State-dependent sequential allostery exhibited by chaperonin TRiC/CCT revealed by network analysis of Cryo-EM maps

Yan Zhang [a, 1], James Krieger [a, 1], Karolina Mikulska-Ruminska [a, 2], Burak Kaynak [a], Carlos Oscar S. Sorzano [b], José-María Carazo [b], Jianhua Xing [a], Ivet Bahar [a, *]

[a] Department of Computational and Systems Biology, University of Pittsburgh, 800 Murdoch Building, 3420 Forbes Avenue, Pittsburgh, PA, 15261, USA
[b] Centro Nacional de Biotecnología (CSIC), Darwin, 3, 28049, Madrid, Spain

ARTICLE INFO

ABSTRACT

The eukaryotic chaperonin TRiC/CCT plays a major role in assisting the folding of many proteins through an ATP-driven allosteric cycle. Recent structures elucidated by cryo-electron microscopy provide a broad view of the conformations visited at various stages of the chaperonin cycle, including a sequential activation of its subunits in response to nucleotide binding. But we lack a thorough mechanistic understanding of the structure-based dynamics and communication properties that underlie the TRiC/CCT machinery. In this study, we present a computational methodology based on elastic network models adapted to cryo-EM density maps to gain a deeper understanding of the structure-encoded allosteric dynamics of this hexadecameric machine. We have analysed several structures of the chaperonin resolved in different states toward mapping its conformational landscape. Our study indicates that the overall architecture intrinsically favours cooperative movements that comply with the structural variabilities observed in experiments. Furthermore, the individual subunits CCT1-CCT8 exhibit state-dependent sequential events at different states of the allosteric cycle. For example, in the ATP-bound state, subunits CCT5 and CCT4 selectively initiate the lid closure motions favoured by the overall architecture; whereas in the apo form of the heteromer, the subunit CCT7 exhibits the highest predisposition to structural change. The changes then propagate through parallel fluxes of allosteric signals to neighbours on both rings. The predicted state-dependent mechanisms of sequential activation provide new insights into TRiC/CCT intra- and inter-ring signal transduction events.

© 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

1.1. Cryo-EM revolution is accompanied by advances in molecular modelling and computations

Cryogenic electron microscopy (cryo-EM) has found increasingly more use in structural biology in recent years as a maturing tool to characterize not only a single structure but the ensemble of conformations, often relevant to biological function, especially for supramolecular assemblies (Callaway, 2020; Nogales, 2016). Thanks to radically new technological advances experienced by this technique in recent years in both microscope hardware and image processing software, termed the resolution revolution, cryo-EM is now capable of capturing information about biomolecular structure in multiple states to near-atomic resolution (Cheng, 2018; Mitra, 2019; Nogales, 2016).

Cryo-EM data on macromolecules consists of a large number of

2D projections of "particles" in an unknown range of orientations and conformations trapped in a thin layer of vitreous ice after plunge freezing (Mitra, 2019). 3D structures are recovered by maximum likelihood image analysis methods to assign orientations to each particle and reconstruct 3D structural information (Cossio and Hummer, 2018; Scheres et al., 2007; Vilas et al., 2018). While methods are emerging for initial model reconstructions (Elmlund et al., 2013; Gomez-Blanco et al., 2019; Lyumkis et al., 2013; Punjani et al., 2017; Xie et al., 2020) and dealing with continuous heterogeneity (Maji et al., 2020; Sorzano et al., 2019), 3D reconstruction is usually achieved by alignment to a reference structure and aided by classification (Cossio and Hummer, 2018; Vilas et al., 2018). The result is a number of electron density maps corresponding to class averages, which can be found in the EM Data Bank (EMDB) or EM Data Resource (Lawson et al., 2016, 2020). Density maps are then analysed further to better understand the underlying structure and structural variability of the protein or complex being studied. This is usually done by building or fitting and refining atomic models to match the density map (Alnabati and Kihara, 2019; Kim et al., 2020; Malhotra et al., 2019; Miyashita and Tama, 2018; Nicholls et al., 2018).

Fitting of existing x-ray crystallographic, nuclear magnetic resonance (NMR), and/or integrative modelling data on the components of a complex structure or assembly resolved by cryo-EM is especially important for low-to-intermediate resolution cryo-EM maps while *ab initio* modelling may be carried out at higher resolutions (~4 Å or better) (DiMaio and Chiu, 2016; Joseph et al., 2017; Koukos and Bonvin, 2020; Leelananda and Lindert, 2020; Nicholls et al., 2018). Given the conformational variability of these components, *flexible fitting* methods have been developed, which use the collective modes of motion predicted by normal mode analysis (NMA) (e.g. NMFF (Tama et al., 2004a; Tama et al., 2004b), NORMA (Suhre et al., 2006) and other tools; see below), or the conformational changes sampled in molecular dynamics (MD) simulations (e.g. MDFF (Trabuco et al., 2008)). Similar approaches are used in replica simulation frameworks for fitting ensembles of structures and inferring dynamics from cryo-EM maps (Bonomi et al., 2018; Bonomi and Vendruscolo, 2019; Miyashita et al., 2017). Accelerated by enhanced sampling schemes (Abrams and Bussi, 2013; Bernardi et al., 2015; Harpole and Delemotte, 2018) and/or dedicated hardware such as the Anton supercomputers (Shaw et al., 2009, 2014), MD simulations are also used for exploring the microseconds to milliseconds dynamics of proteins (Dror et al., 2012; Hollingsworth and Dror, 2018).

However, conducting MD simulations of supramolecular systems at atomic detail is still extremely expensive, both in terms of computing time and memory (Hollingsworth and Dror, 2018). MD simulations suffer from inaccuracies in sampling the conformational space, especially the cooperative changes in structure that are often functional. MD simulations are also sensitive to force field parameters, especially over the long timescales needed to observe biologically relevant events (Bottaro and Lindorff-Larsen, 2018). It is thus desirable to develop physically inspired models and methods that lend themselves to efficient analytical solutions for large data analyses and cooperative/global motions of molecular machines, albeit at low resolution, rather than detailed atomic models that provide precise information but on limited length and time scales.

1.2. Insights into the allosteric dynamics of cryo-EM structures can be gained upon representing their electron density maps by elastic network models

Elastic network models (ENMs) combined with normal mode analysis (NMA) provide a robust framework for efficiently determining collective motions of biological relevance (Bahar et al., 2010a, 2010b; Doruker et al., 2012), as shown for biomolecular systems at multiple scales, from single proteins to supramolecular assemblies (Atilgan et al., 2001; Bahar et al., 2010b; Eyal et al., 2015), and more recently the chromatin (Sauerwald et al., 2017; Zhang et al., 2020a). These models allow for analytical derivations of the spectrum of normal modes of motion, *uniquely* encoded by the architecture. Each normal mode is characterized by a vibrational frequency and *mode shape* (i.e. movements of all nodes/residues of the network/structure along each mode). In particular, the global modes or *softest modes* at the lowest frequency end of the spectrum robustly enable the predisposition of the structure to cooperative changes relevant to function (Bahar et al., 2010b; Zheng et al., 2006). These properties have led to the development of many flexible fitting algorithms using NMA (Costa et al., 2020; Hinsen et al., 2005; Lopez-Blanco and Chacon, 2013; Matsumoto and Ishida, 2009; Miyashita and Tama, 2018; Suhre et al., 2006; Tama et al., 2004a; Velazquez-Muriel and Carazo, 2007; Zheng and Tekpinar, 2014) as well as hybrid methods that utilize ENMs for modelling mechanisms of action, allosteric signalling and evolution (Krieger et al., 2020).

While ENM NMA emerged as an efficient tool for investigating the dynamics of large systems where high resolution atomic models may be lacking, such as those characterised by cryo-EM, a barrier has been that electron density maps could not be readily converted into a network model. Among approaches developed to address this issue, the topology representing network (TRN) method (Martinetz and Schulten, 1994; Ming et al., 2002a) and the closely related quantized elastic deformational model (Martinetz and Schulten, 1994; Ming et al., 2002a) have shown good performance and robustness (Beuron et al., 2003; Chacon et al., 2003; Kong et al., 2003; Ming et al., 2002b) as has the approximation-accuracy control method (Jonic and Sorzano, 2016).

In the present study, we adopt and extend the TRN method to generate a coarse-grained (CG) representation of cryo-EM density maps amenable to ENM analysis at large scale. We also developed (i) a new implementation of adaptive ANM (Yang et al., 2009c), enabling easier access to this method for calculating conformational transitions; (ii) an extension of topology preserving techniques, Spectral Clustering (Lee and Verleysen, 2007; Ma and Fu, 2012) and Self-Organizing Maps (SOM) (Kohonen, 2001) for identification of dynamical domains related to subunits; and (iii) a residue-agnostic nearest neighbour method for aligning CG and atomic models.

We integrated our method into the *ProDy* API, a widely used Python interface for protein dynamics analyses (Bakan et al., 2011, 2014), thus enabling its automated use for the broader community. *ProDy* features a wide range of tools for identifying functional sites and mechanisms and building ensembles of related structures. The integration of our new TRN-based pipeline into *ProDy* allows one to use various ENMs such as the Gaussian network model (GNM) (Bahar et al., 1997), the anisotropic network model (ANM) (Atilgan et al., 2001; Eyal et al., 2015), distance-dependent ENMs (Hinsen, 1998, 2008; Hinsen and Kneller, 1999; Hinsen et al., 2000; Riccardi et al., 2009; Yang et al., 2009a), the rotation and translation of blocks (RTB) model (Tama et al., 2000), and the reduced Hessian framework for modelling environment effects (Hinsen et al., 2000; Ming and Wall, 2005; Zheng and Brooks, 2005) including membrane environment (Lezon and Bahar, 2012).

1.3. The mammalian chaperonin TRiC/CCT is a cryo-EM-resolved allosteric machine that requires multiscale mode analysis

We applied our approach to the chaperonin CCT (also known as TRiC), a molecular machine of critical importance for ensuring correct folding of about 10% of eukaryotic proteins. CCT has been

implicated in a number of diseases including cancer, Huntington's, Alzheimer's and Parkinson's diseases (Jin et al., 2019b), and hence extensively studied, thus providing a wealth of structural and functional data for assessing the significance of our predictions. It is also of sufficiently large size to demonstrate the utility of our new CG methodology, which is still amenable to all-atom ANM analysis, for multiscale methodology. Finally, the availability of structural data at atomic level for CCT, as well as the rapidly growing data on this system, will serve as a testbed for our methodology.

TRiC/CCT, like the bacterial chaperonin GroEL, is composed of two oligomeric rings, stacked back-to-back against each other. However, CCT rings are hetero-octameric with pairwise sequence identities of 20−35% between the subunits (see Fig. S1A), while GroEL rings are homo-heptameric. Furthermore, CCT subunits in a given structure exhibit large structural heterogeneity (more than 6 Å root-mean-square deviation (RMSD) in their coordinates; Fig. S1B), resulting in asynchronous movements between the subunits. Previous studies indeed suggested that each subunit has different ATP binding affinities and hydrolysis capabilities (Amit et al., 2010; Reissmann et al., 2012). Yet, CCT chaperonin activity requires coordinated conformational changes across the eight subunits in a given ring as well as inter-ring couplings, driven by ATP binding and hydrolysis (Cong et al., 2012; Meyer et al., 2003; Reissmann et al., 2012; Skjaerven et al., 2015) for enabling ring closure and formation of a chamber enclosing the substrate. So, instead of moving concertedly, typical of the MWC model for allostery (Changeux, 2012), the conformational changes of CCT subunits are sequential (Gruber and Horovitz, 2016; Gruber et al., 2017; Jin et al., 2019a; Rivenzon-Segal et al., 2005), characterized by the KNF model (Bai et al., 2010; Duke et al., 2001; Koshland et al., 1966).

The allosteric cycle of CCT (Cong et al., 2012) comprises five main conformational states (Fig. 1A−B), whose interconversions are regulated by ATP binding and hydrolysis. The set of structures enclosed in the *green box* includes the apo and ATP-bound CCT (e.g., PDB models 4a0o, 5gw4, 6ks7 and 4a0v). These transition to more closed structures upon ATP hydrolysis including an asymmetric conformer (*purple*; PDB model 4a0w) and fully closed structures (*yellow*; including PDB model 3iyg). Phosphate release leads to the opening of the chaperonin (PDB model 4a13). Recent studies have expanded this structural repertoire considerably (see Fig. 1B−C) but the basic conformational features remain largely unchanged.

In the present study, with the help of NMA and a Markovian stochastic model, both hinging on the ENM representation, we will dissect the space of motions accessible to CCT and identify the dominant mechanisms of structural changes (presented in several movies) at various stages of the allosteric cycle. We will show that the functional changes in structure such as lid closure/opening are favoured by ANM modes intrinsically accessible to the 3D architecture. Our study further shows that each intermediate passage during the allosteric cycle is initiated in a state-dependent manner by different subunits, pointing to the complexity of the overall machinery of this eukaryotic chaperonin. Finally, combination of these results with ANM analysis for full-atomic structures will provide a multiscale view of the intrinsic dynamics of the individual modules (subunits and their domains) and how these are exploited in the global chaperoning machinery.

2. Methods

2.1. Topology representing network (TRN) algorithm

The original TRN algorithm has been outlined earlier (Martinetz and Schulten, 1994), and successfully applied to cryo-EM density maps (Beuron et al., 2003; Chacon et al., 2003; Kong et al., 2003;

Ming et al., 2002b). In this study, we modify the algorithm, as described below, in order to learn a CG representation of the structure from its cryo-EM density map.

The cryo-EM density map can be considered as an array of N real numbers, $\boldsymbol{\rho} = [\rho_1, \rho_2, ..., \rho_N]$, each associated with a 3D coordinate $\boldsymbol{r}_i = [x_i, y_i, z_i]$, indicating the density at the designated position. To reduce noise and remove negative densities, we first set any density that is lower than a threshold (e.g., 1.0 used in this study) to be 0. This essentially gives us a discrete probability distribution function for each voxel at $\boldsymbol{r}_i$:

$$P(\boldsymbol{r}_i) = \rho_i \Big/ \sum_j \rho_j \tag{1}$$

Suppose we use M nodes to represent the density map, whose initial coordinates are $\boldsymbol{r}_m(0)$ with $0 \leq m \leq M$ are sampled from $P(\boldsymbol{r}_i)$. The following adaptation steps are iteratively applied to refine the positions of the nodes: (1) sample $\boldsymbol{r}$ from the distribution $P(\boldsymbol{r}_i)$; (2) determine the distance of each one of the node positions $\boldsymbol{r}_m$ to $\boldsymbol{r}$ and sort these distances into a ranking k_m, such that the closest one to $\boldsymbol{r}$ has rank $k = 0$, and the farthest, $k = M - 1$; and (3) update the coordinates of all m nodes following

$$\boldsymbol{r}_m(t+1) = \boldsymbol{r}_m(t) + \varepsilon e^{-k_m/\lambda}(\boldsymbol{r} - \boldsymbol{r}_m(t)) \tag{2}$$

where t is the number of adaptation steps already applied, ε is the step size, and λ is a parameter effectively controlling the number of updated nodes. Intuitively, the closest node is dragged towards $\boldsymbol{r}_i$ with the distance $\varepsilon \boldsymbol{r}_i - \boldsymbol{r}_m(t)$, and others are dragged towards $\boldsymbol{r}_i$ in the same manner with the distance penalized based on their ranking using the exponential term. With decreasing λ, $e^{-k_m/\lambda}$ goes to zero, and effectively no node other than the closest is moved. Both ε and λ are computed adaptively during the iterations as follows:

$$\lambda(t) = \lambda_0 \left(\frac{\lambda_f}{\lambda_0}\right)^{t/t_f} \tag{3}$$

$$\varepsilon(t) = \varepsilon_0 \left(\frac{\varepsilon_f}{\varepsilon_0}\right)^{t/t_f} \tag{4}$$

so that λ and ε starts with λ_0 and ε_0, and as $t \to t_f$, λ and ε go to λ_f and ε_f, respectively. In this study, we used the following parameters (Martinetz and Schulten, 1994): $\lambda_0 = 0.2M$, $\lambda_f = 0.01$, $\varepsilon_0 = 0.3$, $\varepsilon_f = 0.05$, $t_f = 200M$, $M = 3000$ for EMD-1961 (~1/3 of CCT residues).

Nodes corresponding to unidentified densities in the centre of the CCT chamber (Cong et al., 2012) were removed, resulting in a refined model with $M = 2,908$ nodes. The network connectivity is defined as described below. The algorithm is implemented and incorporated into the *ProDy* API.

2.2. ENM-based analyses

In general, an ENM represents the structure as a bead-and-spring model where the bead positions are identified with those of α-carbons (C^{α}), and elastic springs connect pairs of beads within a cut-off distance characteristic of the range of inter-residue coordination. In the application to cryo-EM data, we used the TRN nodes to build the network. NMA and Markovian analyses were performed using methods previously implemented within *ProDy* (Bakan et al., 2011). For ClustENM, we used a procedure similar to that previously introduced (Kurkcuoglu et al., 2016). We upgraded
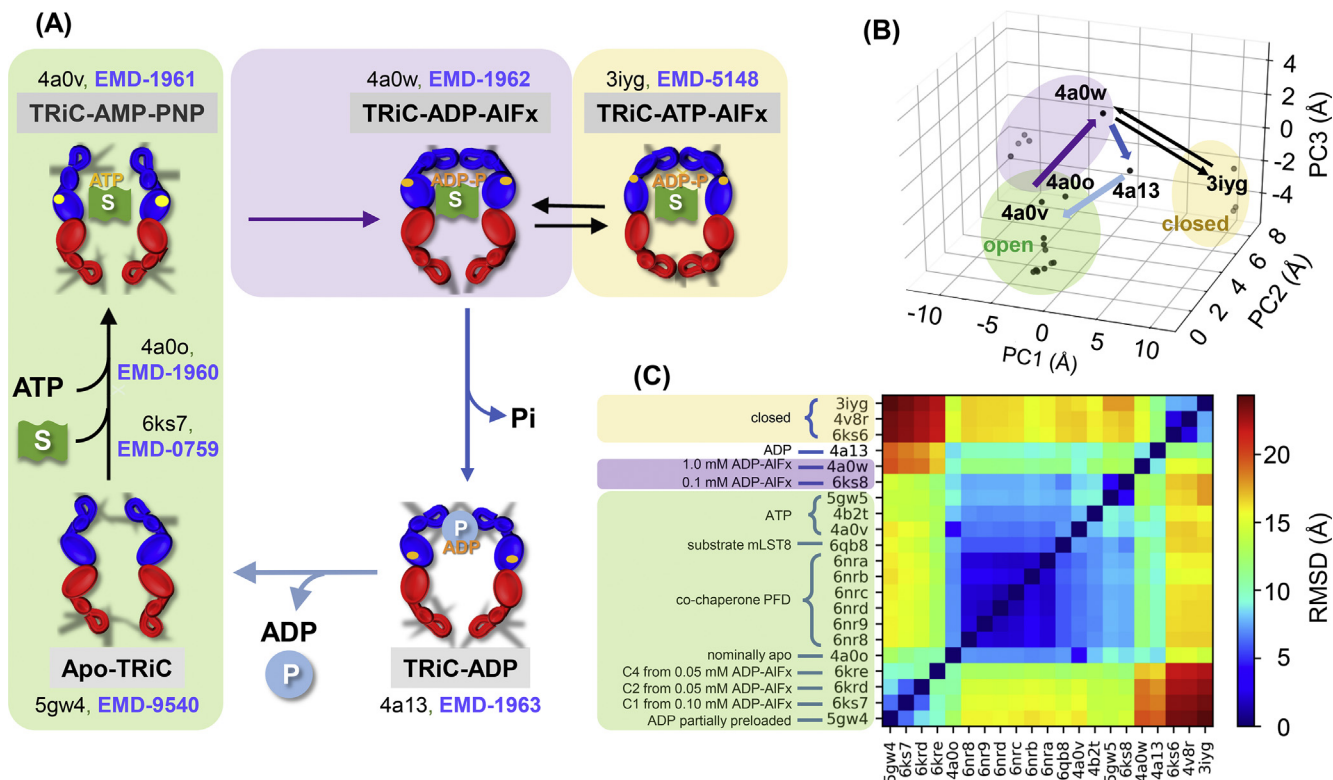
Fig. 1. The allosteric cycle of the eukaryotic chaperonin CCT and conformational states visited during the cycle. (A) Major conformational transitions triggered by ATP and substrate (S; green flag) binding and ATP hydrolysis (upper part). The upper left conformation represents the ATP-bound form as AMP-PNP (PDB: 4a0v) and ATP-γ-S (PDB: 4b2t) are non-hydrolysable ATP analogues. ATP hydrolysis is accompanied by ring closure (upper right); treatment with ADP-AlFx and ATP-AlFx (containing various aluminium fluorides) generates analogues corresponding to different stages in the hydrolysis process. Then phosphate release leads to ring opening again in the presence of the remaining ADP (lower right). This is followed by release of ADP and folded product (P; blue sphere), returning the heteromeric structure to the apo state (lower left). PDB and EMDB IDs are labelled for key resolved states. An ensemble of intermediate structures has been resolved for the ATP binding transition (including PDB: 6ks7) (Jin et al., 2019a). **(B)** Projection of all known structures onto the first three principal components (PC1-PC3) deduced from the principal component analysis (PCA) of the known conformers (each represented by a dot). The purple, dark blue and light blue arrows show the primary pathway between structures with at least one ring open, and the thin black equilibrium arrows display the passages between the kinetically trapped, fully closed conformations. **(C)** RMSDs between all pairs of structures shown in **B**. Green, yellow and purple shaded regions refer to conformers in open, closed and intermediate states in all panels.

the adaptive ANM (aANM) (Yang et al., 2009c) to improve the efficiency of sampling the space between pairs of conformers, also implemented in *ProDy*.

In order to compare the mode spectra predicted by the TRN representation of electron density maps and corresponding all-atomic models in the PDB, we aligned the TRN against the atomic models using a nearest neighbour approach where each TRN bead was matched with the nearest α-carbon; and we treated the non-aligned nodes as the environment of the aligned system using the reduced model framework (Hinsen et al., 2000; Ming and Wall, 2005; Zheng and Brooks, 2005). See **Supplementary Methods** for details.

