the sequence stretch. Unique sequence stretches are highlighted in red characters. For each of the insertion shown in *A*, *C*, *E*, and *G*, we also show a comparison case (*B*, *D*, *F*, and *H*) by fitting CCT2(β) model into the same subunit map as in *A*, *C*, *E*, and *G*. CCT2(β) does not have any significant insertion in the entire sequence, which allows it to serve as a control case. For example, CCT3(γ) fits into the a_{ii} map very well (*A*), whereas fitting CCT2(β) into the same map will have an obvious unoccupied density as shown in *B*, highlighted by orange dotted circle and red arrowhead.

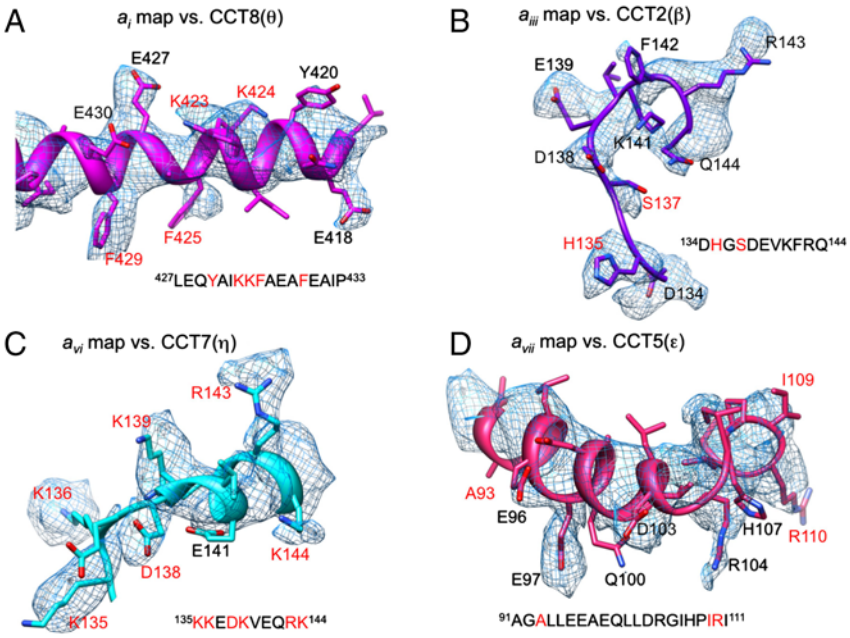

Fig. S4. Examples of a unique fit between the subunit map and the identified model in a characteristic sequence stretch for the four TRiC subunits without unique insertion: CCT8(θ) in *A*, CCT2(β) in *B*, CCT7(η) in *C*, and CCT5(ε) in *D*. The unique sequence stretches are highlighted in red characters in both the sequence and the model, which are also shown in Fig. S1*A*.

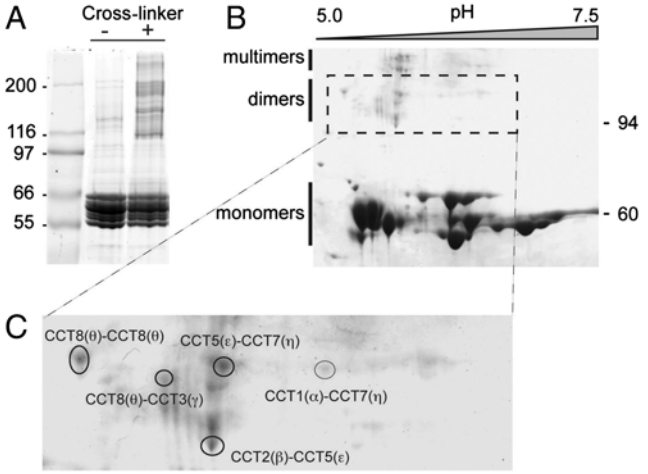

Fig. S5. Cross-linking and nearest-neighbor analysis. (*A*) Formaldehyde cross-linking of TRiC results in the production of dimeric and oligomeric species. Cross-linked (+) and control (–) TRiC samples were incubated in parallel in the presence and absence of cross-linker and separated by 8% acrylamide SDS-PAGE and stained with Sypro ruby. (*B*) Two-dimensional gel electrophoresis was used to separate the cross-linked species. Spots from the region consistent with dimers (~110 kD) were examined to identify linked pairs. (*C*) Magnification of the dimeric region of the gel. Gel spots were excised and examined by MS to identify TRiC subunits. Given the similarity of molecular weight and isoelectric point among TRiC/CCT subunits, most dimeric cross-links were overlapping in the gel. Five spots from the ~110-kD region of the gel corresponded to unique cross-linked subunit dimers, as indicated in the gel. We attribute CCT8(θ)–CCT8(θ) to a homodimeric cross-link as it runs at the expected dimer molecular weight and is the only spot analyzed containing a single TRiC subunit. Of note, CCT8 (θ) also has the lowest isoelectric point of all TRiC subunits, consistent with its position at the acidic end of the gel.

