Subunit conformational variation within individual GroEL oligomers resolved by Cryo-EM

Soung-Hun Roh[a], Corey F. Hryc[a,b], Hyun-Hwan Jeong[c], Xue Fei[d], Joanita Jakana[a], George H. Lorimer[d], and Wah Chiu[a,b,1]

[a]National Center for Macromolecular Imaging, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030; [b]Graduate Program in Structural and Computational Biology and Molecular Biophysics, Baylor College of Medicine, Houston, TX 77030; [c]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030; and [d]Department of Chemistry and Biochemistry, University of Maryland, College Park, MD 20742

Contributed by Wah Chiu, June 13, 2017 (sent for review March 22, 2017; reviewed by Xiaochen Bai and Kenneth H. Downing)

Single-particle electron cryo-microscopy (cryo-EM) is an emerging tool for resolving structures of conformationally heterogeneous particles; however, each structure is derived from an average of many particles with presumed identical conformations. We used a 3.5-Å cryo-EM reconstruction with imposed D7 symmetry to further analyze structural heterogeneity among chemically identical subunits in each GroEL oligomer. Focused classification of the 14 subunits in each oligomer revealed three dominant classes of subunit conformations. Each class resembled a distinct GroEL crystal structure in the Protein Data Bank. The conformational differences stem from the orientations of the apical domain. We mapped each conformation class to its subunit locations within each GroEL oligomer in our dataset. The spatial distributions of each conformation class differed among oligomers, and most oligomers contained 10–12 subunits of the three dominant conformation classes. Adjacent subunits were found to more likely assume the same conformation class, suggesting correlation among subunits in the oligomer. This study demonstrates the utility of cryo-EM in revealing structure dynamics within a single protein oligomer.

cryo-EM | focused classification | chaperonin | atomic-model per-particle | conformational variation

Understanding the function of a molecular machine typically requires determination of the structural components at atomic or near-atomic resolution. Although X-ray crystallography can provide atomic details, the crystal lattice forces may unnaturally constrain the structure. Moreover, solution methods, such as NMR spectroscopy, can reveal short-range dynamic behavior in the absence of information on the entire complex (1, 2). Electron cryo-microscopy (cryo-EM) has made substantial progress toward achieving atomic resolution for a broad range of molecular machines by averaging thousands to millions of single-particle images that are assumed to have identical conformations (3–6).

Cryo-EM has also been used to determine multiple structures of a single specimen preparation containing conformational and compositional heterogeneity (7, 8). The focused classification and refinement approaches (7, 8) have been used for two primary purposes: sorting out localized structural heterogeneity among particles (9, 10) and improving feature resolvability of flexible domains through local refinement (11, 12). These molecular machines frequently are composed of multiple protein subunits that can generate correlated and/or uncorrelated motions (13), and no experimental technique has yet captured the atomic structures of individual subunits within a single molecular machine in solution. GroEL, *Escherichia coli* chaperonin (13, 14), are composed of 14 chemically identical protomers (13, 14) and have been structurally characterized by both X-ray crystallography (15, 16) and cryo-EM (5, 17–19). However, these studies have not revealed any heterogeneity among the protomers in each oligomer. Here we used a 3D classification strategy for cryo-EM data (10) to investigate structural heterogeneity among chemically identical subunits and their conformational variation and correlation in a single GroEL.

Results

Cryo-EM Map and Model of GroEL. We reconstructed GroEL from a set of 260 micrographs of apo-GroEL containing 37,367 particle images with a detectable signal close to 3 Å (Fig. 1 *A* and *B*, Fig. S1 *A*–*C*, and Tables S1 and S2) (*Materials and Methods*). Imposing D7 symmetry resulted in a resolution of 3.5 Å (Fig. 1*C* and Figs. S1*D* and S2*A*). At this resolution, we could unambiguously isolate and segment the 14 individual protein subunits.