2.3. Subunit identification by spectral clustering

Spectral Clustering (SC) is based on the graph Laplacian, which encodes the connectivity between data points. The eigenvector corresponding to the first nonzero eigenvalue of the Laplacian, also called Fiedler vector, and a set of consecutive eigenvectors partition the Laplacian in such a way that it becomes a block diagonal, each block representing a principal subspace associated with a different cluster (Fiedler, 1989; Jianbo and Malik, 2000; Ng et al., 2001). Here we used a Gaussian kernel constructed by a pairwise distance matrix with a soft cut-off of 10 Å, which ensures full connectivity. Using SC, we organized the graph into 8 clusters, each containing two CCT subunits. A second round of SC subdivided each cluster

into four, each containing the equatorial and apical domains of the subunit pairs. Sub-clusters of the same subunit were then combined which led to 16 clusters corresponding to the 16 subunits. We compared the results from SC with those from Self-Organizing Maps (SOM) plus k-means (Fig. S2). We defined the "ground truth" by mapping beads to the nearest α-carbons in the corresponding all-atom model (PDB: 4a0v).

3. Results

3.1. TRN representation of electron density maps coupled with ANM provides an effective description of the CCT architecture and equilibrium dynamics

As schematically depicted in Fig. 2A, the subunits are designated as a_1 to a_8 in one of the rings (termed *cis* or upper ring), and a_1' to a_8' in the other (*trans* or lower), using the nomenclature introduced in cryo-EM studies (Cong et al., 2012; Zang et al., 2016). Supplementary Table S1 lists alternative nomenclature/identifications adopted in different studies, as reference. In the current study, we use either the identifiers a_1-a_8 and a_1'-a_8' for the respective *cis* and *trans* rings, or the corresponding subunit numbers CCT 2, 4, 1, 3, 6, 8, 7, 5 (same in both rings), identified recently (Kalisman et al., 2013; Leitner et al., 2012; Wang et al., 2018; Zang et al., 2018), or both names (a_1/CCT2, a_2/CCT4, …, a_8'/CCT5) for clarity.

The algorithm assigns a network of *N* nodes to a probability
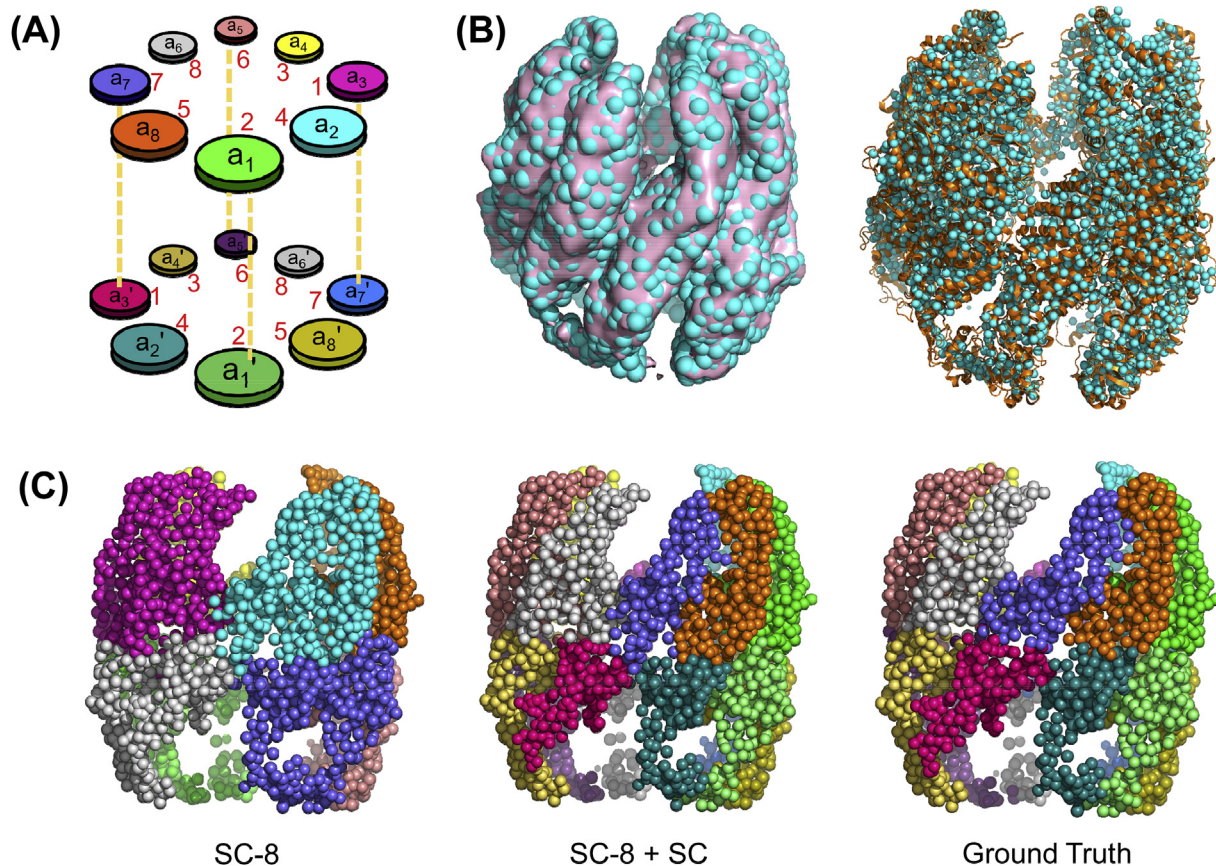
Fig. 2. Model for ATP-bound CCT generated using the Topology Representing Network (TRN) and Spectral Clustering (SC) guided by ANM. (A) Diagram depicting the nomenclature of subunits in the two hetero-octameric rings. The pairs of subunits a_k and a_k' belonging to the respective upper and lower rings are sequentially identical; the pairs (a_1, a_1') and (a_5, a_5') are vertically stacked against each other while the remaining subunits are arranged in a two-fold symmetric way. Numbers in *red* correspond to the biological subunit numbers without the CCT prefix. **(B)** Superposition of the refined network of 2,908 nodes (*cyan, spheres*) with either the electron density map (EMD-1961; *pink surface; left*) or the corresponding all-atom model (*right*; PDB: 4a0v; *orange ribbon*). **(C)** Decomposition of the structure into clusters using spectral clustering, starting with eight clusters (SC-8, *left*). Subdivision of each cluster obtained from SC-8 into two subunits using Spectral Clustering (SC-8 + SC) yields the *middle* diagram. The *right* diagram represents the ground truth where beads are grouped into subunits by mapping them to the nearest α-carbons in the atomic structure (PDB: 4a0v). Corresponding SOM + k-means results are shown in Fig. S2A. The clusters are coloured as in panel A.

distribution based on the EM density map voxels as described in Section 2.1. Information on connectivity is included in the original algorithm, but, in our adaptation, we use the ANM to derive the Hessian H that also accounts for the stiffness of inter-node couplings (Atilgan et al., 2001; Bahar et al., 2017).

Structural superposition of the TRN beads and the original density map (Fig. 2B *left*) as well as the all-atom structure deposited in the PDB (Fig. 2B *right*) (Cong et al., 2012) verified that a network of about $M = 3,000$ nodes adequately represented the density data for CCT-AMP-PNP and provided a good description of overall architecture. A few nodes that protrude with respect to the density maps, often located at highly sparse, peripheral positions with elongated shapes, could be attributed to artefacts from averaging of flexible/disordered substrates, termini and apical loops with variable occupancies and conformations (Cong et al., 2012). Such nodes were removed from the model as their large swinging motions could obscure the cooperative motions of the multimer, a phenomenon known as the tip effect (Lu et al., 2006). The refined network had $M = 2,908$ nodes.

The level of granularity adopted here for the TRN is comparable to that of recent cryo-EM structure analyses (Sorzano et al., 2019; Vilas et al., 2018). On the other hand, it is lower than that of conventional single-node-per-residue ANM, which would contain

more than 8,000 nodes based on 16 subunits of >500 residues each. Previous work demonstrated that the ANM is scalable, i.e. one can retain the accuracy of the results even with 1-2 orders of magnitude fewer nodes provided that the cut-off distance for defining the network connectivity and corresponding Hessian is suitably rescaled (Doruker et al., 2002). This feature was borne out by previous application to bacterial chaperonin GroEL, which showed that soft-partitioning of the structure into 10^2-10^3 nodes (clusters of spatially neighbouring residues) almost invariably retained GroEL global dynamics and allosteric signalling properties otherwise described by a full-residue (~8,000 nodes) model (Chennubhotla and Bahar, 2006). Based on these considerations, in order to adapt the TRN algorithm to the ANM analysis of cryo-EM data, we rescaled the cut-off distance using the relationship between the network model granularity and the optimal cut-off distance (Doruker et al., 2002). This led to a TRN cut-off distance of 20 Å, as opposed to 12–15 Å in the conventional ANM (Fig. S3A).

The TRN representation is agnostic to chain connectivity, or to the identity of the individual subunits. As the network contains intertwined domains which may not lend itself to identification of subunits by standard clustering methods, we adopted two topology-based methods, namely, spectral clustering (SC; implemented in scikit-learn (Pedregosa et al., 2011)), and Self-Organizing

Maps (SOM; implemented in NeuPy (Shevchuk, 2015)) for identifying subunit boundaries. As shown in Fig. 2C, SC organized the CCT nodes into four well-defined clusters in each ring in the first step. We then used a multi-step procedure guided by the following features: (i) CCT-ATP subunits are spatially grouped in pairs (Cong et al., 2012), (ii) subunits also move in pairs as observed in the ANM analysis (see Section 3.2), and (iii) intermediate and equatorial domains of neighbouring subunits are tightly connected. Our multistep procedure led to 16 clusters that closely approximated the 16 subunits of EMD-1961 (Fig. 2C, *middle*).

The SOM method generated similar clusters, in support of SC results, but a few clusters contained disjoint nodes (see *red arrow* in Fig. S2A), so we adhered to SC results. Besides, SC partitions the beads based on the eigenvectors of the graph Laplacian, which bears close resemblance to the Gaussian Network Model (GNM) (Bahar et al., 1997) mode decomposition. The SC result is therefore conceptually rational and consistent with the ENM framework. Finally, either algorithm, SC or SOM, was verified to yield an accuracy of ~90% upon comparison of the identified partitioning (of nodes into subunits) to that obtained by mapping the nodes to the nearest α-carbon atoms in the deposited all-atom structure (*ground truth* in Fig. 2C, *right*).

Overall, this analysis showed that topology-based hierarchical (multi-step) clustering guided by ANM dynamics accurately discerned the subunits. Tight packing of equatorial domains of neighbouring subunits made this a nontrivial process. Since our method is guided by ANM dynamics and SC effectively groups nodes with similar motions together, the generated clusters are dynamics-based domains, which may not necessarily coincide with protomers. However, they are functionally important, and therefore, such a partitioning of the structure is useful. Furthermore, successively smaller domains can be identified by a multistep approach.

Two additional tests were performed for assessing the use of TRN for modelling CCT equilibrium dynamics. First, we evaluated the GNM-predicted mean-square fluctuations (MSFs) of residues for CCT bound to an ATP-analogue, TRiC-AMP-PNP (EMD-1961), using its TRN model. Comparison to x-ray crystallographic B factors observed for CCT bound to another ATP analogue (ATP-γ-S; PDB: 2xsm) (Munoz et al., 2011) showed a good level of agreement despite the low resolution (5.5 Å) of the x-ray structure and the use of electron density maps for constructing the TRN (Fig. S4). Second, we compared the global mode spectra predicted for the TRN representation of TRiC-AMP-PNP, with that computed using the GNM and ANM for its all-atomic model (PDB: 4a0v), which yielded reasonable agreement (Fig. S5). These results confirmed that the TRN provides an adequate framework for analysing the collective dynamics of CCT.

3.2. The architecture of ATP-analogue-bound CCT favours pairwise couplings between intra-ring neighbours and en bloc motions of selected subunits

To explore the machinery of CCT on a global scale, we analysed as a first step the softest 20 non-zero modes accessible to TRiC-AMP-PNP modelled above. The soft modes usually facilitate, if not drive, the collective rearrangements (of all subunits) relevant to biological function (Bahar et al., 2017); therefore it is of interest to determine which particular subunits exhibit a high predisposition to undergo a lid closure (e.g. passage to TRiC-ADP-AlFx in Fig. 1) accompanying ATP hydrolysis.

The mean-square fluctuations (MSFs) of the subunits and the cross-correlations (or *covariance*) driven by these softest modes are presented in Fig. 3. The surface in panel A is colour-coded by MSFs; the regions undergoing the smallest motions are in *blue*, and

intermediate and higher mobility regions in *green* and *orange*, respectively. The equatorial domains are highly stable and closely coupled between the two rings, while the apical domains exhibit the highest mobility consistent with previous computations (Chacon et al., 2003) and experiments (Cong et al., 2012).

Fig. 3B shows the subunits that exhibit the highest tendency to move in the ATP analogue-bound structure driven by the softest 5 (*green*), 10 (*orange*) and 20 (*blue*) modes (see details for modes 1—5 in Fig. S6). In principle the peaks in the slowest modes would indicate the subunits that undergo the earliest movements. Here CCT5, CCT4 and CCT2 in the *cis* ring (*upper panel*), and CCT4, CCT2 and CCT3 in the *trans* ring (*lower panel*) exhibit the highest peaks. The *trans* ring is less constrained than the *cis* ring (compare the ordinate values in the two panels), which could be attributed to stabilization induced upon ATP analogue or substrate binding to the upper ring subunits.

We have also verified that the MSFs and cross-correlations driven by the global modes were insensitive to the inter-node distance-dependency of the spring constants in the ENM. As shown in Figs. S3B—C, conventional ANM (with uniform spring constant for all pairs closer than 15 Å), the ENM with force constant exponentially decaying with distance (Hinsen, 1998), and a parameter-free ANM where the spring constant scales inversely with the distance squared (Yang et al., 2009a) yielded comparable results to those with uniform spring constant for all pairs closer than 20 Å. The major difference was a few excessively high peaks predicted by parameter-free ANM, which signals the possible underestimation of inter-node couplings at specific loci in the latter.

The covariance maps in Fig. 3C—D provide information on the level of coherence/coupling between nodes within individual subunits (*diagonal elements*) and across subunits (*off-diagonal*). The *lower-left* and *upper-right* quadrants correspond to the intra-ring correlations for the upper and lower rings, respectively; and off-diagonal quadrants display the inter-ring correlations. The boundary between the two rings is shown by the *white dashed lines*. *Dark red* and *dark blue* entries refer to the strongest correlations (large movements in the same direction, usually intra-subunit) and anticorrelations (large movements in opposite directions, inter-subunit). Subunits a_2/CCT4, a_8/CCT5 and a_1/CCT2 exhibit the largest and most coherent (*en bloc*) movements in the upper ring; as do subunits a_4'/CCT3, a_2'/CCT4, and a_1'/CCT2 in the lower ring consistent with their distinctive mobilities in panel **B**. A closer look at the diagonal 2 × 2 blocks indicates that the subunits tend to move in pairs. This feature is observed in both upper and lower rings, suggesting that such couplings are not associated with nucleotide binding, but with the asymmetric hetero-octameric nature of the rings.

The covariance map in Fig. 3C reflects both the sizes of subunit motions and their relative orientations (correlation cosines). Further examination of purely orientational correlations (upon normalization with respect to subunit-averaged MSFs; see **Supplemental Methods**) in Fig. 3D confirms the strong coupling between sequential pairs (a_1, a_2), (a_3, a_4), (a_5, a_6), (a_7, a_8) in the *cis* ring, and their counterparts in the *trans* ring. We note the motions of each pair are anticorrelated with respect to others in the same ring, and especially the pairs at the opposite face (*dark blue regions* in the *lower left quadrant*). Similar anticorrelations can be discerned in the lower ring (*upper right quadrant*). Therein, the pair a_5' and a_6' is distinguished by a strong coupling. Finally, the *light green-yellow* diagonal along the inter-ring quadrants shows the couplings between neighbouring subunits across the rings. Calculations confirmed the reproducibility of this map using either uniform or distance-dependent (exponential) spring constants, but not with parameter-free ANM (Fig. S3C).
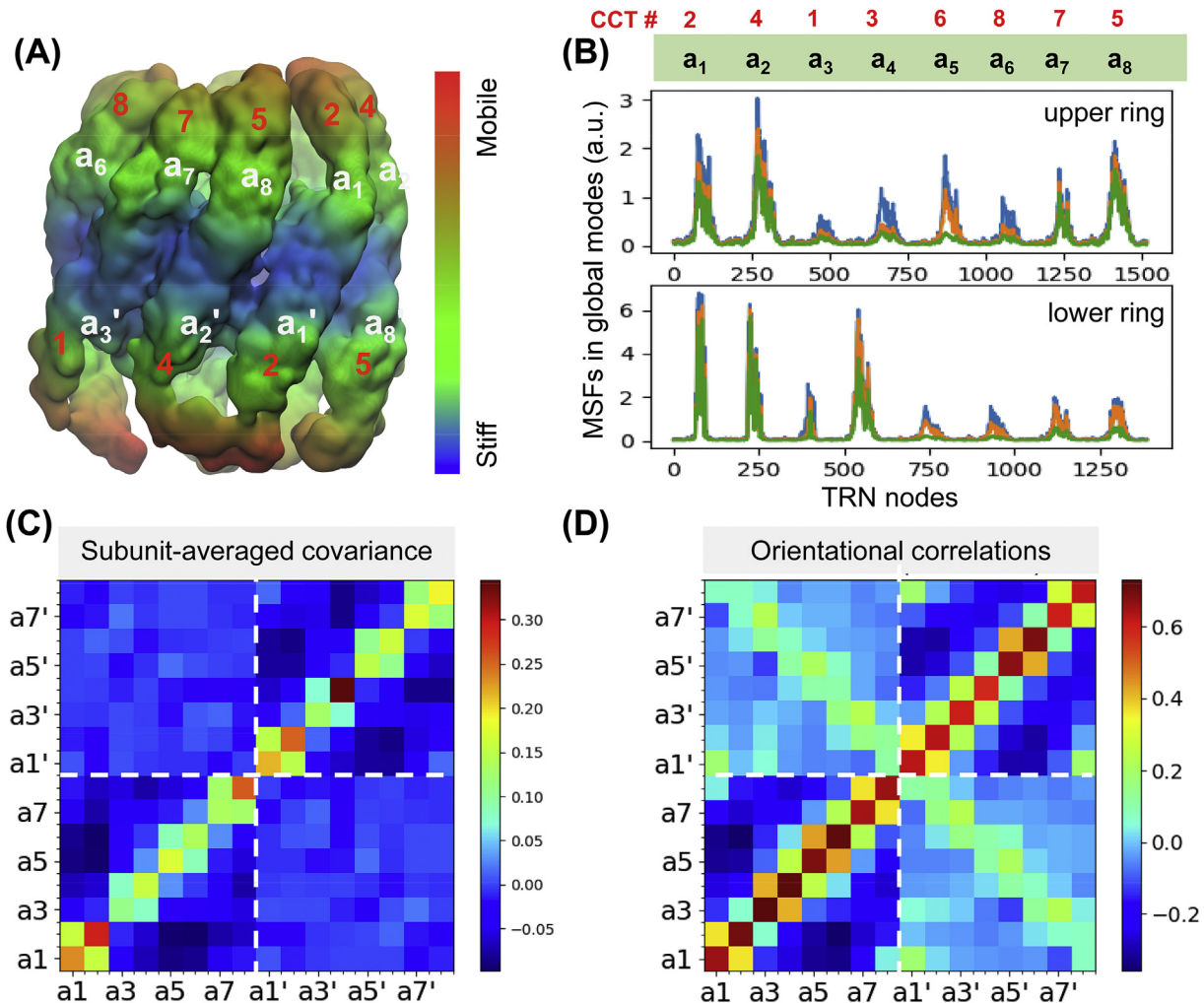
Fig. 3. Global dynamics of TRiC-AMP-PNP. (A) Results from ANM analysis of the TRN (based on EMD-1961), displaying the architecture colour-coded by the MSFs of nodes (in *surface representation*; *blue*: most rigid; *orange*: most mobile) in the softest 20 modes. **(B)** MSFs of the subunits as driven by the subsets of 5 (*green*), 10 (*orange*) and 20 (*blue*) softest modes. Note the high mobility of subunits a_1, a_2 and a_8 in the upper ring, and those of a_1', a_2' and a_4' in the lower ring. Two sets of labels are listed in panels A and B: a_1-a_8 (or a_1'-a_8' for the lower panel) and corresponding biological subunit names CCT1-8 (in *red*). **(C)** Covariance between the global motions of the subunits based on the softest 20 modes. Strong coupling between pairs of adjacent subunits on a given ring are seen. **(D)** Orientational correlations between the global movements of the subunits (same as covariance but normalized with respect to MSFs).

3.3. Differential mobilities of the subunits in the ATP-bound form reveals the predisposition of CCT4 and CCT5 to enable functional changes

By definition, the collective modes of motions predicted by the ANM describe the thermal fluctuations accessible to an equilibrium structure under physiological conditions (Atilgan et al., 2001; Bahar et al., 2010b). They also underlie the adaptability of the structure to changes induced by intermolecular interactions (Haliloglu and Bahar, 2015; Zhang et al., 2020b). Perturbations caused by any energy input, for example ATP hydrolysis, would "push" the equilibrium structure away from its energy minimum along the most easily accessible modes, starting from *mode 1, 2*, etc. The individual modes represent axes of overall reconfiguration away from the equilibrium state in the multidimensional conformational space, and many structural changes are indeed effectuated by a combination of selected modes (Meireles et al., 2011). Among the accessible modes, those requiring a lower energy ascent (lower frequency or smaller uphill curvature along the energy landscape) will be naturally preferred in response to perturbations, e.g. ATP binding or hydrolysis. These soft modes naturally show more

significant movements than the ones with higher frequency, in response to the same amount of energy allocated to each mode as dictated by the equipartition law. Figs. 3B and S6C show that subunits a_2/CCT4, a_8/CCT5 and a_1/CCT2 in the upper ring and a_1', a_2' and a_4' (CCT2, CCT4 and CCT3) in the lower ring are predisposed to exhibit the largest movements in the softest modes. As such, these subunits are suggested to be most readily adaptable to subsequent structural changes that the ATP-bound state will undergo, i.e. those taking place upon ATP hydrolysis, if not driving the changes that facilitate hydrolysis. We also note that subunit a_4' is distinguished by a large *en bloc* movement (Fig. 3C); and adjacent pairs in each ring tend to move in tandem (Fig. 3C–D).