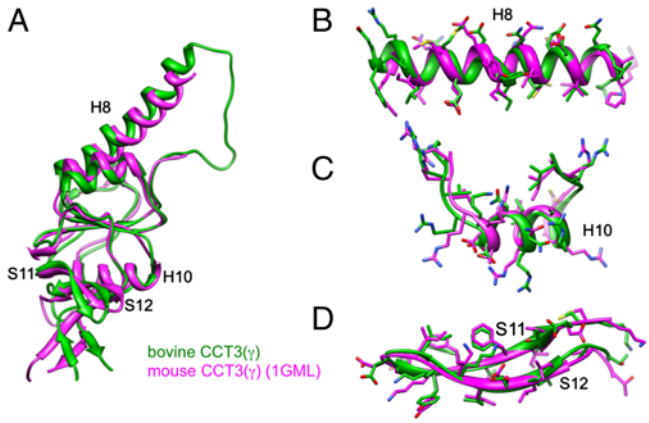

Fig. S6. Model quality validation. (*A*) Conformation comparison of mouse CCT3(γ) apical domain crystal structure (1GML chain A, *Purple*) with our corresponding optimized model of bovine CCT3(γ) (*Green*). (*B–D*) Detailed comparison between the two models in the H8 (*B*) and H10 (*C*) α-helix regions, and S11 and S12 (*D*) β-strand regions with side chains being illustrated. We zoomed in these helices and beta strands for comparison because these are well resolved in our map. Our model agrees very well with 1GML. Of note, 1GML is the crystal structure of apical domain only without the constrains of intermediate domain, therefore, the bottom beta strands underneath S11 and S12 in *A* appear to have a conformation variation from our model, which is from a complete CCT3(γ) subunit.

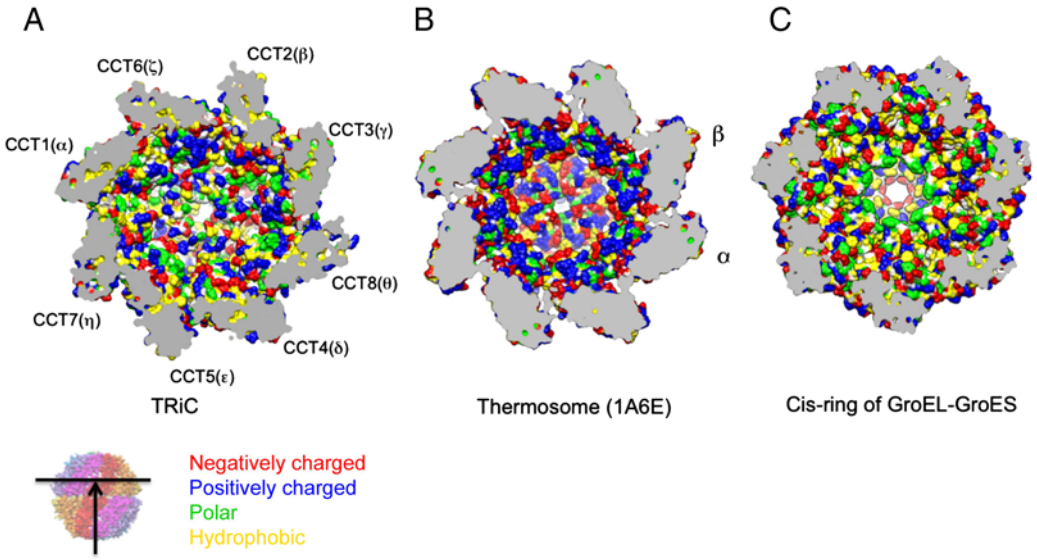

Fig. S7. Central cavity residue properties of TRiC (*A*), thermosome (1A6E, in *B*), and the *cis*-ring of GroEL-GroES (1AON, in *C*). Cutaway views of the apical and intermediate domains of the three complexes showing various side chains: positively charged (*Blue*), negatively charged (*Red*), polar (*Green*), hydrophobic (*Yellow*), and main chain (*White*). The viewing angle is illustrated in the lower left small panel.

Other Supporting Information Files

Table S1 (DOCX)