To analyze the map, we rigid-body fit an atomic model [Protein Data Bank (PDB) ID code 1SS8] (15) to the 3.5-Å cryo-EM density map and then optimized it to improve the fit-to-density while maintaining proper stereochemistry (Fig. 1*D*, Fig. S2*B*, and Tables S1 and S2) (*Materials and Methods*) (20). The optimized model was structurally similar to previously published crystal structures (15, 21, 22) (PDB ID codes 1SS8, 1GRL, and 1OEL), with an rmsd of ∼0.7 Å at the C-alpha backbone level. The local resolution (23) of the equatorial domain of the new model was close to 3 Å, whereas that of the apical domain was ∼4 Å (Fig. S3*A*). Furthermore, the nonuniform local resolution of the new model was similar to the variation in atomic displacement parameters (ADPs) in the 1OEL model derived from X-ray crystallography (Fig. S3*B*). High apical-domain ADP distributions have been found in multiple crystal structures (Fig. S3*B*) (15, 24) and molecular dynamics simulations of apo-GroEL (22). Thus, localized flexibility of the GroEL apical domain is an inherent

Significance

Using cryo-EM and expanding on focused classification allowed for the bacterial chaperone, GroEL, to be structurally resolved at atomic detail one particle at a time. This approach differs from generating the traditional structural ensemble average of presumably identical particle images. Three dominant subunit conformations were identified, and their spatial distributions were found to be different in each individual GroEL oligomer. This strategy demonstrates the application of cryo-EM to characterize the biophysical properties of structural dynamics of individual molecular machines one at a time.

Author contributions: S.R. and W.C. designed research; S.R., C.F.H., H.J., X.F., and J.J. performed research; S.R. and C.F.H. analyzed data; and S.R., C.F.H., H.J., G.H.L., and W.C. wrote the paper.

Reviewers: X.B., University of Texas Southwestern Medical Center; and K.H.D., Lawrence Berkeley National Laboratory.

Conflict of interest statement: W.C. and K.H.D. are coauthors of a 2015 paper, which was a report of a community effort of different cryo-EM experts to evaluate the contrast transfer function parameter determination of different datasets of cryo-EM images using different software. They did not have any direct collaboration on this work.

Freely available online through the PNAS open access option.

Data deposition: The 3D cryo-EM density map has been deposited in the Electron Microscopy Data Bank (accession no. EMD-8750). The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 5W0S).

[1]To whom correspondence should be addressed. Email: wah@bcm.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1704725114/-/DCSupplemental.

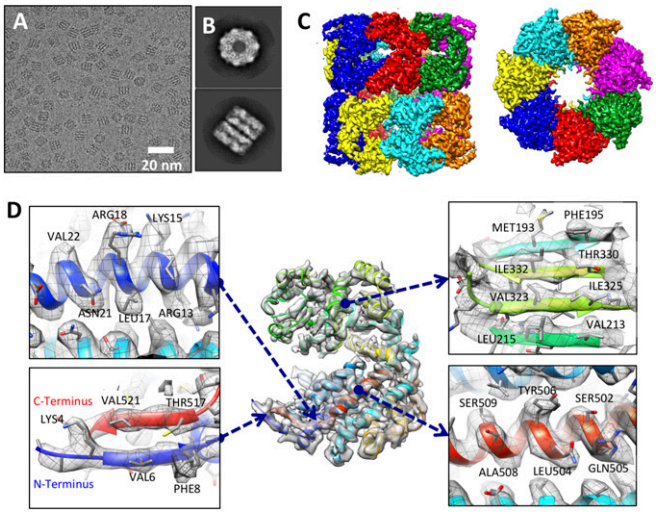

Fig. 1. Cryo-EM map and model of apo-GroEL. (*A*) A typical motion-corrected raw image of vitrified GroEL particles recorded on a Gatan K2 direct detector. (*B*) Reference-free 2D class averages for top and side views. (*C*) Side and end-on views of the 3.5-Å 3D cryo-EM map of apo-GroEL. Each subunit is colored differently (EMD-8750). (*D*) Segmented cryo-EM map for a single subunit with the model superimposed (PDB ID code 5W0S). Representative areas revealing high-resolution features are highlighted in the side panels.

structural characteristic rather than an artifact of crystallization or cryo-specimen freezing.

Structural Variation of the Apical Domain Within and Among GroEL Particles. The high flexibility of the GroEL apical domain raised the question of whether any specific pattern of structural variation exists among the 14 compositionally identical subunits of the GroEL oligomer. A symmetry-free reconstruction would not be sufficiently sensitive to resolve small differences because of overwhelming signals from the entire GroEL complex. Most importantly, the subunit variation might differ from particle to particle, where an asymmetric reconstruction is unlikely to yield an informative solution. Thus, we used a localized 3D classification approach (10, 25) to analyze each of the particle images and determine whether computationally isolated subunits could be resolved into distinct conformations.