### 3.4. Conformers along the allosteric cycle exhibit state-dependent dynamics manifested by differential sequences of subunit motions in soft modes

The results in Subsections 3.2-3.3 are uniquely defined by the architecture of CCT bound to non-hydrolysable ATP analogues. Computations performed for TRiC/CCT structures in different states along the allosteric cycle revealed, on the other hand, the high

mobility of other subunits, in a *state-dependent* manner. Fig. 4 displays the results for three additional conformers visited at different stages of the allosteric cycle: (**A**) TRiC resolved (Zang et al., 2016) in a nucleotide-partially-preloaded (NPP) state equivalent to the apo state (*green shaded* in Fig. 1), (**B**) another newly resolved structure at relatively low (0.1 mM) ADP-AlFx level, called conformation 1 (Jin et al., 2019a), representative of an intermediate during ATP binding; and (**C**) the TRiC structure at 1 mM ADP-AlFx (Cong et al., 2012) which may be viewed as an intermediate stabilized

during hydrolysis (*purple shaded* in Fig. 1). The plots *on the right* display the subunit displacements driven by the softest modes of motion. Peaks therein indicate the subunits with highest propensity to undergo conformational changes, like Fig. 3B for the ATP-analogue-bound TRiC (except for displaying the results here for the two rings by consecutive residue numbers indicated by the respective *blue* and *grey* bars along the upper abscissa).

Based on the mobility profiles driven by the softest modes, we conclude that the ability of subunits to readily undergo large-scale
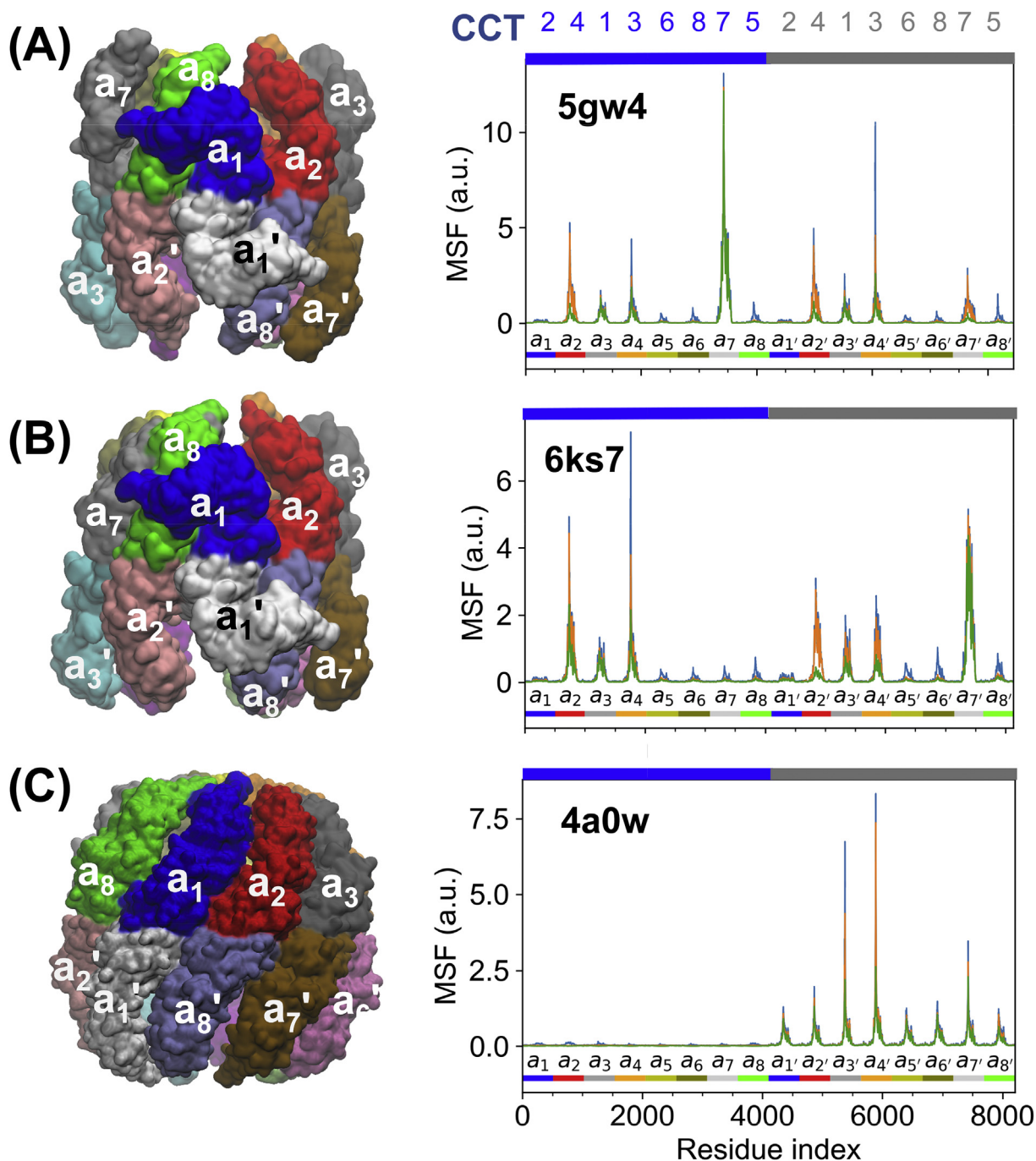
Fig. 4. Different conformational states of TRiC visited during its allosteric cycle and corresponding mobility profiles. Panels **A** and **B** display results for intermediate structures between apo and ATP-bound forms, namely (**A**) ADP partially preloaded form (PDB: 5gw4) and (**B**) Conformation 1 at 0.1 mM ADP-AlFx concentration (PDB: 6ks7). Panel **C** shows TRiC at 1 mM ADP-AlFx (an intermediate stabilized during hydrolysis; PDB: 4a0w). The MSF curves on the *right* display the movements driven by 5 (*green*), 10 (*orange*), and 20 (*blue*) ANM modes. Labels along the upper abscissa indicate the subunit numbers (CCT1-8). The *colour code* along the lower abscissa indicates the consecutive subunits, with the colour scheme matching the upper ring on the *left*.

movements driven by these modes depends on the conformational state. Mainly, the subunits that are predicted to move first are: CCT7 (*cis*/upper ring) in (**A**); subunits CCT7 (*trans*), followed by CCT4 (*cis*) and CCT3 (*cis*) in (**B**); and CCT3 (*trans*) followed by CCT1 (*trans*) in (**C**).

Complemented by the results presented in Fig. 3 for the ATP analogue-bound state, which is an intermediate state between conformers (**B**) and (**C**), the present analysis shows that as the overall eukaryotic chaperonin proceeds from apo to ATP-bound, and then to ADP-bound state, different subunits exhibit a high propensity to undergo conformational rearrangements. Subunits CCT6, CCT8 and CCT3 reportedly bind nucleotides at low concentrations (Jin et al., 2019a). The former two subunits (as well as CCT2 and 5) are highly stable in the low nucleotide concentration conformers (**A**) and (**B**) in Fig. 4; they practically do not move in the soft modes; but CCT3 retains some intrinsic flexibility. Our analysis suggests that the initial conformational change toward the ATP-bound state is facilitated by consecutive motions of CCT7 (both rings), and then CCT3 and CCT4 (both rings) (Fig. 4A—B); the conformational changes accompanying the early stages of hydrolysis in ATP-bound TRiC are mainly undergone by CCT5, CCT4 and CCT2 in the *cis* ring, and CCT2 and CCT4 in the *trans* ring (Fig. 3B), followed by the concerted motions of CCT3 and CCT1 (*trans*) that ensure the *trans* ring closure/opening near the completion of hydrolysis (Fig. 4C).

3.5. Soft modes define pre-existing conformational paths that comply with the allosteric transitions of CCT

The above analysis provided information on specific subunits that are distinguished by their high conformational mobility/adaptability at various stages of the chaperonin cycle. We now examine the mechanisms of their motions.

Collective motions of ATP-analogue-bound TRiC. We first focus on the softest ten modes intrinsically favoured in this state, toward gaining mechanistic insights into ATP-regulated events. These modes induce distributed motions across the subunits. Due to the heteromeric nature of the rings, they depart from perfect symmetry. Yet, we discern movements reminiscent of cylindrical symmetry, such as radial breathing, biaxial stretching/contraction across the circular cross-section, or overall rolling movements of subunits within rings, schematically depicted in Fig. 5A—C. For example, *mode 6* induces a concerted opening/closing or bending of the eight apical domains that approximates a breathing motion when viewed from the top, but the circular symmetry is broken due to the heterogeneity of the subunits (Fig. 5D and Movie 1A).

The most cooperative modes, *modes 1 and 2,* tend to deform the circular cross-section into an ellipse by bringing closer together the pairs of subunits (a_1/CCT2, a_2/CCT4) and (a_5/CCT6, a_6/CCT8) that are oppositely located on the ring, while (a_7/CCT5, a_8/CCT7) and (a_3/CCT1, a_4/CCT3) splay apart, and *vice versa* (see Fig. 5E and Movie
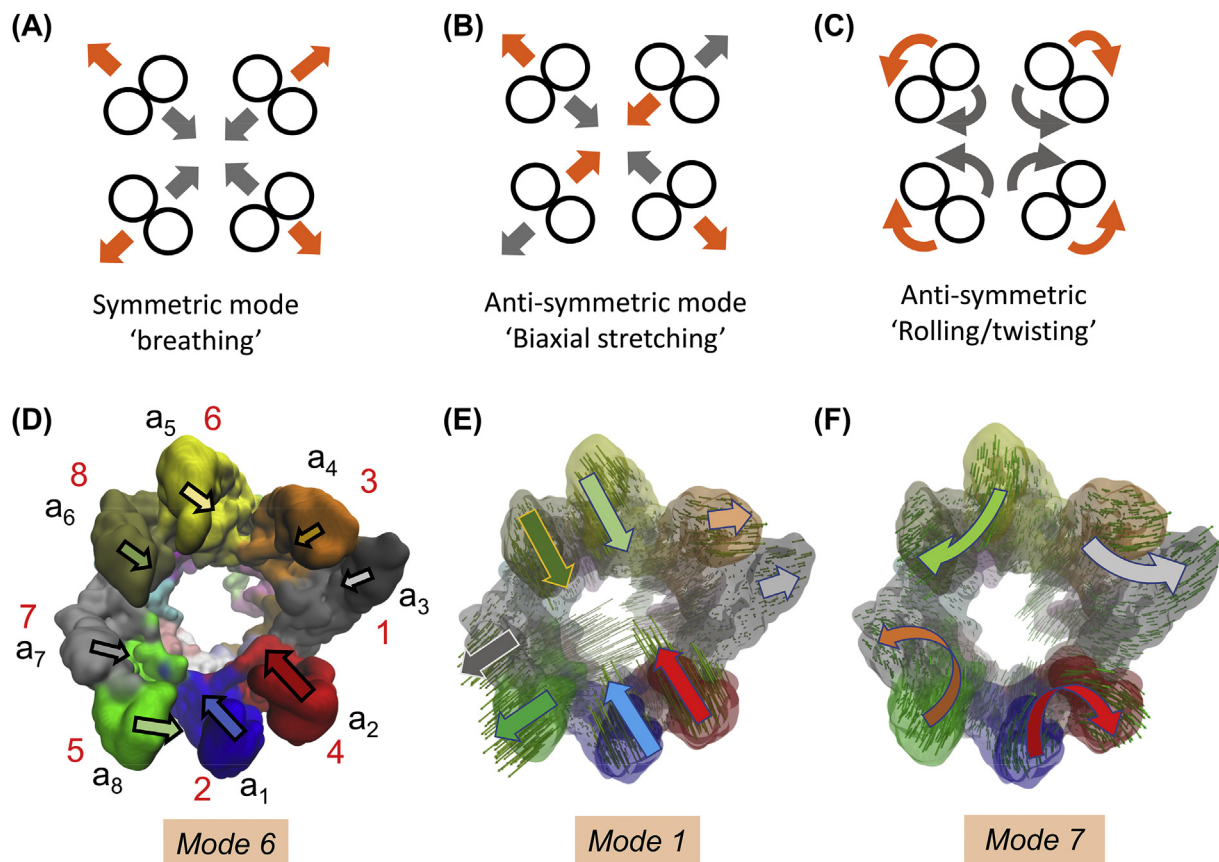
Fig. 5. Schematic description of symmetric or antisymmetric movements approximated by CCT-ATP (EMD-1961). (**A-C**) Symmetric and anti-symmetric movements for circularly symmetric shapes. Note that *grey* and *orange* arrows indicate the opposite direction movements undergone by the subunits in those modes, i.e. the motions alternate between symmetric expansion and compression in **A**, stretching and contraction along orthogonal directions in **B**, and opposite direction rotations in **C**. The latter two are double degenerate in a circularly symmetric structure. (**D**) *Mode 6* approximates the symmetric breathing mode, although it is deformed due to the heterogeneity of the subunits. The arrows display the approximate orientation and size of the motions of individual subunits in the upper ring (labelled). See also Movie 1A. (**E**) *Mode 1* is reminiscent of anti-symmetric stretching depicted in (**B**); *mode 2* shows similar features. (**E**) Modes 7 (and 8, not shown) approximate anti-symmetric rolling/twisting illustrated in (**C**).

2A). *Modes 7 and 8* exhibit a concerted rolling of subunit pairs (Fig. 5F and Movie 2B). The latter two modes have similar frequencies (Fig. S2B), as do the pairs of *modes 3—4* (Movie 1C), and *9—10* (Movie 2C), suggesting that they are almost doubly-degenerate by virtue of the close-to-cylindrical symmetry of the architecture (Na and Song, 2016). Such complementary modes may play essential roles in enabling functional transitions, similar to those observed in the cylindrically symmetric bacterial chaperonin GroEL and icosahedral viral capsids (Chennubhotla et al., 2005; Rader et al., 2005; Yang et al., 2009c).

Movies 1B-C illustrate the motions driven by *modes 4* and *5*. Both modes are distinguished by large coupled movements of a_1 and a_2 to enable the closure/opening of the upper ring by bending of the apical domains. In *mode 5*, their motion is accompanied by in-phase opening/closing of the adjacent pair a_8-a_7, while other subunits are practically rigid. In *mode 4*, on the other hand, the two pairs undergo out-of-phase opening/closure, accompanied by alternating anticorrelated movements of a_4 and a_6 with respect to both pairs. Fig. S6A depicts the subunit movements in *mode 4*, and Fig. S6B those in *mode 5*. Figs. S6C—D display the contributions of *modes 1—5* to global movements and orientational cross-correlations, respectively. Fig. S6C suggests a sequential mechanism in the ATP-bound form wherein lid closure, presumably driven by an ATP-affinity gradient-induced power stroke (Cong et al., 2012; Reissmann et al., 2012), is initiated by the movements of a_2, a_8 and a_1 (*modes 1—2*) succeeded by a_7 (*modes 4—5*). Notably, *mode 5* plays a major role in driving the displacements of a_8 and a_7.

Transitions are enabled by movements along soft modes. To test if these soft modes intrinsically accessible to the upper ring facilitate, if not enable, the transition toward the ADP-bound form where the upper ring is closed, we performed an adaptive ANM (aANM) analysis (Yang et al., 2009c) using the 1 mM ADP-AlFx-bound, hydrolysis-state CCT (PDB: 4a0w) as target structure (Fig. 6A and Movie 3). This approach generated a sequence of conformers whose RMSD from the target gradually decreased at each cycle (Fig. 6A—B). Larger decreases took place at earlier cycles that recruited global modes, supporting their dominant role in enabling this transition. The generated sequence of conformers is shown by the *purple dots* in the conformational landscape in Fig. 6B.

We also repeated the aANM analysis for the remaining core transitions of the functional cycle (*light blue* and *dark blue arrows* in Fig. 1C; *same colour dots* in Fig. 6B) using the all-atom models (PDB: 4a0w, 4a13 and 4a0o) and confirmed that all these transitions were also largely driven by global modes (Fig. 6B).

To explore the conformational space in an unbiased way, we employed the ClustENM approach (Kurkcuoglu et al., 2016) described in Methods subsection 2.2. We observed that the ATP analogue-bound structure explored a region in the vicinity of the starting point on a thin plane in the principal component space similar to that sampled during aANM (Fig. 6B). This confirms that the aANM transitions are intrinsically accessible to the TRN derived from the ATP analogue-bound state. Further, the apo state lies within the ClustENM-sampled region, revealing that there is an easily accessible (via soft modes) path intrinsically favoured by the overall architecture between the apo and ATP-bound conformers.

3.6. TRiC trans ring closure is enabled by collective bending of the apical domains intrinsically accessible to individual subunits

Finally, toward exploring the ability of the hydrolysis transition state analogue-bound structure (1.0 mM ADP-AlFx-bound CCT;
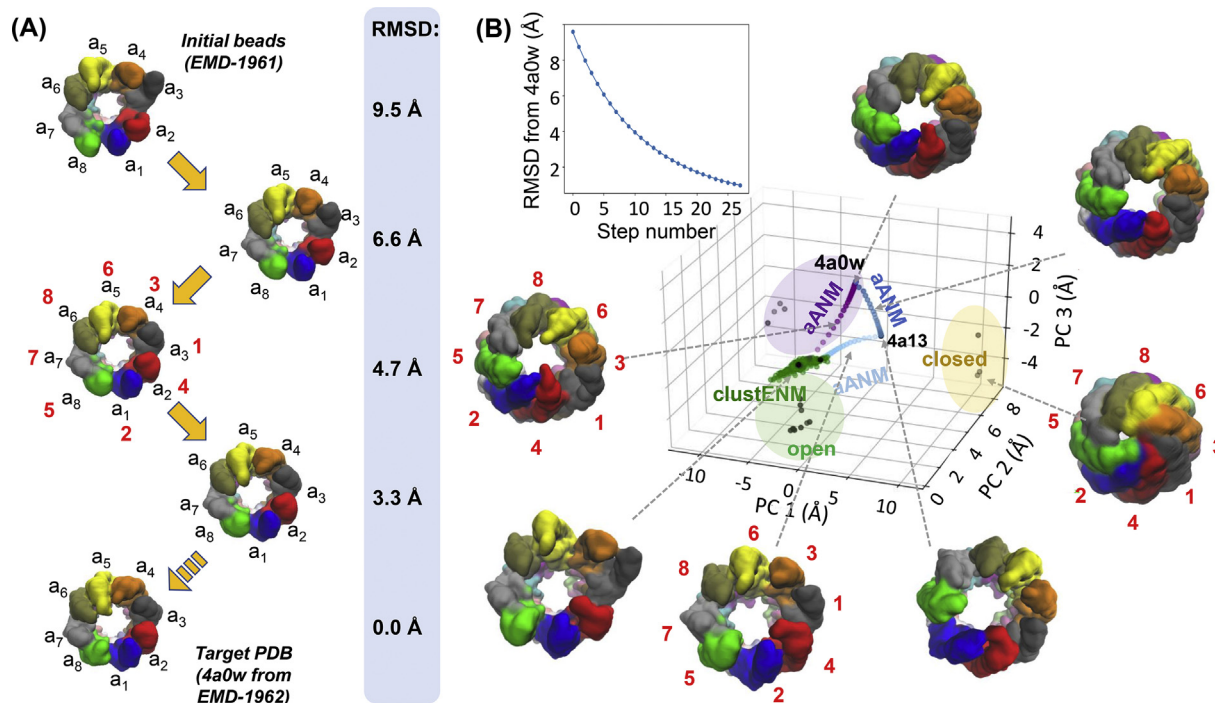

Fig. 6. Transitions between open and closed ring forms are enabled by soft modes. (**A**) The snapshots depict a path from the fully open (EMD-1961) ATP analogue-bound CCT to the asymmetric closed (EMD-1962) hydrolysis transition state analogue-bound form of CCT, driven by collective motions along soft mode coordinates. See corresponding movie, Movie S4, in the Supplementary Material. The conformational changes are generated using adaptive ANM (aANM). The conformation is shown after every 4 steps up until step 12 (4th structure), which reaches an RMSD of 3.6 Å (as compared to the original RMSD from target of 9.5 Å), as well as the target structure. (**B**) Comparison of aANM and ClustENM conformers with experimental structures via projection of generated conformers onto the 3D subspace defined by experimental structures in Fig. 1C (*main panel*). The upper plot shows the decrease in RMSD with respect to the target structure (PDB: 4a0w; *inset*). Three aANM runs are shown in *purple, dark blue* and *light blue*, and ClustENM conformers are *green*.

PDB: 4a0w) to sample the symmetrically closed structures (PDB: 3iyg, 4v8r and 6ks6; *yellow-shaded* region in Fig. 1), we examined the ANM-predicted soft modes intrinsically accessible to the this asymmetric conformer. Movies 4-6 display the softest three modes. In all three cases, consistent with the suppressed mobility of the *cis* ring seen in Fig. 4C, the *cis* ring subunits retained their closed form with minimal change in conformation; whereas the *trans* ring subunits undergo highly cooperative motions. These motions are enabled by hinge-bending at their intermediate domain, which allows for the apical domains to come into close proximity. *Modes 1—2* induce an alternating biaxial stretching/extension in the apical domains located at diametrically opposed ends of the ring when viewed from the *bottom* (Movies 4 and 5); whereas *mode 3* induces a highly symmetric motion where all apical domains concertedly close and open the ring (Movie 6). The apical and equatorial domains act as the "lid" and "stave" for the complex.

We further analysed the intrinsic dynamics of the subunits,

which enables the bending of the apical domains. ANM analysis for an isolated subunit showed that subunit bending that cooperatively enables the closure of the octameric ring is the softest mode (*mode 1*) favoured by the inter-residue contact topology of the subunit. This movement is mediated by a set of hinge residues in the intermediate domain (e.g. L169-V171 and I347-A357 in subunit a_1 (CCT2); see Fig. 7A). The second softest mode induces an overall torsion, mediated by several residues acting as anchors (labelled in Fig. 7B), which is also observed in structural studies (Cong et al., 2012; Jin et al., 2019a). Movies 7A and B display the motions of the subunit, modelled as a TRN, along these two modes. Finally, the inter-residue cross-correlation map for the subunit reveals the critical role of the intermediate domain hinge residues which divide the subunit into two blocks (equatorial and apical domains) that undergo anticorrelated motions (Fig. 7C).

In order to gain further insights into the biological significance of specific residues, we focused on the subunit a_1 in TRiC-AMP-PNP,
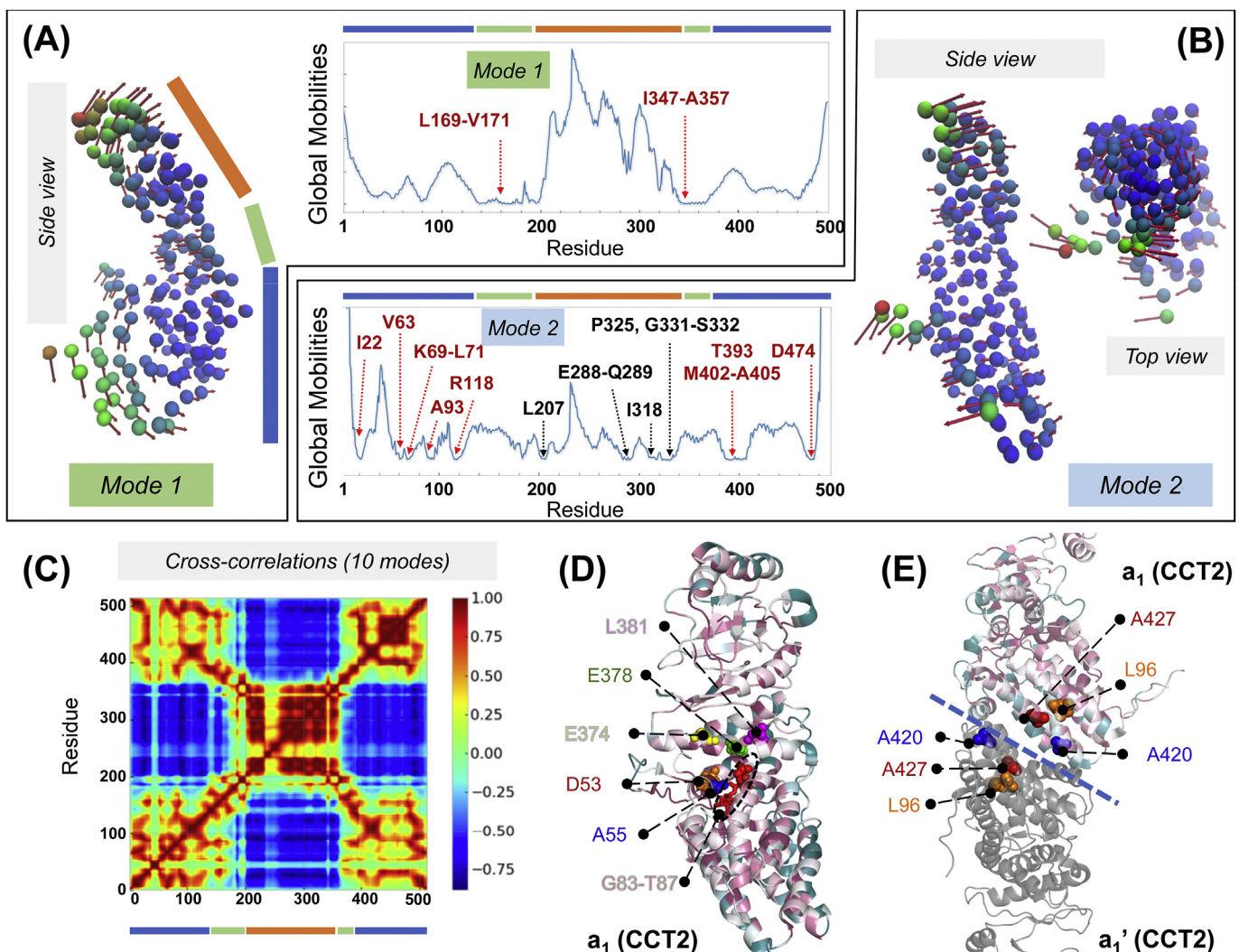
Fig. 7. Intrinsic dynamics of a single subunit in CCT-ATP. (A, B) ANM *modes 1* and *2* of subunit a_1 from the TRN model of TRiC-AMP-PNP. Nodes are colour-coded based on their mobility. Side view arrows on the *left* show the movements in mode 1, and side and top views on the right show the movements in *mode 2*. Corresponding mobility profiles are shown in the middle two panels. Note that the intermediate domain (indicated by the *green bar*) exhibits minimal movements consistent with its hinge role. **(C)** Intrasubunit cross-correlations computed for subunit a_1 using the GNM, coloured from highly anticorrelated (*dark blue*) to highly correlated (*dark red*). Domains are indicated by *blue, green* and *orange* bars. **(D)** Cartoon representation of subunit a_1 (PDB: 4a0v), colour-coded by conservation scores from 1 (*white*) to 9 (*purple; most conserved*) predicted by Consurf (Ashkenazy et al., 2016). Potential functional sites predicted by *DynOmics* are labelled. Red sticks represent the ATP-binding/hydrolysis motif GDGTT, which consists of fully conserved residues G83-T87 (*red sticks*). **(E)** The transverse dimer of subunits a_1 and a_1' with a_1 subunit coloured by conservation score. Potential functional sites predicted by *DynOmics* with conservation scores ≥ 7 (L96, A420 and A427) are shown by spheres. The blue dashed line indicates inter-ring interface.

and checked whether the hinge and functional sites predicted by *ProDy* (Bakan et al., 2011, 2014) and *DynOmics* (Li et al., 2017), respectively, were evolutionarily conserved. The hinge residues (V171, Q368, D371-E374, H378) in the softest GNM mode indeed had conservation scores $\geq$ 7 evaluated by the ConSurf server (Ashkenazy et al., 2016). As to the functional residues predicted by *DynOmics* using the slowest two modes from the PCA_NEST algorithm (Yang et al., 2009b), D53, A55, D84, E374, H378 and L381 were found to be conserved and are shown as spheres in Fig. 7D. Specifically, D84 participates in a highly conserved ATP-binding/hydrolysis motif, GDGTT, next to the equatorial domain. This motif is known to be crucial for the function of *both* bacterial (GroEL) and mammalian (CCT) chaperonins (Kabir et al., 2011). Thus, binding of ATP at this region would clearly impact the motions of the apical domain relative to the equatorial domain.

Overall, this analysis shows that the closure of the *trans* ring of the ADP-bound chaperonin is favoured by its overall architecture, enabled by the intrinsic dynamics of the individual subunits. While several structural changes along the allosteric cycle evolve by sequential, highly asymmetrical motions, it is interesting to note that this particular step simultaneously engages all eight subunits in the ring, at least at its initiation (driven by the slowest modes).

Further analysis of the coupled dynamics of the transverse dimer of subunits a_1 and a_1', presented in Movies 8 and 9, showed that the apical domains gain greater relative mobility, especially at the protrusion helices, when the equatorial domains are constrained by inter-ring contacts. Such large motions have been noted by comparing the subunit conformations in the open (ATP-free)

and closed (ATP-bound) forms (Skjaerven et al., 2015). The "bending" and "twisting" modes are preserved and even amplified at the apical domains upon complexation of the equatorial domains, mediated by highly conserved residues, L96, A420 and A427, at the dimeric interface (Movies 8-9 and Fig. 7E).

3.7. Markovian modelling of information flow between subunits reveals the differential rate of signal transfer across subunits and two fluxes of signal flow

Network modelling permits us to evaluate the hitting time, $H(j, i)$, a quantitative measure of the efficiency of signal propagation across subunits in supramolecular systems, based on a Markovian stochastics model (Chennubhotla and Bahar, 2007). $H(j, i)$ is the expected number of steps required to transmit a signal from node i to node j, averaged over all possible connections/paths. It is asymmetric, i.e. $H(j, i)$ is not necessarily equal to $H(i, j)$. Node i serves as *broadcaster* of signal/perturbation, and j as *receiver*. We evaluated the subunit-averaged hitting times, $\overline{H}(b, a)$ between each pair of subunits a and b (see Supplementary Methods) for the 16 CCT subunits in the ATP analogue-bound state, presented in Fig. 8A. The map is colour-coded from *blue* (fast communication, short hitting time) to *red* (slow communication, long hitting time). Rows indicate the ability/efficiency of subunits to receive signals; columns indicate their ability to send signals. Diagonal elements refer to the communication within subunits, which are always much faster than those between subunits (off-diagonal elements).
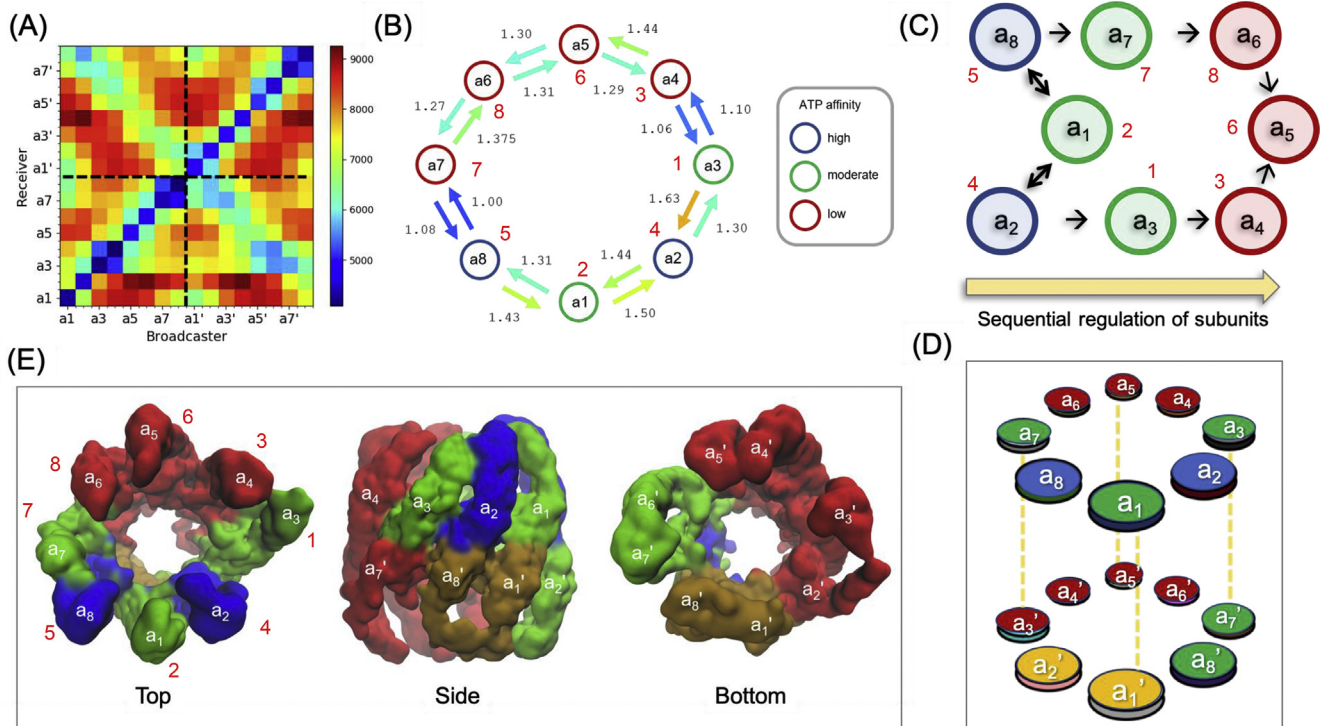
Fig. 8. Intersubunit signal propagation behaviour of TRiC/CCT in the presence of ATP analogues. (A) Efficiency of information flow between subunits, represented by the subunit-averaged hitting time matrix. Each entry represents the hitting time from a broadcaster (horizontal axis) to a receiver (vertical axis), colour-coded from *blue* (fast communication) to *red* (slow communication). **(B)** Diagram illustrating the subunit-averaged hitting time between neighbouring subunits in the upper ring. Arrows are coloured based on the hitting times in the matrix shown in **(A)**. Circles stand for the eight subunits, coloured according to their ATP affinities (Reissmann et al., 2012). The numbers are the hitting times relative to the fastest communication event $\overline{H}(a_7, a_8)$. **(C)** Sequence of events deduced from combined ANM and Markovian stochastics analysis. In *dark blue* are the subunits that initiate ATP-regulated events, first communicated to subunits in *green*, followed by subunits in *red*. Note that the time evolution (indicated by the *arrow* from *left* to *right*) conforms to ATP-binding affinities observed experimentally. Double headed arrows are shown for pairs with comparable hitting times in both directions. **(D)** Information flow across the entire chaperonin, colour-coded by the sequence of events. Subunits a_7' and a_8' are the first responders to a_2. The subunits a_1' and a_2' coloured *yellow* show an intermediate behaviour. **(E)** *Top, side* and *bottom* views of the structure colour-coded by the sequence of events depicted in panel **D**.

We also see that the hitting times between neighbouring subunits (*blue blocks* near the diagonal) are shorter than those between non-neighbouring subunits. The diagonal from *top left* to *bottom right* (secondary diagonal) corresponds to inter-ring neighbours. Certain subunits (e.g. a_1/CCT2 and a_2/CCT4) exhibit low signal-receiving rates from their neighbours. In contrast, a_8/CCT5 is distinguished by its highly efficient communication with a_7/CCT7. Likewise, a_3/CCT1 communicates efficiently with a_4/CCT3.

The emerging scheme of intra-ring allosteric signal transduction is depicted in Fig. 8B. The subunits (circles) are colour-coded by their experimentally observed propensity to bind and presumably hydrolyse ATP (see Discussion). The signal transduction efficiencies between neighbouring subunits are indicated by colour-coded arrows. These results, combined with intrinsic mobilities described in Fig. 3, lead to the following picture for possible time evolution of ATP-regulated signal transduction events across the upper ring subunits. The most adaptable/mobile subunits a_8/CCT5 and a_2/CCT4, followed by a_1/CCT2, are likely to lead the succession of events based on their intrinsic high mobility (see e.g. Fig. 3B or the diagonal elements of the covariance matrix in Fig. 3C). CCT5 and CCT4 are most likely to transduce signals to their neighbours a_7/CCT7 and a_4/CCT3, which in turn propagate the signals to a_6/CCT8 and a_5/CCT6. The sequence of parallel events is described by the scheme in Fig. 8C. Subunits are depicted by circles colour-coded by ATP-binding affinity and arrows indicate preferred information flow directions. Subunits with higher ATP-binding affinity appear to lead the process, the flow of information taking place from high- to low-affinity subunits.

The above analysis describes the flow of information within the *cis* ring. Of interest is to observe how information flows in the *transverse* direction, i.e. between rings. The latter can be inferred from the elements around the secondary diagonal in Fig. 8A. The *blue* (fastest signalling) entries in the *upper left quadrant* show that a_8' immediately receives signals from a_2, consistent with their juxtaposition; and a_1' receives from a_1, and a_2' from a_8. Notably, these inter-ring signals are even faster than those, intra-ring, between a_1 and a_2, attesting to the close coupling between the two rings. These findings are depicted in Fig. 8D–E where subunits are coloured based on relatively fast (*green*), moderate (*orange*) and slow (*red*) transfer of information.

Finally, we note that the Markovian diffusion of allosteric signals revealed here is specific to the ATP-bound state, and in line with the state-dependent dynamics reported above (see Fig. 4), state-dependent signalling mechanisms operate at different stages of the chaperonin cycle, as illustrated in Fig. S7. For example, TRiC in the NPP state (Zang et al., 2016) (approximating an apo state) is distinguished by a remarkable communication between CCT3, CCT6 and CCT8 (Fig. S7A), which could explain their simultaneous pre-loading. As to the next structure in the cycle (represented by PDB model 6ks7 in Fig. S7B), it is interesting to distinguish that the communication between the low affinity half of the ring (CCT5-CCT8) is much faster than that occurring in the other half (composed of CCT1-CCT4) consistent with the propagation of nucleotide loading to moderate affinity subunits CCT1 and CCT7, and completion of the response of high-affinity subunits (Jin et al., 2019a).

Overall, this analysis again shows the state-dependency of allosteric signalling events at various stages of the chaperonin cycle, and reconciles several experimental observations that, at first sight, might seem contradictory.

4. Discussion and conclusion

The present study has been carried out with two major goals: one methodological and the other biological, enabled by the wealth of structural data collected in recent years on eukaryotic chaperonin TRiC/CCT depicted in Fig. 1. We present both aspects below.

4.1. A computational pipeline for automated evaluation of conformational dynamics from electron density maps

First, we developed and generalized a computational methodology for modelling the structure and dynamics of cryo-EM resolved systems. Structural studies of supramolecular dynamics are limited in their ability to gain insights into the mechanistic basis of allosteric events as a result of (i) limitations in cryo-EM data and their interpretation (Cianfrocco and Kellogg, 2020; Herzik et al., 2019; Sorzano et al., 2019), (ii) approximations in the structure fitting algorithms applied to their outputs (Cianfrocco and Kellogg, 2020; Kim et al., 2020; Malhotra et al., 2019), and (iii) sampling and force field inaccuracies in full-atomic simulations (Bottaro and Lindorff-Larsen, 2018). Hence, we adopted a low-resolution model (ENM/TRN) along with analytical approaches rooted in statistical mechanical and spectral graph theoretical methods, which enabled efficient analysis of global dynamics.

Our computational pipeline comprises three main steps: generation of a pseudoatomic structure from cryo-EM density maps via TRN, identification of subunits or domains via multistep spectral clustering, and analyses of the TRN using ENM-based methods. The pipeline has been implemented in the *ProDy* API, enabling users to take advantage of a wide range of tools developed for ENMs (Fuglebakk et al., 2012; Lopez-Blanco and Chacon, 2016; Tobi and Bahar, 2005; Zheng et al., 2009), including the prediction of the signature dynamics of protein families (Mikulska-Ruminska et al., 2019; Zhang et al., 2019). The *ProDy* API permits us (i) to reconcile physics-based theories of polymer statistical mechanics and machine learning algorithms of spectral graph theory for inferring signalling properties (Chennubhotla and Bahar, 2007); (ii) evaluate sensors and effectors of allosteric signals using perturbation-response theory (Atilgan and Atilgan, 2009; Atilgan et al., 2010; General et al., 2014); (iii) bridge between sequence conservation/coevolution and structural dynamics (Liu and Bahar, 2012); and (iv) visualize molecular motions through an interactive GUI, normal mode wizard *NMWiz*, that interoperates with VMD (Humphrey et al., 1996).

In principle, the pipeline can be applied to any cryo-EM density maps. The availability of the new tools developed here through our GitHub repository (https://github.com/prody/ProDy) enables application to a wide range of biological systems. The methodology is most beneficial for maps of large, dynamic biomolecular complexes at low resolution. Since ENM methods depend mostly on the overall architecture and are highly robust to small differences in atomic positions, TRN/ENM generates mode shapes comparable to those of ENM based on full-atomic models. The availability of all-atom models for CCT helped us test and validate this feature (see Figs. 2 and S5).

However, the accuracy of the current method is limited by that of electron density maps. For example, highly mobile regions such as the apical loops of TRiC/CCT can be missing in the resolved cryo-EM density maps and consequently, the structure and dynamics of those regions cannot be accurately modelled. However, because of their high flexibility/adaptability it is unlikely that these regions would alter the global dynamics. Another limitation is the lack of sequence information. Dynamical domains identified with spectral clustering (shown in Fig. 2) group residues with similar dynamics together but these are not necessarily sequentially contiguous segments. Yet, the overall variability in the 3D shape of the examined system is insensitive to chain connectivity, as bonded and non-bonded neighbours are not distinguished in the neighbour representation. As such, even without knowledge of the sequence, one

can still understand the structure-encoded dynamics, and subunit- or domain-averaged matrices and animations.

4.2. New insights into the conformational space and signalling mechanisms of the supramolecular machine TRiC/CCT

CCT is known to undergo sequential rather than all-or-none changes in the conformations of the subunits in each ring (Kafri and Horovitz, 2003; Reissmann et al., 2012; Rivenzon-Segal et al., 2005). The functional and dynamic asymmetry of this 16-meric machine (Gestaut et al., 2019; Jin et al., 2019a), originates from the sequence and structure heterogeneities of its subunits (see Fig. S1). Our study further revealed that the sequence of these events depends on the 'state' temporarily stabilized during the allosteric cycle of the chaperonin. The sequence of conformational changes undergone by the NPP state, for example, is different from that of the ATP-bound form, which is also different from the hydrolysis transition state, as delineated in Figs. 3-4.

Likewise, the different states exhibit different communication patterns between subunits (Fig. 8 and S7), as predicted by our Markovian stochastics model. Particular subunits exhibited higher propensity to initiate motions and signalling, reconciling previous structural, functional and kinetic studies (Amit et al., 2010; Gestaut et al., 2019; Gruber et al., 2017; Kabir et al., 2011; Kalisman et al., 2013; Leitner et al., 2012). Overall a functional division of the ring into two halves was apparent: one (containing a_1/CCT2, a_2/CCT4 and a_8/CCT5) with high ATP-binding affinity and early hydrolysis capabilities and the other (containing a_4/CCT3, a_5/CCT6 and a_6/CCT8) involved in binding substrate proteins.

In the case of the unbound or NPP state resolved under nominally nucleotide-free conditions (PDB: 5gw4) (Zang et al., 2016), the first subunit predicted to undergo a conformational change in the *cis* ring conformation is a_7 (Fig. 4A); whereas Fig. 4C reveals mobilities in selected *trans* ring subunits in perfect accord with the observations made at slight levels (0.1 mM) of ATP analogue (Jin et al., 2019a). Furthermore, our Markovian stochastics analysis demonstrated a high cooperativity and efficient communication between subunits a_4, a_5, and a_6 (CCT3, CCT6 and CCT8) (Fig. S7A) in this partially preloaded state, and this is also in agreement with the experimentally observed pre-loaded state of these subunits.

TRiC resolved at a slightly higher concentration of ATP analogue (Jin et al., 2019a), on the other hand, exhibits a distinctive peak in the *trans* ring subunit a_7'/CCT7 (Fig. 4B), also consistent with the propensity of this subunit to move after the *cis* ring subunit a_7 in the cryo-EM ensemble resolved at this ADP-AlFx level (Jin et al., 2019a). The structure in which subunit a_7 has moved from its NPP state position (PDB: 6ks7) exhibits a remarkable communication between subunits a_5/CCT6, a_6/CCT8, a_7/CCT7 and a_8/CCT5 (Fig. S7B), i.e. the original group of closely communicating subunits now expands to include subunits a_8 and a_7, prior to transitioning to the fully ATP (analogue)-bound state.

With regard to ATP hydrolysis, a number of studies proposed that the closure events stimulated by ATP-binding and hydrolysis start on the high ATP-binding affinity subunits a_2/CCT4 and a_8/CCT5 (Jin et al., 2019a; Reissmann et al., 2012; Rivenzon-Segal et al., 2005; Yamamoto et al., 2017). Notably, our current analysis also points to these two subunits as the most malleable/adaptable ones *in the pre-hydrolysis ATP analogue-bound structure* (see the *green curves* in Fig. 3B), succeeded by a_1/CCT2 and then a_7/CCT7. These are not necessarily the subunits that bind nucleotide first in the abovementioned studies, but they lead the hydrolysis process, enabled by the topology of contacts assumed in the ATP-bound form. Their high mobility in the global modes accessible to the ATP-bound state enable, if not drive, the movements that accompany the hydrolysis of ATP molecules.

Our Markovian information diffusion analysis (Fig. 8) suggested that the initial movements of CCT4/a_2 and CCT5/a_8 would propagate in both directions through two fast-communicating interfaces (a_3/a_4 and a_7/a_8; Fig. 8B) toward the low ATP affinity subunits. A recent study of hydrolysis kinetics (Gruber et al., 2017) indicates, however, that hydrolysis events start with the subunits that have shown first nucleotide binding behaviour, i.e. CCT3/a_4, CCT6/a_5 and CCT8/a_6, and that there are two conformational waves emanating from these subunits, clockwise and counterclockwise. This type of behaviour is plausible given that a_4/CCT3, a_5/CCT6 and a_6/CCT8 may have bound nucleotide earlier, and if the successive steps of allosteric changes from NPP (PDB: 5gw4) (Zang et al., 2016), to 0.1 mM ADP-AlFx-bound (PDB: 6ks7) (Jin et al., 2019a), and to fully loaded (1 mM) TRiC-ADP-AlFx (PDB: 4a0v) (Cong et al., 2012) (i.e. successive steps including the transitions within the *green box* of Fig. 1A—B and passage to TRiC-ADP-AlFx in *purple box*) are considered. When the individual steps are dissected and the propensity of the state TRiC-AMP-PNP *per se* is examined, however, our analysis reveals a signal feedback from a_2/CCT4 and a_8/CCT5 to a_4/CCT3, a_5/CCT6 and a_6/CCT8, via soft modes of motions (see Movies S1 and S2). Soft modes indeed enable the transition from ATP (analogue)-bound to ADP (analogue)-bound state as illustrated in Fig. 6 and Movie S3.

Next, we examined the intrinsic dynamics of the hydrolysis transition state. Our study also provides insights into motions in the *trans* ring presumably facilitating (Fig. 3B, *lower panel*) or succeeding (Fig. 4C) ATP hydrolysis. The hydrolysis intermediate structure (PDB: 4a0w) exhibits a relatively more ordered or symmetric organization of the subunits compared to that in the apo or partially preloaded states. This suggests that the closure of the protein folding chamber may also involve more symmetrically distributed movements. Our analysis indeed reveals the coupled involvement of all *trans* ring subunits in the softest modes (Movies 4-6) and in particular a highly cooperative mode (mode 3, Movie 6) that enables chamber closure by the collective bending of the apical domains of all subunits a_1'-a_8' in the *trans* ring. This is in contrast to other states where movements of the subunits occurred in a sequential manner. This analysis thus reveals that the 'order' induced in the hydrolysis transition state may drive cooperative movements close to all-or-none in nature at the beginning of this particular step of the allosteric cycle. We note that the *trans* ring subunits a_3'/CCT3 and a_4'/CCT1 are distinguished by their higher mobilities after the deployment of these initial, uniformly distributed movements (high modes in the *orange* and *blue* curves in Fig. 4C).

In order to understand how the intrinsic movements of individual subunits assist in the global machinery, we looked at the softest modes predicted by the GNM for an isolated subunit, a_1/CCT2, of TRiC/CCT. Our analysis revealed that the apical domain's bending ability is encoded in the 3D topology of the individual subunits (see also Movies 7-9) — a phenomenon also observed earlier in GroEL (Yang et al., 2009c) and emphasized in a recent review (Zhang et al., 2020b). The importance of the predicted hinge sites was underscored by their high evolutionary conservation. At the same time, the hinge sites were located next to (and partially coincide with) the highly conserved ATP binding/hydrolysis motif **GDGTT** near the equatorial domain. The close proximity of this motif to the hinge centre attests to the mechanochemical nature of ATP-regulated chaperonin activity. Further examination of the transverse dimers showed how the stabilization of the equatorial domains as anchors enhanced the mobility of the apical domains.

Although the focus of this study was on the intrinsically accessible mechanism of CCT structural change triggered before and upon ATP binding and hydrolysis, the mechanism of substrate loading is equally important for understanding how the CCT

structure/dynamics is coupled to substrate folding, and release. Based on the structures, one possible scheme is that, ATP binding stabilizes a particular conformation of the lid (storing elastic energy for later), which helps unfolded and misfolded substrates bind to the apical domain of the low ATP affinity sides of the upper ring (Cuellar et al., 2019; Leitner et al., 2012; Llorca et al., 1999). Substrate binding then triggers ATP hydrolysis and consequently lid closure in one ring, which releases the elastic energy and facilitates substrate folding by destabilizing the metastable misfolded state. The correctly folded substrate has relatively weaker binding affinity to CCT and is released after lid opening. The sequential firing of ATP molecules might be important for using the ATP energy efficiently, as it might cause less energy dissipation directly as heat and give time for the polypeptide to search and relax within the conformation space — annealing. More experimental and computational data and studies are needed to clarify these aspects of the chaperoning process.

The present computations are exclusively based on the overall network topology representative of electron density maps; no information on sequence properties or chemical interactions is included. The observed mechanisms of events, many in line with experimental observations, reveal that the mechanical properties of the chaperonin have been tailored to comply with its chemical character and allosteric function, presumably optimized by evolution.

Author statement

Yan Zhang: Conceptualization, Methodology, Software, Writing - original draft, Formal analysis, Investigation, Visualization, Writing - review & editing, Validation. **James Krieger**: Formal analysis, Investigation, Visualization, Writing - original draft, Writing - review & editing. **Karolina Mikulska-Ruminska**: Formal analysis, Visualization, Investigation. **Burak Kaynak**: Methodology, Software, Formal analysis, Investigation, Visualization, Writing - review & editing, Validation. **Carlos Oscar S. Sorzano**: Writing - review & editing, Funding acquisition. **José-María Carazo**: Writing - review & editing, Funding acquisition. **Jianhua Xing**: Supervision, Writing - review & editing, **Ivet Bahar**: Conceptualization, Formal analysis, Supervision, Writing - original draft, Writing - review & editing, Project administration, Funding acquisition.

Acknowledgment

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pbiomolbio.2020.08.006.

References

Abrams, C., Bussi, G., 2013. Enhanced sampling in molecular dynamics using metadynamics, replica-exchange, and temperature-acceleration. Entropy 16, 163—199.

Alnabati, E., Kihara, D., 2019. Advances in structure modeling mMethods forcryo-electron microscopy maps. Molecules 25.

Amit, M., Weisberg, S.J., Nadler-Holly, M., McCormack, E.A., Feldmesser, E., Kaganovich, D., Willison, K.R., Horovitz, A., 2010. Equivalent mutations in the eight subunits of the chaperonin CCT produce dramatically different cellular and gene expression phenotypes. J. Mol. Biol. 401, 532—543.

Ashkenazy, H., Abadi, S., Martz, E., Chay, O., Mayrose, I., Pupko, T., Ben-Tal, N., 2016. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res. 44, W344—W350.

Atilgan, A.R., Durell, S.R., Jernigan, R.L., Demirel, M.C., Keskin, O., Bahar, I., 2001. Anisotropy of fluctuation dynamics of proteins with an elastic network model. Biophys. J. 80, 505—515.

Atilgan, C., Atilgan, A.R., 2009. Perturbation-response scanning reveals ligand entry-exit mechanisms of ferric binding protein. PLoS Comput. Biol. 5, e1000544.

Atilgan, C., Gerek, Z.N., Ozkan, S.B., Atilgan, A.R., 2010. Manipulation of conformational change in proteins by single-residue perturbations. Biophys. J. 99, 933—943.

Bahar, I., Atilgan, A.R., Erman, B., 1997. Direct Evaluation of Thermal Fluctuations in Proteins Using a Single-Parameter Harmonic Potential. Fold Des. 2, 173—181.

Bahar, I., Jernigan, R.L., Dill, K.A., 2017. Protein Actions: Principles and Modeling. Garland Science.

Bahar, I., Lezon, T.R., Bakan, A., Shrivastava, I.H., 2010a. Normal mode analysis of biomolecular structures: functional mechanisms of membrane proteins. Chem. Rev. 110, 1463—1497.

Bahar, I., Lezon, T.R., Yang, L.W., Eyal, E., 2010b. Global dynamics of proteins: bridging between structure and function. Annu. Rev. Biophys. 39, 23—42.

Bai, F., Branch, R.W., Nicolau Jr., D.V., Pilizota, T., Steel, B.C., Maini, P.K., Berry, R.M., 2010. Conformational spread as a mechanism for cooperativity in the bacterial flagellar switch. Science 327, 685—689.

Bakan, A., Dutta, A., Mao, W., Liu, Y., Chennubhotla, C., Lezon, T.R., Bahar, I., 2014. Evol and ProDy for bridging protein sequence evolution and structural dynamics. Bioinformatics 30, 2681—2683.

Bakan, A., Meireles, L.M., Bahar, I., 2011. ProDy: protein dynamics inferred from theory and experiments. Bioinformatics 27, 1575—1577.

Bernardi, R.C., Melo, M.C.R., Schulten, K., 2015. Enhanced sampling techniques in molecular dynamics simulations of biological systems. Biochim. Biophys. Acta 1850, 872—877.

Beuron, F., Flynn, T.C., Ma, J., Kondo, H., Zhang, X., Freemont, P.S., 2003. Motions and negative cooperativity between p97 domains revealed by cryo-electron microscopy and quantised elastic deformational model. J. Mol. Biol. 327, 619—629.

Bonomi, M., Pellarin, R., Vendruscolo, M., 2018. Simultaneous determination of protein structure and dynamics using cryo-electron microscopy. Biophys. J. 114, 1604—1613.

Bonomi, M., Vendruscolo, M., 2019. Determination of protein structural ensembles using cryo-electron microscopy. Curr. Opin. Struct. Biol. 56, 37—45.

Bottaro, S., Lindorff-Larsen, K., 2018. Biophysical experiments and biomolecular simulations: a perfect match? Science 361, 355—360.

Callaway, E., 2020. Revolutionary cryo-EM is taking over structural biology. Nature 578, 201.

Chacon, P., Tama, F., Wriggers, W., 2003. Mega-Dalton biomolecular motion captured from electron microscopy reconstructions. J. Mol. Biol. 326, 485—492.

Changeux, J.P., 2012. Allostery and the Monod-Wyman-Changeux model after 50 years. Annu. Rev. Biophys. 41, 103—133.

Cheng, Y., 2018. Single-particle cryo-EM-How did it get here and where will it go. Science 361, 876—880.

Chennubhotla, C., Bahar, I., 2006. Markov propagation of allosteric effects in biomolecular systems: application to GroEL-GroES. Mol. Syst. Biol. 2, 36.

Chennubhotla, C., Bahar, I., 2007. Signal propagation in proteins and relation to equilibrium fluctuations. PLoS Comput. Biol. 3, 1716—1726.

Chennubhotla, C., Rader, A.J., Yang, L.W., Bahar, I., 2005. Elastic network models for understanding biomolecular machinery: from enzymes to supramolecular assemblies. Phys. Biol. 2, S173—S180.

Cianfrocco, M.A., Kellogg, E.H., 2020. What could go wrong? A practical guide to single-particle cryo-EM: from biochemistry to atomic models. J. Chem. Inf. Model. 60, 2458—2469.

Cong, Y., Schroder, G.F., Meyer, A.S., Jakana, J., Ma, B., Dougherty, M.T., Schmid, M.F., Reissmann, S., Levitt, M., Ludtke, S.L., Frydman, J., Chiu, W., 2012. Symmetry-free cryo-EM structures of the chaperonin TRiC along its ATPase-driven conformational cycle. EMBO J. 31, 720—730.

Cossio, P., Hummer, G., 2018. Likelihood-based structural analysis of electron microscopy images. Curr. Opin. Struct. Biol. 49, 162—168.

Costa, M.G.S., Fagnen, C., Venien-Bryan, C., Perahia, D., 2020. A new strategy for atomic flexible fitting in cryo-EM maps by molecular dynamics with excited normal modes (MDeNM-EMfit). J. Chem. Inf. Model. 60, 2419—2423.

Cuellar, J., Ludlam, W.G., Tensmeyer, N.C., Aoba, T., Dhavale, M., Santiago, C., Bueno-Carrasco, M.T., Mann, M.J., Plimpton, R.L., Makaju, A., Franklin, S., Willardson, B.M., Valpuesta, J.M., 2019. Structural and functional analysis of the role of the chaperonin CCT in mTOR complex assembly. Nat. Commun. 10, 2865.

DiMaio, F., Chiu, W., 2016. Tools for model building and optimization into near-atomic resolution electron cryo-microscopy density maps. Methods Enzymol. 579, 255—276.

Doruker, P., Jernigan, R.L., Bahar, I., 2002. Dynamics of large proteins through hierarchical levels of coarse-grained structures. J. Comput. Chem. 23, 119—127.

Doruker, P., Liu, Y., Yang, Z., Bahar, I., 2012. In silico coarse-grained approaches to structural dynamics and function of proteins and their assemblies. In: Egelman, E.-i.-C.E. (Ed.), Comprehensive Biophysics. Academic Press, Elsevier, pp. 27—52.

Dror, R.O., Dirks, R.M., Grossman, J.P., Xu, H., Shaw, D.E., 2012. Biomolecular simulation: a computational microscope for molecular biology. Annu. Rev. Biophys. 41, 429—452.

Duke, T.A., Le Novere, N., Bray, D., 2001. Conformational spread in a ring of proteins: a stochastic approach to allostery. J. Mol. Biol. 308, 541—553.

Elmlund, H., Elmlund, D., Bengio, S., 2013. PRIME: Probabilistic initial 3D model generation for single-particle cryo-electron microscopy. Structure 21, 1299—1306.

Eyal, E., Lum, G., Bahar, I., 2015. The anisotropic network model web server at 2015 (ANM 2.0). Bioinformatics 31, 1487–1489.

Fiedler, M., 1989. Laplacian of Graphs and Algebraic Connectivity. Banach Center Publications, Warsaw.

Fuglebakk, E., Echave, J., Reuter, N., 2012. Measuring and comparing structural fluctuation patterns in large protein datasets. Bioinformatics 28, 2431–2440.

General, I.J., Liu, Y., Blackburn, M.E., Mao, W., Gierasch, L.M., Bahar, I., 2014. ATPase subdomain IA is a mediator of interdomain allostery in Hsp70 molecular chaperones. PLoS Comput. Biol. 10, e1003624.

Gestaut, D., Limatola, A., Joachimiak, L., Frydman, J., 2019. The ATP-powered gymnastics of TRiC/CCT: an asymmetric protein folding machine with a symmetric origin story. Curr. Opin. Struct. Biol. 55, 50–58.

Gomez-Blanco, J., Kaur, S., Ortega, J., Vargas, J., 2019. A robust approach to ab initio cryo-electron microscopy initial volume determination. J. Struct. Biol. 208, 107397.

Gruber, R., Horovitz, A., 2016. Allosteric mechanisms in chaperonin machines. Chem. Rev. 116, 6588–6606.

Gruber, R., Levitt, M., Horovitz, A., 2017. Sequential allosteric mechanism of ATP hydrolysis by the CCT/TRiC chaperone is revealed through Arrhenius analysis. Proc. Natl. Acad. Sci. U. S. A. 114, 5189–5194.

Haliloglu, T., Bahar, I., 2015. Adaptability of protein structures to enable functional interactions and evolutionary implications. Curr. Opin. Struct. Biol. 35, 17–23.

Harpole, T.J., Delemotte, L., 2018. Conformational landscapes of membrane proteins delineated by enhanced sampling molecular dynamics simulations. Biochim. Biophys. Acta Biomembr. 1860, 909–926.

Herzik Jr., M.A., Fraser, J.S., Lander, G.C., 2019. A multi-model approach to assessing local and global cryo-EM map quality. Structure 27, 344–358 e3.

Hinsen, K., 1998. Analysis of domain motions by approximate normal mode calculations. Proteins 33, 417–429.

Hinsen, K., 2008. Structural flexibility in proteins: impact of the crystal environment. Bioinformatics 24, 521–528.

Hinsen, K., Kneller, G.R., 1999. A simplified force field for describing vibrational protein dynamics over the whole frequency range. J. Chem. Phys. 111, 10766–10769.

Hinsen, K., Petrescu, A., Dellerue, S., Bellissent-Funel, M., Kneller, G.R., 2000. Harmonicity in slow protein dynamics. Chem. Phys. 261, 25–37.

Hinsen, K., Reuter, N., Navaza, J., Stokes, D.L., Lacapere, J.J., 2005. Normal mode-based fitting of atomic structure into electron density maps: application to sarcoplasmic reticulum Ca-ATPase. Biophys. J. 88, 818–827.

Hollingsworth, S.A., Dror, R.O., 2018. Molecular dynamics simulation for all. Neuron 99, 1129–1143.

Humphrey, W., Dalke, A., Schulten, K., 1996. VMD: visual molecular dynamics. J. Mol. Graph. 14 (33–8), 27–28.

Jianbo, S., Malik, J., 2000. Normalized cuts and image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 22, 888–905.

Jin, M., Han, W., Liu, C., Zang, Y., Li, J., Wang, F., Wang, Y., Cong, Y., 2019a. An ensemble of cryo-EM structures of TRiC reveal its conformational landscape and subunit specificity. Proc. Natl. Acad. Sci. U. S. A. 116, 19513–19522.

Jin, M., Liu, C., Han, W., Cong, Y., 2019b. TRiC/CCT chaperonin: structure and function. Subcell. Biochem. 93, 625–654.

Jonic, S., Sorzano, C.O., 2016. Versatility of approximating single-particle electron microscopy density maps using pseudoatoms and approximation-accuracy control. BioMed Res. Int. 2016, 7060348.

Joseph, A.P., Polles, G., Alber, F., Topf, M., 2017. Integrative modelling of cellular assemblies. Curr. Opin. Struct. Biol. 46, 102–109.

Kabir, M.A., Uddin, W., Narayanan, A., Reddy, P.K., Jairajpuri, M.A., Sherman, F., Ahmad, Z., 2011. Functional subunits of eukaryotic chaperonin CCT/TRiC in protein folding. J. Amino Acids 2011, 843206.

Kafri, G., Horovitz, A., 2003. Transient kinetic analysis of ATP-induced allosteric transitions in the eukaryotic chaperonin containing TCP-1. J. Mol. Biol. 326, 981–987.

Kalisman, N., Schroder, G.F., Levitt, M., 2013. The crystal structures of the eukaryotic chaperonin CCT reveal its functional partitioning. Structure 21, 540–549.

Kim, D.N., Gront, D., Sanbonmatsu, K.Y., 2020. Practical considerations for atomistic structure modeling with cryo-EM maps. J. Chem. Inf. Model. 60, 2436–2442.

Kohonen, T., 2001. Self-Organizing Maps, third ed. Springer-Verlag, Berlin.

Kong, Y., Ming, D., Wu, Y., Stoops, J.K., Zhou, Z.H., Ma, J., 2003. Conformational flexibility of pyruvate dehydrogenase complexes: a computational analysis by quantized elastic deformational model. J. Mol. Biol. 330, 129–135.

Koshland Jr., D.E., Nemethy, G., Filmer, D., 1966. Comparison of experimental binding data and theoretical models in proteins containing subunits. Biochemistry 5, 365–385.

Koukos, P.I., Bonvin, A., 2020. Integrative modelling of biomolecular complexes. J. Mol. Biol. 432, 2861–2881.

Krieger, J.M., Doruker, P., Scott, A.L., Perahia, D., Bahar, I., 2020. Towards gaining sight of multiscale events: utilizing network models and normal modes in hybrid methods. Curr. Opin. Struct. Biol. 64, 34–41.

Kurkcuoglu, Z., Bahar, I., Doruker, P., 2016. ClustENM: ENM-based sampling of essential conformational space at full atomic resolution. J. Chem. Theor. Comput. 12, 4549–4562.

Lawson, C.L., Berman, H.M., Chiu, W., 2020. Evolving data standards for cryo-EM structures. Struct. Dyn. 7, 014701.

Lawson, C.L., Patwardhan, A., Baker, M.L., Hryc, C., Garcia, E.S., Hudson, B.P., Lagerstedt, I., Ludtke, S.J., Pintilie, G., Sala, R., Westbrook, J.D., Berman, H.M., Kleywegt, G.J., Chiu, W., 2016. EMDataBank unified data resource for 3DEM.

Nucleic Acids Res. 44, D396–D403.

Lee, J.A., Verleysen, M., 2007. Nonlinear Dimensionality Reduction. Springer.

Leelananda, S.P., Lindert, S., 2020. Using NMR chemical shifts and cryo-EM density restraints in iterative rosetta-MD protein structure refinement. J. Chem. Inf. Model. 60, 2522–2532.

Leitner, A., Joachimiak, L.A., Bracher, A., Monkemeyer, L., Walzthoeni, T., Chen, B., Pechmann, S., Holmes, S., Cong, Y., Ma, B., Ludtke, S., Chiu, W., Hartl, F.U., Aebersold, R., Frydman, J., 2012. The molecular architecture of the eukaryotic chaperonin TRiC/CCT. Structure 20, 814–825.

Lezon, T.R., Bahar, I., 2012. Constraints imposed by the membrane selectively guide the allosteric access dynamics of the glutamate transporter GltPh. Biophys. J. 102, 1331–1340.

Li, H., Chang, Y.Y., Lee, J.Y., Bahar, I., Yang, L.W., 2017. DynOmics: dynamics of structural proteome and beyond. Nucleic Acids Res. 45, W374–W380.

Liu, Y., Bahar, I., 2012. Sequence evolution correlates with structural dynamics. Mol. Biol. Evol. 29, 2253–2263.

Llorca, O., McCormack, E.A., Hynes, G., Grantham, J., Cordell, J., Carrascosa, J.L., Willison, K.R., Fernandez, J.J., Valpuesta, J.M., 1999. Eukaryotic type II chaperonin CCT interacts with actin through specific subunits. Nature 402, 693–696.

Lopez-Blanco, J.R., Chacon, P., 2013. iMODFIT: efficient and robust flexible fitting based on vibrational analysis in internal coordinates. J. Struct. Biol. 184, 261–270.

Lopez-Blanco, J.R., Chacon, P., 2016. New generation of elastic network models. Curr. Opin. Struct. Biol. 37, 46–53.

Lu, M., Poon, B., Ma, J., 2006. A new method for coarse-grained elastic normal-mode analysis. J. Chem. Theor. Comput. 2, 464–471.

Lyumkis, D., Vinterbo, S., Potter, C.S., Carragher, B., 2013. Optimod–an automated approach for constructing and optimizing initial models for single-particle electron microscopy. J. Struct. Biol. 184, 417–426.

Ma, Y., Fu, Y., 2012. Manifold Learning Theory and Applications. CRC Press, New York.

Maji, S., Liao, H., Dashti, A., Mashayekhi, G., Ourmazd, A., Frank, J., 2020. Propagation of conformational coordinates across angular space in mapping the continuum of states from cryo-EM data by manifold embedding. J. Chem. Inf. Model. 60, 2484–2491.

Malhotra, S., Trager, S., Dal Peraro, M., Topf, M., 2019. Modelling structures in cryo-EM maps. Curr. Opin. Struct. Biol. 58, 105–114.

Martinetz, T., Schulten, K., 1994. Topology representing networks. Neural Network. 7, 507–522.

Matsumoto, A., Ishida, H., 2009. Global conformational changes of ribosome observed by normal mode fitting for 3D Cryo-EM structures. Structure 17, 1605–1613.

Meireles, L., Gur, M., Bakan, A., Bahar, I., 2011. Pre-existing soft modes of motion uniquely defined by native contact topology facilitate ligand binding to proteins. Protein Sci. 20, 1645–1658.

Meyer, A.S., Gillespie, J.R., Walther, D., Millet, I.S., Doniach, S., Frydman, J., 2003. Closing the folding chamber of the eukaryotic chaperonin requires the transition state of ATP hydrolysis. Cell 113, 369–381.

Mikulska-Ruminska, K., Shrivastava, I., Krieger, J., Zhang, S., Li, H., Bayir, H., Wenzel, S.E., VanDemark, A.P., Kagan, V.E., Bahar, I., 2019. Characterization of differential dynamics, specificity, and allostery of lipoxygenase family members. J. Chem. Inf. Model. 59, 2496–2508.

Ming, D., Kong, Y., Lambert, M.A., Huang, Z., Ma, J., 2002a. How to describe protein motion without amino acid sequence and atomic coordinates. Proc. Natl. Acad. Sci. U. S. A. 99, 8620–8625.

Ming, D., Kong, Y., Wakil, S.J., Brink, J., Ma, J., 2002b. Domain movements in human fatty acid synthase by quantized elastic deformational model. Proc. Natl. Acad. Sci. U. S. A. 99, 7895–7899.

Ming, D., Wall, M.E., 2005. Allostery in a coarse-grained model of protein dynamics. Phys. Rev. Lett. 95, 198103.

Mitra, A.K., 2019. Visualization of biological macromolecules at near-atomic resolution: cryo-electron microscopy comes of age. Acta Crystallogr. F Struct. Biol. Commun. 75, 3–11.

Miyashita, O., Kobayashi, C., Mori, T., Sugita, Y., Tama, F., 2017. Flexible fitting to cryo-EM density map using ensemble molecular dynamics simulations. J. Comput. Chem. 38, 1447–1461.

Miyashita, O., Tama, F., 2018. Hybrid methods for macromolecular modeling by molecular mechanics simulations with experimental data. Adv. Exp. Med. Biol. 1105, 199–217.

Munoz, I.G., Yebenes, H., Zhou, M., Mesa, P., Serna, M., Park, A.Y., Bragado-Nilsson, E., Beloso, A., de Carcer, G., Malumbres, M., Robinson, C.V., Valpuesta, J.M., Montoya, G., 2011. Crystal structure of the open conformation of the mammalian chaperonin CCT in complex with tubulin. Nat. Struct. Mol. Biol. 18, 14–19.

Na, H., Song, G., 2016. The effective degeneracy of protein normal modes. Phys. Biol. 13, 036002.

Ng, A.Y., Jordan, M.I., Weiss, Y., 2001. On spectral clustering: analysis and an algorithm. In: NIPS'01: Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic, pp. 849–856.

Nicholls, R.A., Tykac, M., Kovalevskiy, O., Murshudov, G.N., 2018. Current approaches for the fitting and refinement of atomic models into cryo-EM maps using CCP-EM. Acta Crystallogr. D Struct. Biol. 74, 492–505.

Nogales, E., 2016. The development of cryo-EM into a mainstream structural biology technique. Nat. Methods 13, 24–27.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A.,

Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: ma-chine learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

Punjani, A., Rubinstein, J.L., Fleet, D.J., Brubaker, M.A., 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods 14, 290–296.

Rader, A.J., Vlad, D.H., Bahar, I., 2005. Maturation dynamics of bacteriophage HK97 capsid. Structure 13, 413–421.

Reissmann, S., Joachimiak, L.A., Chen, B., Meyer, A.S., Nguyen, A., Frydman, J., 2012. A gradient of ATP affinities generates an asymmetric power stroke driving the chaperonin TRIC/CCT folding cycle. Cell Rep. 2, 866–877.

Riccardi, D., Cui, Q., Phillips Jr., G.N., 2009. Application of elastic network models to proteins in the crystalline state. Biophys. J. 96, 464–475.

Rivenzon-Segal, D., Wolf, S.G., Shimon, L., Willison, K.R., Horovitz, A., 2005. Sequential ATP-induced allosteric transitions of the cytoplasmic chaperonin containing TCP-1 revealed by EM analysis. Nat. Struct. Mol. Biol. 12, 233–237.

Sauerwald, N., Zhang, S., Kingsford, C., Bahar, I., 2017. Chromosomal dynamics predicted by an elastic network model explains genome-wide accessibility and long-range couplings. Nucleic Acids Res. 45, 3663–3673.

Scheres, S.H., Gao, H., Valle, M., Herman, G.T., Eggermont, P.P., Frank, J., Carazo, J.M., 2007. Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. Nat. Methods 4, 27–29.

Shaw, D.E., Dror, R.O., Salmon, J.K., Grossman, J.P., Mackenzie, K.M., Bank, J.A., Young, C., Deneroff, M.M., Batson, B., Bowers, K.J., Chow, E., Eastwood, M.P., Ierardi, D.J., Klepeis, J.L., Kuskin, J.S., Larson, R.H., Lindorff-Larsen, K., Maragakis, P., Moraes, M.A., Piana, S., Shan, Y., Towles, B.P., 2009. Millisecond-scale molecular dynamics simulations on Anton. In: SC '09: Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis, 65, pp. 1–11.

Shaw, D.E., Grossman, J.P., Bank, J.A., Batson, B., Butts, J.A., Chao, J.C., Deneroff, M.M., Dror, R.O., Even, A., Fenton, C.H., Forte, A., Gagliardo, J., Gill, G., Greskamp, B., Ho, C.R., Ierardi, D.J., Iserovich, L., Kuskin, J.S., Larson, R.H., Layman, T., Lee, L.-S., Lerer, A.K., Li, C., Killebrew, D., Mackenzie, K.M., Mok, S.Y.-H., Moraes, M.A., Mueller, R., Nociolo, L.J., Peticolas, J.L., 2014. Anton 2: raising the bar for per-formance and programmability in a special-purpose molecular dynamics su-percomputer. In: SC '14: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, pp. 41–53.

Shevchuk, Y., 2015. NeuPy: neural networks in Python. Retrieved from. http://neupy.com. (Accessed 14 November 2019).

Skjaerven, L., Cuellar, J., Martinez, A., Valpuesta, J.M., 2015. Dynamics, flexibility, and allostery in molecular chaperonins. FEBS Lett. 589, 2522–2532.

Sorzano, C.O.S., Jimenez, A., Mota, J., Vilas, J.L., Maluenda, D., Martinez, M., Ramirez-Aportela, E., Majtner, T., Segura, J., Sanchez-Garcia, R., Rancel, Y., Del Cano, L., Conesa, P., Melero, R., Jonic, S., Vargas, J., Cazals, F., Freyberg, Z., Krieger, J., Bahar, I., Marabini, A., Carazo, J.M., 2019. Survey of the analysis of continuous conformational variability of biological macromolecules by electron micro-scopy. Acta Crystallogr. F Struct. Biol. Commun. 75, 19–32.

Suhre, K., Navaza, J., Sanejouand, Y.H., 2006. NORMA: a tool for flexible fitting of high-resolution protein structures into low-resolution electron-microscopy-derived density maps. Acta Crystallogr. Sect. D Biol. Crystallogr. 62, 1098–1100.

Tama, F., Gadea, F.X., Marques, O., Sanejouand, Y.H., 2000. Building-block approach for determining low-frequency normal modes of macromolecules. Proteins 41, 1–7.

Tama, F., Miyashita, O., Brooks 3rd, C.L., 2004a. Flexible multi-scale fitting of atomic structures into low-resolution electron density maps with elastic network normal mode analysis. J. Mol. Biol. 337, 985–999.

Tama, F., Miyashita, O., Brooks 3rd, C.L., 2004b. Normal mode based flexible fitting of high-resolution structure into low-resolution experimental data from cryo-EM. J. Struct. Biol. 147, 315–326.

Tobi, D., Bahar, I., 2005. Structural changes involved in protein binding correlate with intrinsic motions of proteins in the unbound state. Proc. Natl. Acad. Sci. U. S. A. 102, 18908–18913.

Trabuco, L.G., Villa, E., Mitra, K., Frank, J., Schulten, K., 2008. Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. Structure 16, 673–683.

Velazquez-Muriel, J.A., Carazo, J.M., 2007. Flexible fitting in 3D-EM with incomplete data on superfamily variability. J. Struct. Biol. 158, 165–181.

Vilas, J.L., Tabassum, N., Mota, J., Maluenda, D., Jimenez-Moreno, A., Majtner, T., Carazo, J.M., Acton, S.T., Sorzano, C.O.S., 2018. Advances in image processing for single-particle analysis by electron cryomicroscopy and challenges ahead. Curr. Opin. Struct. Biol. 52, 127–145.

Wang, H., Han, W., Takagi, J., Cong, Y., 2018. Yeast inner-subunit PA-NZ-1 labeling strategy for accurate subunit identification in a macromolecular complex through cryo-EM analysis. J. Mol. Biol. 430, 1417–1425.

Xie, R., Chen, Y.X., Cai, J.M., Yang, Y., Shen, H.B., 2020. SPREAD: a fully automated toolkit for single-particle cryogenic electron microscopy data 3D reconstruction with image-network-aided orientation assignment. J. Chem. Inf. Model. 60, 2614–2625.

Yamamoto, Y.Y., Uno, Y., Sha, E., Ikegami, K., Ishii, N., Dohmae, N., Sekiguchi, H., Sasaki, Y.C., Yohda, M., 2017. Asymmetry in the function and dynamics of the cytosolic group II chaperonin CCT/TRiC. PloS One 12, e0176054.

Yang, L., Song, G., Jernigan, R.L., 2009a. Protein elastic network models and the ranges of cooperativity. Proc. Natl. Acad. Sci. U. S. A. 106, 12347–12352.

Yang, L.W., Eyal, E., Bahar, I., Kitao, A., 2009b. Principal component analysis of native ensembles of biomolecular structures (PCA_NEST): insights into functional dynamics. Bioinformatics 25, 606–614.

Yang, Z., Majek, P., Bahar, I., 2009c. Allosteric transitions of supramolecular systems explored by network models: application to chaperonin GroEL. PLoS Comput. Biol. 5, e1000360.

Zang, Y., Jin, M., Wang, H., Cui, Z., Kong, L., Liu, C., Cong, Y., 2016. Staggered ATP binding mechanism of eukaryotic chaperonin TRiC (CCT) revealed through high-resolution cryo-EM. Nat. Struct. Mol. Biol. 23, 1083–1091.

Zang, Y., Wang, H., Cui, Z., Jin, M., Liu, C., Han, W., Wang, Y., Cong, Y., 2018. Devel-opment of a yeast internal-subunit eGFP labeling strategy and its application in subunit identification in eukaryotic group II chaperonin TRiC/CCT. Sci. Rep. 8, 2374.

Zhang, S., Chen, F., Bahar, I., 2020a. Differences in the intrinsic spatial dynamics of the chromatin contribute to cell differentiation. Nucleic Acids Res. 48, 1131–1145.

Zhang, S., Li, H., Krieger, J.M., Bahar, I., 2019. Shared signature dynamics tempered by local fluctuations enables fold adaptability and specificity. Mol. Biol. Evol. 36, 2053–2068.

Zhang, Y., Doruker, P., Kaynak, B., Zhang, S., Krieger, J., Li, H., Bahar, I., 2020b. Intrinsic dynamics is evolutionarily optimized to enable allosteric behavior. Curr. Opin. Struct. Biol. 62, 14–21.

Zheng, W., Brooks, B.R., 2005. Probing the local dynamics of nucleotide-binding pocket coupled to the global dynamics: myosin versus kinesin. Biophys. J. 89, 167–178.

Zheng, W., Brooks, B.R., Thirumalai, D., 2006. Low-frequency normal modes that describe allosteric transitions in biological nanomachines are robust to sequence variations. Proc. Natl. Acad. Sci. U. S. A. 103, 7664–7669.

Zheng, W., Brooks, B.R., Thirumalai, D., 2009. Allosteric transitions in biological nanomachines are described by robust normal modes of elastic networks. Curr. Protein Pept. Sci. 10, 128–132.

Zheng, W., Tekpinar, M., 2014. High-resolution modeling of protein structures based on flexible fitting of low-resolution structural data. Adv. Protein Chem. Struct. Biol. 96, 267–284.

Supplementary file for

State-Dependent Sequential Allostery Exhibited by Chaperonin TRiC/CCT Revealed by Network Analysis of Cryo-EM Maps

Yan Zhang[1a], James Krieger[1a], Karolina Mikulska-Ruminska[1b], Burak Kaynak[1], Carlos Oscar S. Sorzano[2], José-María Carazo[2], Jianhua Xing[1], and Ivet Bahar[1c]

1. *Department of Computational and Systems Biology, University of Pittsburgh, 800 Murdoch Building, 3420 Forbes Avenue, Pittsburgh, PA 15261, USA*
2. *Centro Nacional de Biotecnología (CSIC), Darwin, 3, 28049 Madrid, Spain*

[a] Equal contribution

[b] Present address: Institute of Physics, Department of Biophysics and Medical Physics, Nicolaus Copernicus University, 87-100 Torun, Poland

[c] Corresponding author

Lead Contact (corresponding author): I Bahar
Department of Computational & Systems Biology
School of Medicine, University of Pittsburgh
800 Murdoch Building, 3420 Forbes Avenue, Pittsburgh, PA 15213

bahar@pitt.edu http://www.csb.pitt.edu/cms/
Lab: http://www.csb.pitt.edu/Faculty/bahar/

# Supplementary Figures
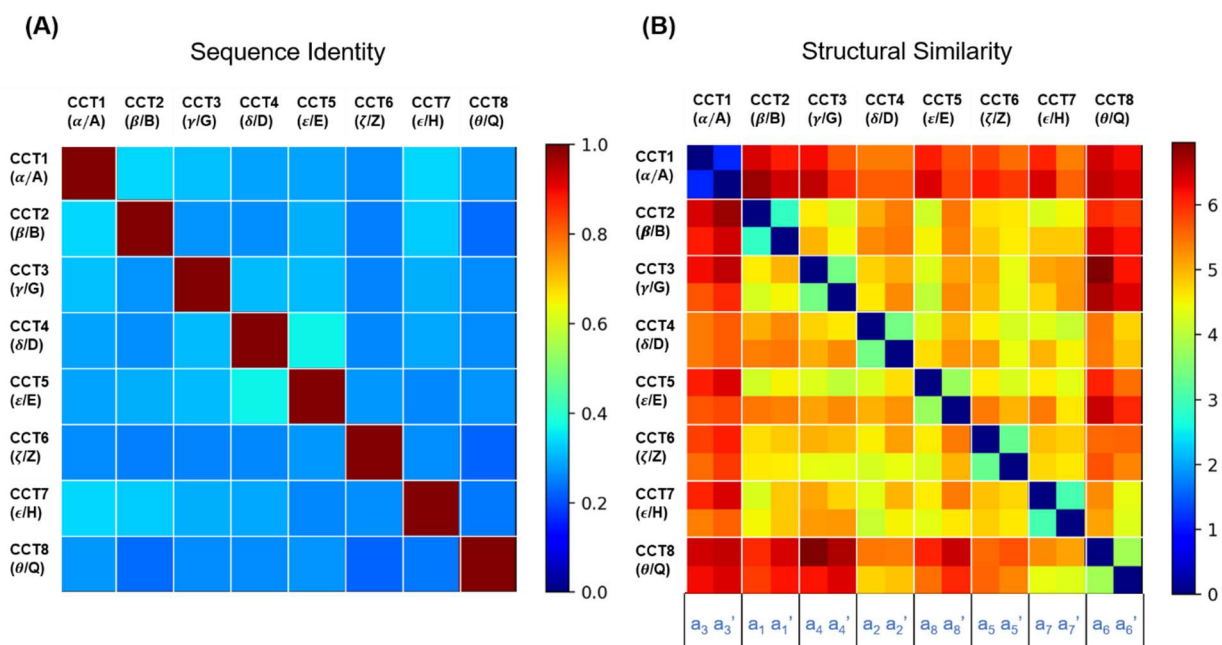
Figure S1. The subunits of CCT are sequence diverse but structurally similar. (A) The sequence identity matrix shows that the sequence identity between subunits of a ring is ~ 20-35 %. Results are shown for bovine CCT. Both rings are hetero-octameric, with the same subunits arranged in opposite order (see **Figure 2**). The eight subunits in a given ring are labelled as CCT1-CCT8, or (α, β, γ, δ, ϵ, ζ, ϵ, θ) or (A, B, G, D, E, Z, H, Q), or a_1-a_8 (and a_1'-a_8'). See more details in **Table S1**. In the present study, we distinguish the *cis* and *trans* ring subunits by adopting the nomenclature a_1-a_8 and a_1'-a_8' for subunits in the respective rings. (B) The pairwise root-mean-square deviation (RMSD) matrix for subunits in yeast CCT resolved in the presence of AMP-PNP (PDB: 5gw5). Most pairs exhibit RMSDs lower than 7 Å. The subunits are arranged in pairs corresponding to matching ones from the upper and lower rings, indicated along the lower abscissa. The *colour bar* shows the RMSD in Ångstroms.
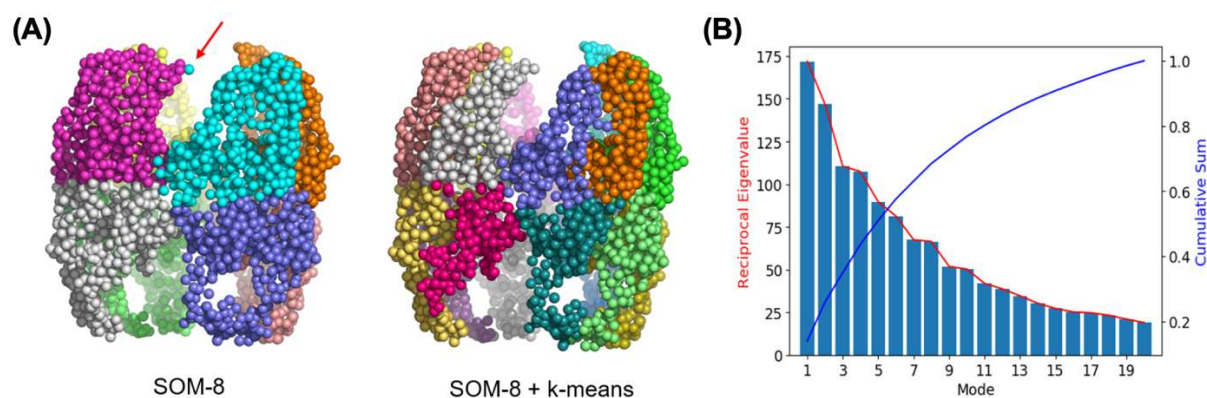
Figure S2. **Supplemental information on ANM soft mode spectrum and SOM-predicted clusters**. **(A)** The organization of the CCT structure in 8 clusters identified by the SOM methodology, where the *red arrow* indicates a typical outlier that needs to be corrected manually (*left*). Identification of the 16 subunits by dividing each SOM-8 cluster further using k-means (*right*). This TRN model refers to the ATP-bound form (EMD-1961). **(B)** The frequency distribution of soft modes predicted by the ANM, which represents the weights of these modes in the collective dynamics of the chaperonin (*bars*; *left ordinate*) and the normalized cumulative sum of the weights (*blue curve, right ordinate*).
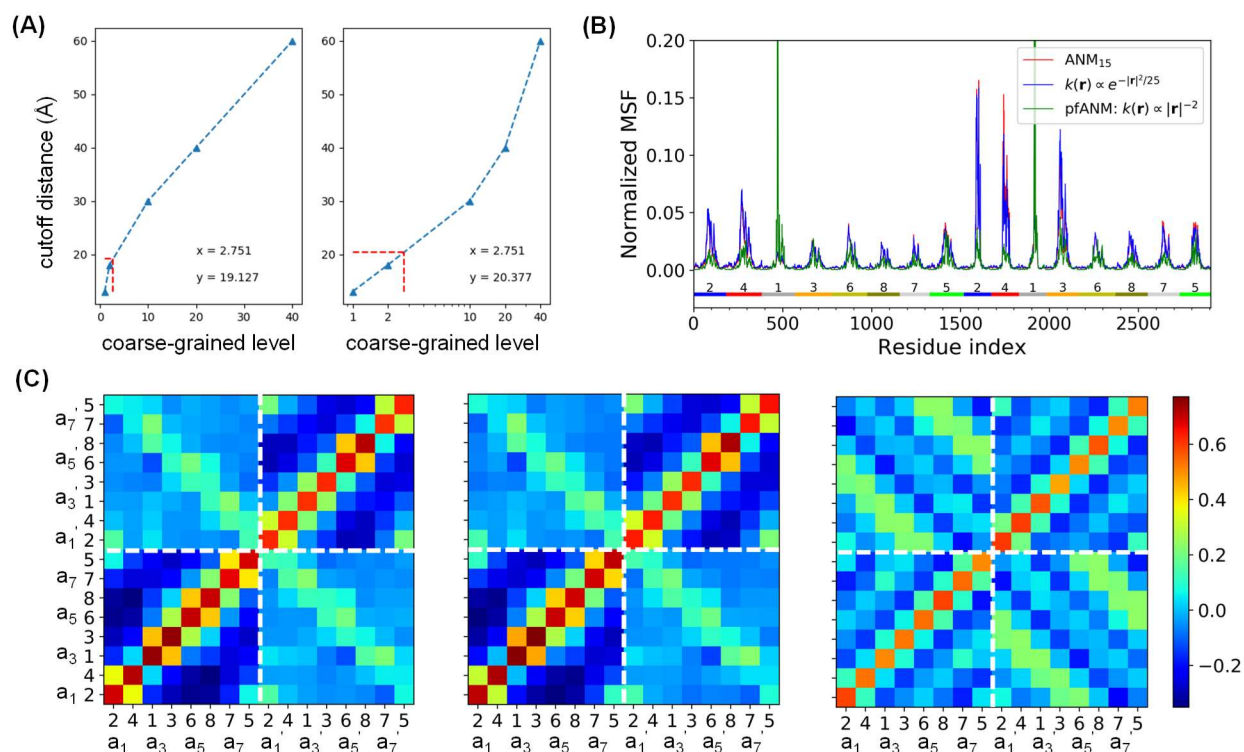
Figure S3. Choice of ANM cutoff distance for placing springs between nodes, and sensitivity of results to the distance dependency of spring constants. **(A)** Interpolation of ANM cutoff distance as a function of the coarse-graining level based on results from (Doruker et al., 2002). Coarse-graining level is defined as the ratio of the number residues in the protein to the number of beads used in the model, higher number indicating an increased sparsity of network nodes to represent the structure. Interpolation on both linear and logarithmic scale yielded a cutoff of ~20 Å corresponding to a coarse-graining level of ~8,000/3,000 ≈ 2.7, which is the ratio of CCT residues to the number of beads in the TRN model. **(B)** Normalized MSFs over the first 20 modes of CCT-ATP (EMD1961) with: a cutoff of 15 Å (*red*); a distance-dependent spring constant in exponential form (*blue*); and parameter-free ANM (pfANM, *green*). **(C)** Subunit-averaged orientational cross-correlations over the first 20 modes of CCT-ATP (EMD1961) using the conventional ANM (Atilgan et al., 2001) with a cutoff of 15 Å (*left*), ANM with exponential distance-dependent spring constant (*middle*), and pfANM (*right*). The results on the *left* and *middle* panels demonstrate similar behaviors as we observe in **Figure 3D**. In the *right* panel, we observe a different trend across subunits on the upper-left or lower-right quadrants, which represent inter-ring cross correlations. Both structural (Cong et al., 2012; Wang et al., 2018; Zang et al., 2016; Zang et al., 2018) and biological (Kalisman et al., 2013; Leitner et al., 2012; Wang et al., 2018; Zang et al., 2018) subunit numbering scheme are labeled on the two axes. White dashed lines separate *cis* and *trans* rings.
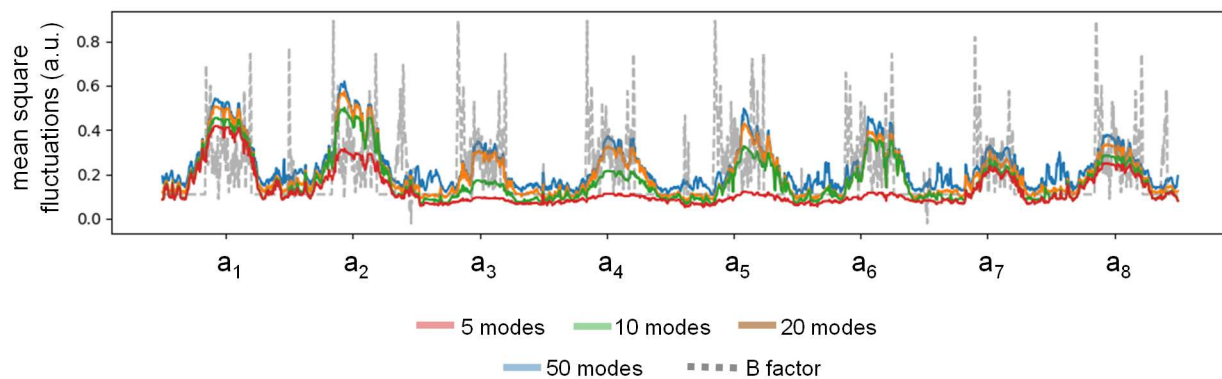
Figure S4. Comparison between mean square fluctuations of the TRN model predicted by GNM and experimental B factor. The *solid curves* are MSFs obtained with 5, 10, 20, and 50 GNM modes, and the *dashed gray line* shows the X-ray crystallographic B factors from the structure deposited by Munoz et al (PDB: 2xsm) (Munoz et al., 2011). The lower ring was incomplete, and therefore is not shown here.
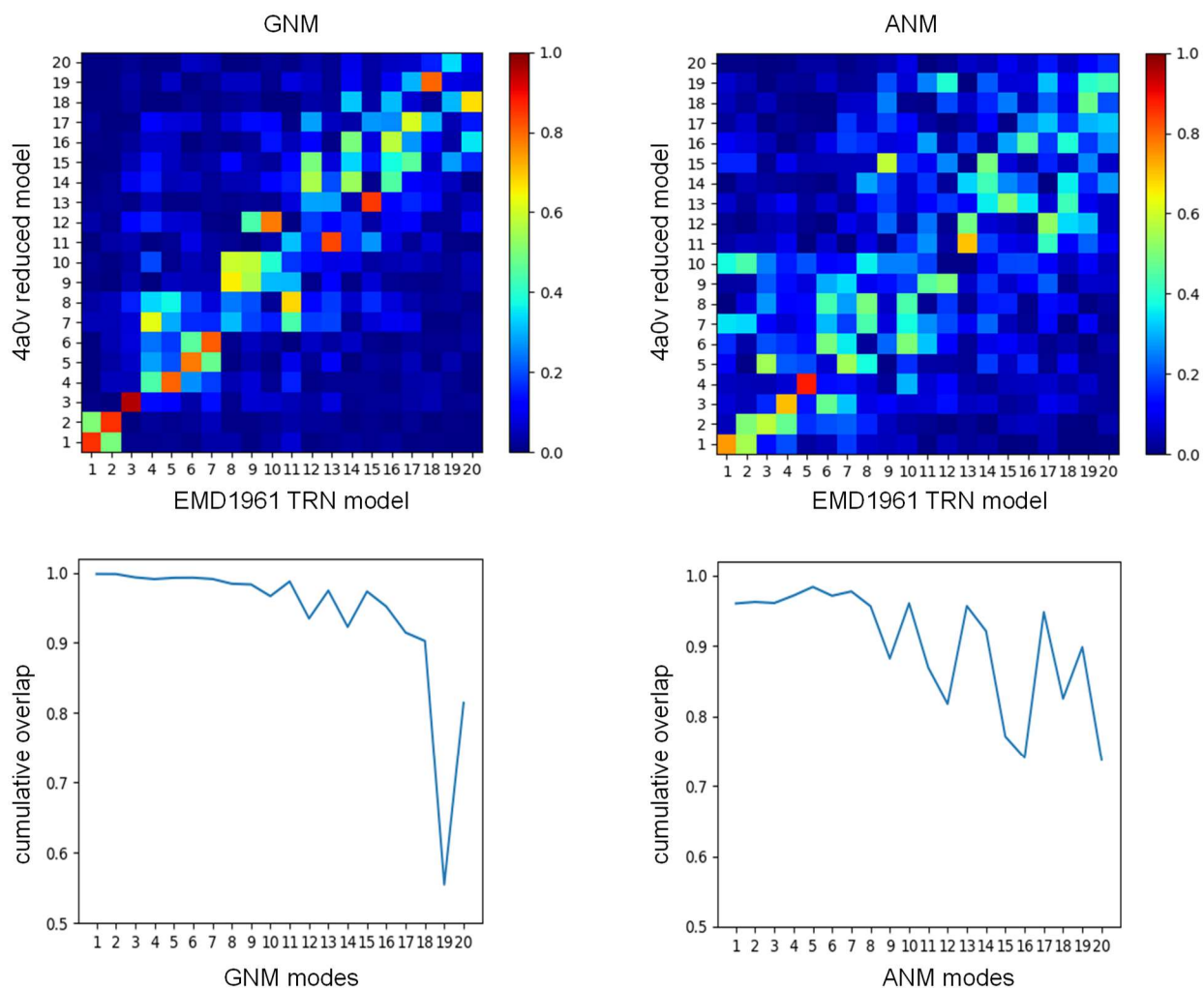
Figure S5. Comparison of global mode spectra predicted full atomic and electron-density-based models for TRiC-AMP-PNP. The softest 20 modes were evaluated by the GNM (*left*) and ANM (*right*) analysis using TRN model derived from electron density maps (EMD1961) and the full-atomic model (PDB: 4a0v) constructed for the same structure. The mode-mode correlation maps *on the top* display the correlation cosines between the two sets of modes (see color scale on the right), the ordinate and abscissa referring to the respective full-atomic elastic network and TRN models. The plots underneath represent the cumulative overlap between each of the global modes k ($1 \leq k \leq 20$) predicted based on the TRN and the softest 20 modes predicted based on the atomic model, showing that each of the softest 20 modes are well accounted for by either model, even though there may be minor changes in the order (relative frequencies) or shape (eigenvectors) of the modes. The all-atom model is aligned to the TRN model with a nearest-neighbour method, and non-aligned residues in the all-atom structure are treated as the environment in the reduced model. Cosine similarity is computed as the measure of mode similarity. Cumulative overlaps are obtained by summing up the squared cosines in each column and taking the square root of the sum.
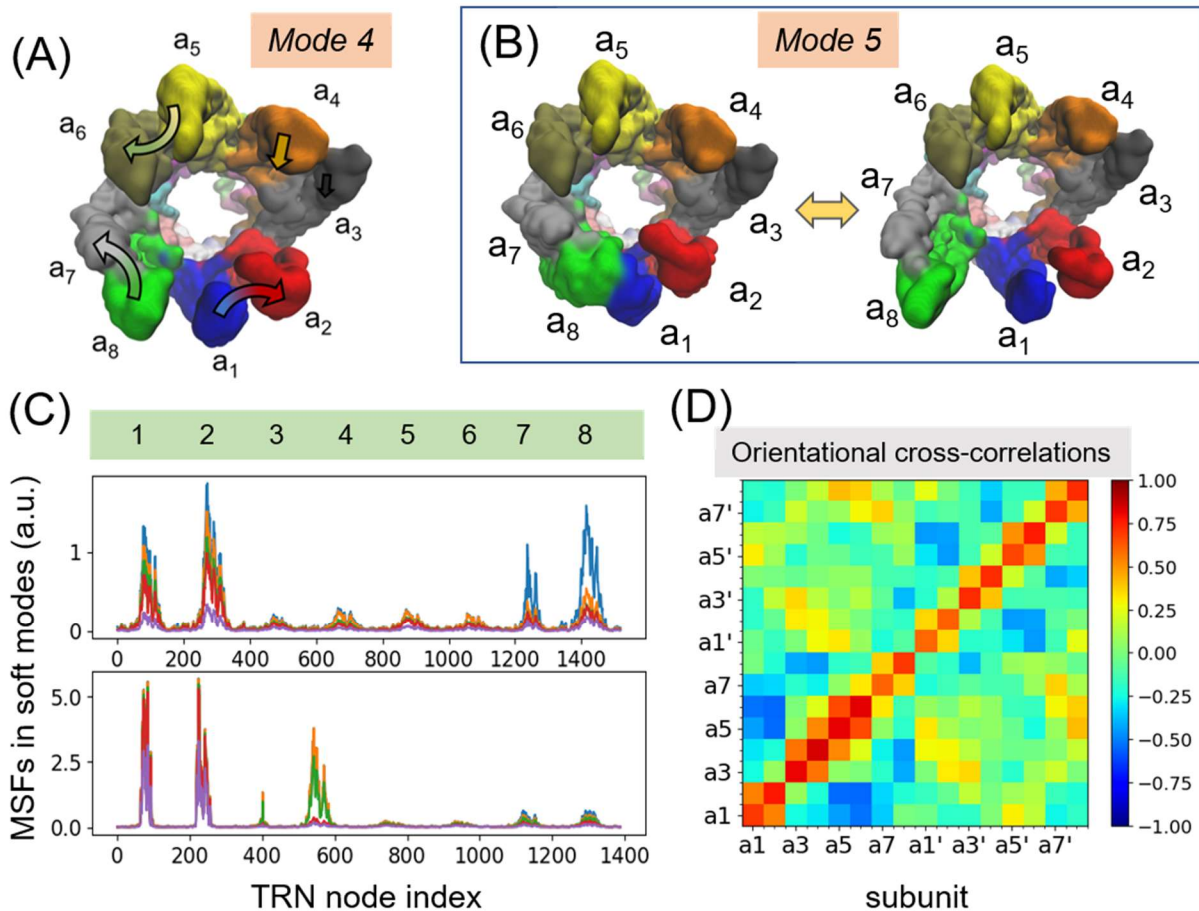
Figure S6. Collective dynamics of CCT-ATP and transition from open to closed state of the upper ring. (A) A highly cooperative mode, *mode 4*, accessible to the cryo-EM structure TRiC-ADP-AIFX (EMDB-1961). The motion is almost uniformly distributed across all subunits in both rings**. (B)** Fluctuations between open and closed conformers in *mode 5*. Note the decoupling of a_8 from of a_1 and its significant fluctuations between open and closed forms. (**C**) MSFs driven *by mode 1* (*violet*), *modes 1-2* (*red*), *modes 1-3* (*green*), *modes 1-4* (orange) and *modes 1-5* (*blue*). Both *modes 4* and *5*, as well as *mode* 6, contribute to the closure of the upper ring (see **Supplementary Movie S1 A-C**). (**D**) Orientational cross-correlations between all pairs of subunits average driven by *modes 1-5* (averaged over all TRN nodes in each subunit), color-coded by the strength and type (negative/anticorrelated, *blue*; or positive/correlated, *red*) of intersubunit cross-correlations.
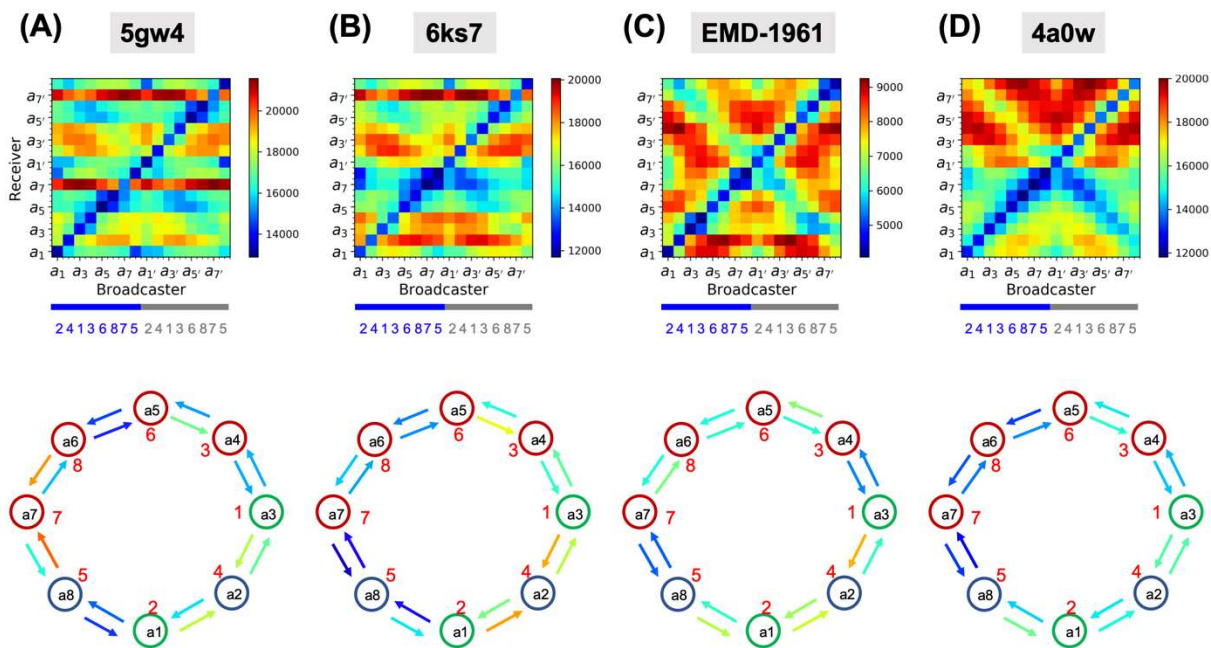
Figure S7. **Subunit-averaged hitting time matrices and corresponding communication patterns for four different states.** Each panel corresponds to different states visited along the allosteric cycle (see **Figure 1**; and also **Figs 3** and **4** for the results for the same structures): **(A)** nucleotide partially preloaded yeast CCT at minimal nucleotide conditions (PDB: 5gw4) (Zang et al., 2016), **(B)** slightly more nucleotide loaded yeast CCT conformer 1 from 0.1 mM ADP-AlFx conditions (PDB: 6ks7) (Jin et al., 2019b), **(C)** fully AMP-PNP loaded bovine CCT based on TRN (corresponding PDB: 4a0v; same as **Figure 8A-B**, reproduced here for easier comparison), and **(D)** hydrolysis transition state asymmetrically-closed bovine CCT from 1 mM ADP-AlFx conditions (PDB: 4a0w) (Cong et al., 2012).
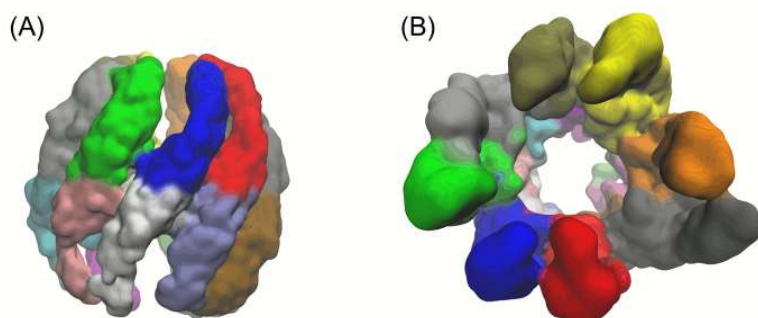
Supplementary Movies
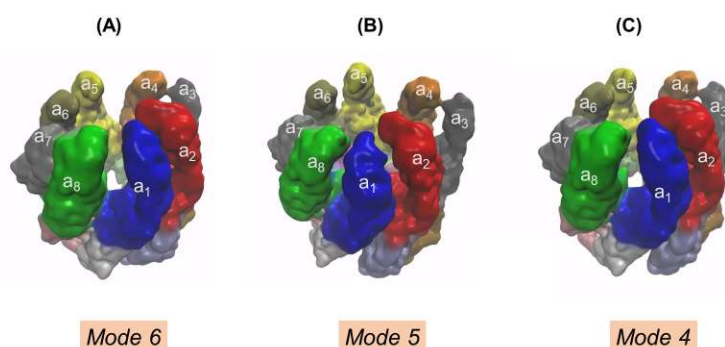
Supplementary Movie 1A-C. Lid closure movements revealed by three ANM modes of TRiC-ATP. The beads reconstructed from EMD-1961 are displaced along ANM *modes 6, 5,* and *4* for an RMSD of 6 Å in both directions to show motions related to upper ring closure.
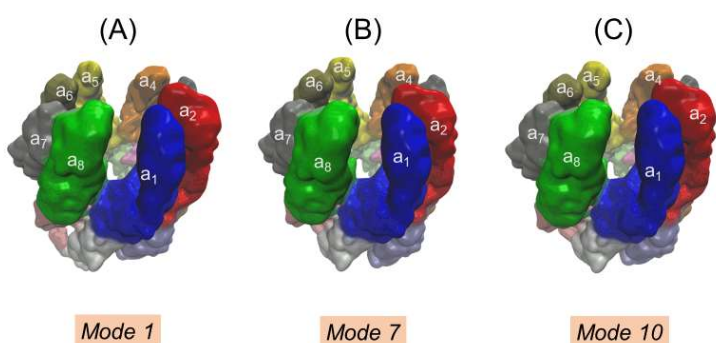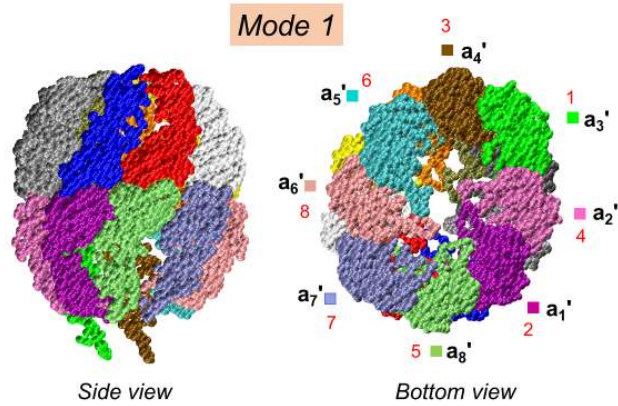

Supplementary Movie 2A-C. ANM *modes 1, 7,* and *10* of TRiC-ATP. The animations have been created using the same method and RMSD as **Supplementary Movie 1A-C**.
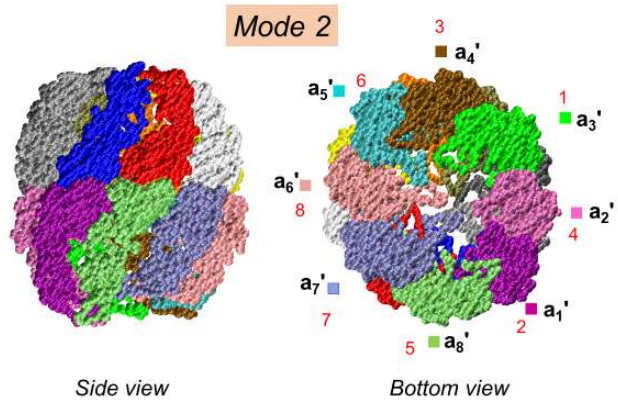

Supplementary Movie 3. Transition from open to closed state generated by adaptive ANM applied to TRiC-ATP. The adaptive ANM is performed starting from the TRN model of *open*, ATP-bound TRiC (EMD-1961), and undergoing gradual reconfigurations in the space of soft modes accessible at various stages of the transition, to reach the half-closed form (PDB: 4a0w), viewed from **(A)** the *side* and **(B)** the *top*.
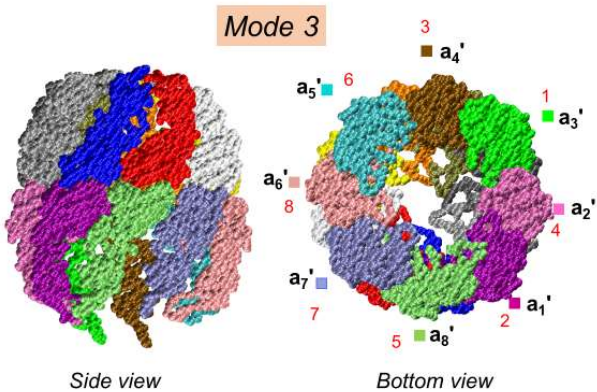
TRIC-ADP-AIFx

Mode 1

Side view   Bottom view

TRIC-ADP-AIFx

Mode 2

Side view   Bottom view

TRIC-ADP-AIFx

Mode 3

Side view   Bottom view
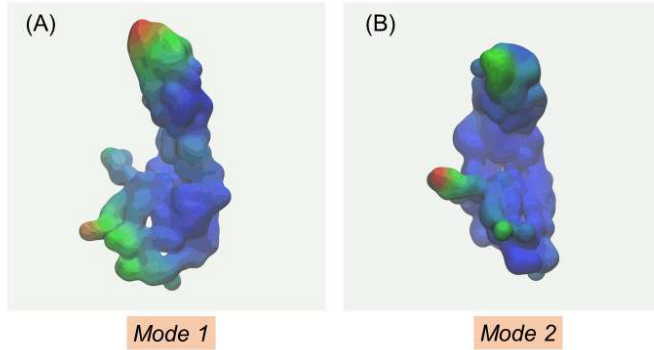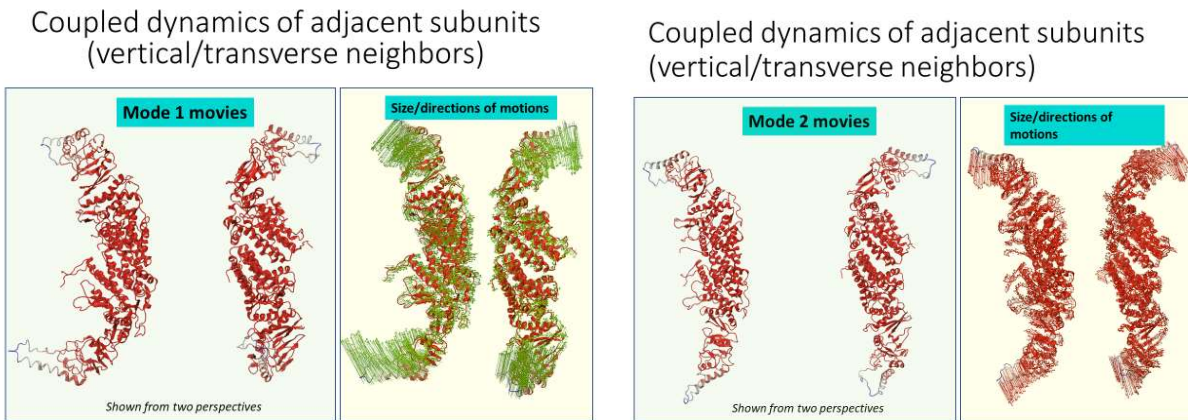
Supplementary Movie 4-6. ANM *modes 1, 2,* and *3* of TRiC-ADP-AIFx. The movies have been generated using the deposited atomic model (PDB: 4a0w). Fluctuations with RMSD of 6 Å with respect to the starting structure are displayed in both directions for each case. The *side* and *bottom views* are shown.

Intrinsic dynamics of CCT subunits: two softest modes

Mode 1   Mode 2

Supplementary Movie 7 A and B. Intrinsic dynamics of CCT subunits. ANM results were generated for the isolated subunit a_1 reconstructed from EMD-1961, in the absence of other subunits. TRN beads are displaced along ANM *modes 1* and *2* by an RMSD of ~6 Å with respect to the starting structure in both directions, shown from two different views that best illustrate the respective motions.

Coupled dynamics of adjacent subunits (vertical/transverse neighbors)

Coupled dynamics of adjacent subunits (vertical/transverse neighbors)

Mode 1 movies   Size/directions of motions

Shown from two perspectives

Mode 2 movies   Size/directions of motions

Shown from two perspectives

Supplementary Movies 8 and 9. Coupled dynamics of pairs of CCT subunits (vertical/transverse dimers). ANM is performed for subunit a_1 and a_1' of the apo TRiC structure (PDB: 4a0o), in the absence of all other subunits, and then the residues are displaced along ANM modes 1 and 2 by an RMSD of 6 Å in both directions. Two different views are shown in each case. The right panels display the directions of motions (*arrows*) in each case.

Supplementary Table

Table S1. Nomenclature adopted for TRiC subunits in different studies/structures (*)

	CCT4/δ antibody-labelled difference EM (Rivenzon-Segal et al., 2005) Figure 4B	CCT1/α antibody-labelled difference EM (Rivenzon-Segal et al., 2005) Figure 4F	4Å resolution cryo-EM (Cong et al., 2010)	Closed form crystal structure (Dekker et al., 2011)	Symmetry-free cryo-EM structure (Cong et al., 2012)	Associated PDB models e.g. 4A0V (Cong et al., 2012)	XL-MS (Leitner et al., 2012)	Associated PDB models e.g. 4V94 (Leitner et al., 2012)	Yeast cryo-EM models e.g. 5GW4 (Jin et al., 2019a; Zang et al., 2016)	Structures with PFD e.g. 6NRA (Gestaut et al., 2019)
Upper ring/cis ring	CCT7/η	CCT6/ζ	CCT1	CCT4/δ	a_1	A	CCT2/β	B	B	B
	CCT4/δ	CCT5/ε	CCT6	CCT8/θ	a_2	C	CCT4/δ	D	D	D
	CCT8/θ	CCT1/α	CCT2	CCT3/γ	a_3	H	CCT1/α	A	A	A
	CCT3/γ	CCT7/η	CCT3	CCT2/β	a_4	D	CCT3/γ	C	G	C
	CCT2/β	CCT4/δ	CCT8	CCT6/ζ	a_5	E	CCT6/ζ	F	Z	F
	CCT6/ζ	CCT8/θ	CCT4	CCT5/ε	a_6	G	CCT8/θ	H	Q	H
	CCT5/ε	CCT3/γ	CCT5	CCT1/α	a_7	B	CCT7/η	G	H	G
	CCT1/α	CCT2/β	CCT7	CCT7/η	a_8	F	CCT5/ε	E	E	E
Lower ring/trans ring	CCT7/η	CCT6/ζ	CCT1	CCT4/δ	a_1'	M	CCT2/β	b	b	J
	CCT4/δ	CCT5/ε	CCT6	CCT8/θ	a_2'	O	CCT4/δ	d	d	L
	CCT8/θ	CCT1/α	CCT2	CCT3/γ	a_3'	I	CCT1/α	a	a	I
	CCT3/γ	CCT7/η	CCT3	CCT2/β	a_4'	N	CCT3/γ	c	g	K
	CCT2/β	CCT4/δ	CCT8	CCT6/ζ	a_5'	J	CCT6/ζ	f	z	N
	CCT6/ζ	CCT8/θ	CCT4	CCT5/ε	a_6'	K	CCT8/θ	h	q	P
	CCT5/ε	CCT3/γ	CCT5	CCT1/α	a_7'	P	CCT7/η	g	h	O
	CCT1/α	CCT2/β	CCT7	CCT7/η	a_8'	L	CCT5/ε	e	e	M

(*) Highlighted are the two nomenclatures used in this study.

Supplementary Methods

ANM Analysis

The Hessian matrix **H** was built based on either TRN nodes from density maps or C^α-atoms from atomic structures, using a cutoff of 20 Å or 15 Å, respectively, as appropriate for the level of coarse-graining (Atilgan et al., 2001; Doruker et al., 2002). In the case of the asymmetric ADP-AlFx-bound structure (PDB: 4a0w), we used homology modelling to fill in the missing residues with the assistance of the SWISS-MODEL server (Waterhouse et al., 2018).

Aside from the conventional ANM, where uniform spring constants are used (Atilgan et al., 2001), we also studied two other ANMs with distance dependent force constant (Hinsen, 1998; Yang et al., 2009a). The distance-dependent spring constant used by Hinsen is given by an exponential function (referred to as the "exponential spring constant" in the following) of the form $k(\boldsymbol{r}) \propto \exp(-|\boldsymbol{r}|^2/r_0)$, whilst Yang et al. used an inverse square function of the form $k(\boldsymbol{r}) \propto 1/|\boldsymbol{r}|^2$, in their so-called parameter-free ANM (pfANM).

Eigenvalue decomposition of **H** for a system of N particles yields $3N$-6 non-zero eigenvalues, the first (smallest) k of which are designated as

$$\lambda_1, \lambda_2, \dots, \lambda_k \tag{S1}$$

and corresponding eigenvectors, i.e., ANM modes, denoted as $3N$ dimensional column vectors:

$$\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k \tag{S2}$$

Each eigenvector is of the form:

$$\vec{v}_i = \begin{bmatrix} x_i^{(1)} & y_i^{(1)} & z_i^{(1)} & x_i^{(2)} & y_i^{(2)} & z_i^{(2)} & \dots & x_i^{(3N)} & y_i^{(3N)} & z_i^{(3N)} \end{bmatrix}^T \tag{S3}$$

where the superscripts denote the indices of particles, and the elements designate the normalized displacements along the principal axis defined by mode i.

For a single mode i, the $3N$ x $3N$ covariance matrix $\mathbf{\Sigma}_i$ is computed as follows:

$$\mathbf{\Sigma}_i = \frac{1}{\lambda_i} \vec{v}_i \otimes \vec{v}_i \qquad \text{(S4)}$$

where $\otimes$ is the outer product. $\mathbf{\Sigma}_i$ can be written as a block matrix of N x N blocks, such that each block, or super-element, corresponds to a 3 x 3 matrix for a pair of residues α and β:

$$\mathbf{\Sigma}_i^{(\alpha\beta)} = \begin{bmatrix} x_i^{(\alpha)} x_i^{(\beta)} & x_i^{(\alpha)} y_i^{(\beta)} & x_i^{(\alpha)} z_i^{(\beta)} \\ y_i^{(\alpha)} x_i^{(\beta)} & y_i^{(\alpha)} y_i^{(\beta)} & y_i^{(\alpha)} z_i^{(\beta)} \\ z_i^{(\alpha)} x_i^{(\beta)} & z_i^{(\alpha)} y_i^{(\beta)} & z_i^{(\alpha)} z_i^{(\beta)} \end{bmatrix} \qquad \text{(S5)}$$

The trace of this super-element is the covariance of residues α and β. By taking the trace of all the super-elements in $\mathbf{\Sigma}_i$, we obtain the N x N covariance matrix $\mathbf{C}_i$ describing the cross-correlations between the motions of all residue pairs along mode i, written in concise form as

$$\mathbf{C}_i = \left[\mathbf{C}_i^{(\alpha\beta)} \right]_{\alpha,\beta=1}^{N} = \left[\mathrm{Tr}\left(\mathbf{\Sigma}_i^{(\alpha\beta)} \right) \right]_{\alpha,\beta=1}^{N} \qquad \text{(S6)}$$

The N x N covariance matrix of the first k nonzero modes, is computed as:

$$\mathbf{C} = \sum_{i=1}^{k} \mathbf{C}_i \qquad \text{(S7)}$$

The sum of square-fluctuations of the first k modes is simply the diagonal elements of $\mathbf{C}$. Its normalization results in orientational cross-correlations.

In order to compare normal modes between atom-based models and TRN models, which have different numbers of atoms, we used the reduced model framework (Hinsen et al., 2000; Ming and Wall, 2005; Zheng and Brooks, 2005) as implemented in *ProDy*, allowing the motions of a subset of nodes (system) to be calculated while incorporating the effect of the remainder (environment).

First, the full Hessian $\mathbf{H}$ is divided into blocks:

$$H = \begin{pmatrix} H_{SS} & H_{SE} \\ H_{ES} & H_{EE} \end{pmatrix} \qquad \text{(S8)}$$

where H_{SS} is the portion corresponding to the system, H_{EE} is the portion corresponding to the environment, and H_{SE} and H_{ES} are the cross-terms corresponding to their interactions.

These blocks are used to create an effective Hessian for the system $H_{SS}^{eff} = H_{SS} - H_{SE}H_{EE}^{-1}H_{ES}$, which is used in place of the regular Hessian $\mathbf{H}$.

Alignment of TRN and Atomic models

We used a nearest neighbour approach, which relies simply on the 3D coordinates of the two sets of nodes and is therefore applicable to any two sets of nodes in 3D space. Each TRN node was matched with the nearest atom, in this case a subset of just under a third of the ~8000 α-carbons atoms for the ~3000 TRN beads. The copying of residue names and numbers, atom names and numbers, and chain IDs from the atoms to the beads allowed us to reorder the beads and interface with the existing mapping tools within ProDy and use all the capabilities of that API.

Markovian Network Analysis

We assume a discrete-time, discrete-state Markov process of allosteric information across the network of nodes, based on which the conditional probability of transmitting information from node i to j in one timestep is defined as

$$m_{ij} = \frac{a_{ij}}{d_j} \tag{S9}$$

where a_{ij} is an entry in the adjacency matrix, and d_j is the degree of the node j. The adjacency matrix was built using the GNM (C^α atom- or TRN node-based) Kirchhoff matrix with a cutoff of 10 Å (Bahar et al., 1997).

Hitting time, $H(j, i)$, is defined as the expected number of steps required to transduce a signal from node i to node j. It has been shown that $H(j, i)$ can be evaluated recursively (Chennubhotla and Bahar, 2006):

$$H(j, i) = \sum_{k=1}^{n}[1 + H(j, k)]m_{ki} = 1 + \sum_{k=1, k \neq j}^{n} H(j, k)m_{ki} \tag{S10}$$

Since we are interested in the signal transduction between subunits, which are collections of nodes in a network, we define the subunit-averaged hitting time between subunits a and b, as an average over the hitting time associated with all pairs of nodes belonging to these subunits:

$$\overline{H}(b, a) = \frac{\sum_{j \in b} \sum_{i \in a} H(j,i)}{\sum_{j \in b} \sum_{i \in a} 1} \tag{S11}$$

The TRN results are the average from replica runs of the algorithm, producing three similar models.

Subunit Identification using Self-Organizing Map

In the first step, we used SOM to cluster the beads into 8 clusters, and then the k-means algorithm (implemented in scikit-learn (Pedregosa et al., 2011)) was applied to each cluster which subdivided them into 4 sub-clusters, containing the equatorial and apical domains of the paired subunits, consistent with the intrinsic behaviour of the single subunit (**Figure 7B**). The two sub-clusters of the same subunit were then combined into one cluster, which led to a total 16 clusters. The results are similar to those from SC, however SOM generated clusters with disjoint nodes at the apical domains (see *red* arrow in **Figure S2A**), which needs to be corrected manually before applying k-means.

Adaptive ANM (aANM) and ClustENM

A modified aANM has been adopted, which enabled increased computational efficiency. Accordingly, each initial structure was mapped to a corresponding target structure, which we refer to as initial and target conformations. These end conformations were used to calculate a deformation vector $\boldsymbol{d}$, describing the differences in the node positions between them. Conformation A (either generated by TRN or taken from the PDB) was used for evaluating ANM modes (20 in the first cycle). The overlaps (correlation cosines) between normal modes and $\boldsymbol{d}$ were calculated, and normal modes were ranked based on these values with the highest overlapping modes having the highest ranking. Then a number of top-ranking modes were selected which achieved a cumulative correlation threshold of 0.4. The structure was then deformed along a combination of

these modes that maximally brought it towards the target structure scaled by the fraction 0.2 that avoids unrealistic deformations as previously described (Yang et al., 2009b). This process was repeated with new deformation vectors and ANM modes being calculated in each cycle, providing adaptive selection of new modes at each cycle. Later cycles involved increasing larger numbers of (higher frequency) modes each time modes being selected towards the end of the frequency window previously used, e.g. when modes with indices near 20 were used the number of modes calculated (and hence selected) in the next cycle was increased. This allowed for higher computational efficiency than the original algorithm where the same number of modes were calculated in each cycle; lower frequency, energetically favourable modes drive much of the transition and higher frequency ones then provide local optimization towards the end. aANM was run for about 25-30 cycles until the RMSD between the newest intermediate structure and the target fell below 1 Å. When using the TRN beads as the initial conformation, we assigned residue labels to the beads based on the closest α-carbons in the deposited atomic model (PDB: 4a0v) as discussed above, which allowed us to identify corresponding nodes in the target.

In ClustENM generation of new conformers, the known structure was deformed along a set of linear combinations of the first five slowest ANM modes to generate 3^5 conformers (considering positive and negative direction motions, and no motion, along a given mode, in each case). We obtained 6 clusters using the k-means algorithm from these conformers and the conformers with the lowest RMSD to the cluster centroids were used as representative conformations for each cluster. The representatives were deformed along their slowest modes again in the next generation, and more conformers were generated, which were again divided into 6 clusters. We performed three such successive generations and obtained a total of $6^3 = 216$ conformers. The energy minimization procedure of the original algorithm was omitted here, since the conformers were not generated at atomic detail. We adopted small RMSDs (of 1 Å) at each generation to avoid unrealistic deformations.

17

Supplementary References

Atilgan, A.R., Durell, S.R., Jernigan, R.L., Demirel, M.C., Keskin, O., and Bahar, I. (2001). Anisotropy of fluctuation dynamics of proteins with an elastic network model. Biophys J *80*, 505-515.

Bahar, I., Atilgan, A.R., and Erman, B. (1997). Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. Fold Des *2*, 173-181.

Chennubhotla, C., and Bahar, I. (2006). Markov propagation of allosteric effects in biomolecular systems: application to GroEL-GroES. Mol Syst Biol *2*, 36.

Cong, Y., Baker, M.L., Jakana, J., Woolford, D., Miller, E.J., Reissmann, S., Kumar, R.N., Redding-Johanson, A.M., Batth, T.S., Mukhopadhyay, A.*, et al.* (2010). 4.0-A resolution cryo-EM structure of the mammalian chaperonin TRiC/CCT reveals its unique subunit arrangement. Proc Natl Acad Sci U S A *107*, 4967-4972.

Cong, Y., Schroder, G.F., Meyer, A.S., Jakana, J., Ma, B., Dougherty, M.T., Schmid, M.F., Reissmann, S., Levitt, M., Ludtke, S.L.*, et al.* (2012). Symmetry-free cryo-EM structures of the chaperonin TRiC along its ATPase-driven conformational cycle. EMBO J *31*, 720-730.

Dekker, C., Roe, S.M., McCormack, E.A., Beuron, F., Pearl, L.H., and Willison, K.R. (2011). The crystal structure of yeast CCT reveals intrinsic asymmetry of eukaryotic cytosolic chaperonins. EMBO J *30*, 3078-3090.

Doruker, P., Jernigan, R.L., and Bahar, I. (2002). Dynamics of large proteins through hierarchical levels of coarse-grained structures. J Comput Chem *23*, 119-127.

Gestaut, D., Roh, S.H., Ma, B., Pintilie, G., Joachimiak, L.A., Leitner, A., Walzthoeni, T., Aebersold, R., Chiu, W., and Frydman, J. (2019). The Chaperonin TRiC/CCT Associates with Prefoldin through a Conserved Electrostatic Interface Essential for Cellular Proteostasis. Cell *177*, 751-765 e715.

Hinsen, K. (1998). Analysis of domain motions by approximate normal mode calculations. Proteins *33*, 417-429.

Hinsen, K., Petrescu, A., Dellerue, S., Bellissent-Funel, M., and Kneller, G.R. (2000). Harmonicity in slow protein dynamics. Chem Phys *261*, 25-37.

Jin, M., Han, W., Liu, C., Zang, Y., Li, J., Wang, F., Wang, Y., and Cong, Y. (2019a). An ensemble of cryo-EM structures of TRiC reveal its conformational landscape and subunit specificity. Proc Natl Acad Sci U S A *116*, 19513-19522.

Jin, M., Liu, C., Han, W., and Cong, Y. (2019b). TRiC/CCT Chaperonin: Structure and Function. Subcell Biochem *93*, 625-654.

Kalisman, N., Schroder, G.F., and Levitt, M. (2013). The crystal structures of the eukaryotic chaperonin CCT reveal its functional partitioning. Structure *21*, 540-549.

Leitner, A., Joachimiak, L.A., Bracher, A., Monkemeyer, L., Walzthoeni, T., Chen, B., Pechmann, S., Holmes, S., Cong, Y., Ma, B.*, et al.* (2012). The molecular architecture of the eukaryotic chaperonin TRiC/CCT. Structure *20*, 814-825.

Ming, D., and Wall, M.E. (2005). Allostery in a coarse-grained model of protein dynamics. Phys Rev Lett *95*, 198103.

Munoz, I.G., Yebenes, H., Zhou, M., Mesa, P., Serna, M., Park, A.Y., Bragado-Nilsson, E., Beloso, A., de Carcer, G., Malumbres, M.*, et al.* (2011). Crystal structure of the open conformation of the mammalian chaperonin CCT in complex with tubulin. Nat Struct Mol Biol *18*, 14-19.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V.*, et al.* (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research *12*, 2825-2830.

Rivenzon-Segal, D., Wolf, S.G., Shimon, L., Willison, K.R., and Horovitz, A. (2005). Sequential ATP-induced allosteric transitions of the cytoplasmic chaperonin containing TCP-1 revealed by EM analysis. Nat Struct Mol Biol *12*, 233-237.

Wang, H., Han, W., Takagi, J., and Cong, Y. (2018). Yeast Inner-Subunit PA-NZ-1 Labeling Strategy for Accurate Subunit Identification in a Macromolecular Complex through Cryo-EM Analysis. J Mol Biol *430*, 1417-1425.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L.*, et al.* (2018). SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res *46*, W296-W303.

Yang, L., Song, G., and Jernigan, R.L. (2009a). Protein elastic network models and the ranges of cooperativity. Proc Natl Acad Sci U S A *106*, 12347-12352.

Yang, Z., Majek, P., and Bahar, I. (2009b). Allosteric transitions of supramolecular systems explored by network models: application to chaperonin GroEL. PLoS Comput Biol *5*, e1000360.

Zang, Y., Jin, M., Wang, H., Cui, Z., Kong, L., Liu, C., and Cong, Y. (2016). Staggered ATP binding mechanism of eukaryotic chaperonin TRiC (CCT) revealed through high-resolution cryo-EM. Nat Struct Mol Biol *23*, 1083-1091.

Zang, Y., Wang, H., Cui, Z., Jin, M., Liu, C., Han, W., Wang, Y., and Cong, Y. (2018). Development of a yeast internal-subunit eGFP labeling strategy and its application in subunit identification in eukaryotic group II chaperonin TRiC/CCT. Sci Rep *8*, 2374.

Zheng, W., and Brooks, B.R. (2005). Probing the local dynamics of nucleotide-binding pocket coupled to the global dynamics: myosin versus kinesin. Biophys J *89*, 167-178.