We extended the focused classification and localized reconstruction protocol (10, 25) to the analysis of the subunit conformations, as described in *Materials and Methods*. This method allowed us to computationally isolate individual subunits per particle image and divide them into 10 classes according to their conformational similarity (Fig. 2). In this classification process, 71% of the subunits from the entire particle dataset fell into one of the three dominant conformational classes (Fig. S4). We computed 3D maps from each of these three sets of particle images consisting of one subunit from the dominant class that had been computationally translocated to the same location within the oligomer. This resulted in a map with one subunit corresponding to the specific conformation class and the remaining 13 subunits corresponding to the averaged conformation (Fig. 3*A*).

The resulting maps were well resolved, with visible features such as α helices and β strands. Because of their good resolvability, molecular models were built and optimized for the specific subunit in each of the three maps (Fig. 3*B*). The remaining 29% of the subunits fell into seven other classes whose derived maps were insufficiently resolved at the secondary structure level (Fig. 2*D*) and were not included for further analysis, although they may be functionally or chemically relevant. As a control of our computational approach, we carried out a simulated GroEL dataset with a single rigid and identical subunit conformation to the same classification data-processing steps. The end result yielded one predominant conformation, as expected (Fig. S4*B*). To validate our analysis, we performed the following simulations: we generated a GroEL oligomer model with a mixed conformation of our three dominant subunit models at a ratio of 4:5:5. Our 3D classification analysis resulted in three dominant conformations of subunits, corresponding to the three input models. In the absence of noise, the input and output conformational classes were nearly identical to anticipated ratios (Fig. S4*C*); however, when noise was included in the simulated models, additional minor classes were detected.

The isolated subunit maps computed from the images of the three dominant classes displayed almost identical equatorial domains but had visually distinct rotation in their apical domains (Fig. 3*C*). We confirmed these observations by superimposing the three corresponding models and computed the deviation of the C-alpha positions. For instance, the deviation was ∼0.5 Å for the equatorial domain, a value similar to that of published crystal structures (15, 22), but was ∼4 Å in the apical domain, with the greatest deviation of ∼5 Å at the substrate-binding site, specifically helixes H (residues 230–243) and I (residues 257–268) (Fig. 3*C*). These conformational variations derive from a hinge-like motion around the intermediate region of the subunit (Movie S1), affecting the orientation of the apical domain with respect to the

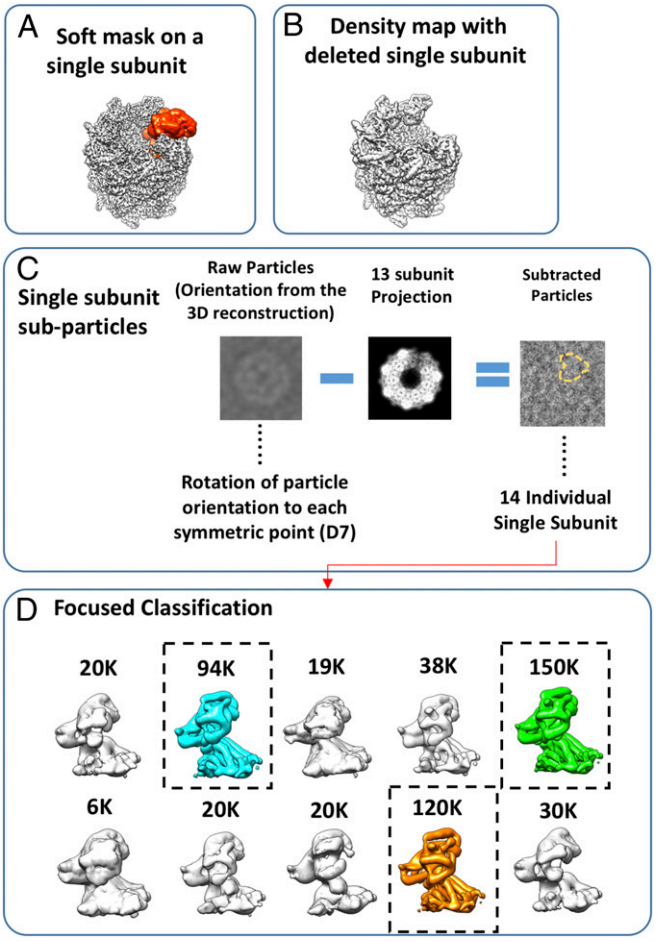

Fig. 2. Single subunit focused 3D classification. (*A*) Location of the soft mask on a single subunit. (*B*) Density map with deleted single subunit. (*C*) Subtraction of the projection including 13 subunits from original particles resulted in single subunit subparticles. (*D*) Subunit-focused classification into 10 groups.

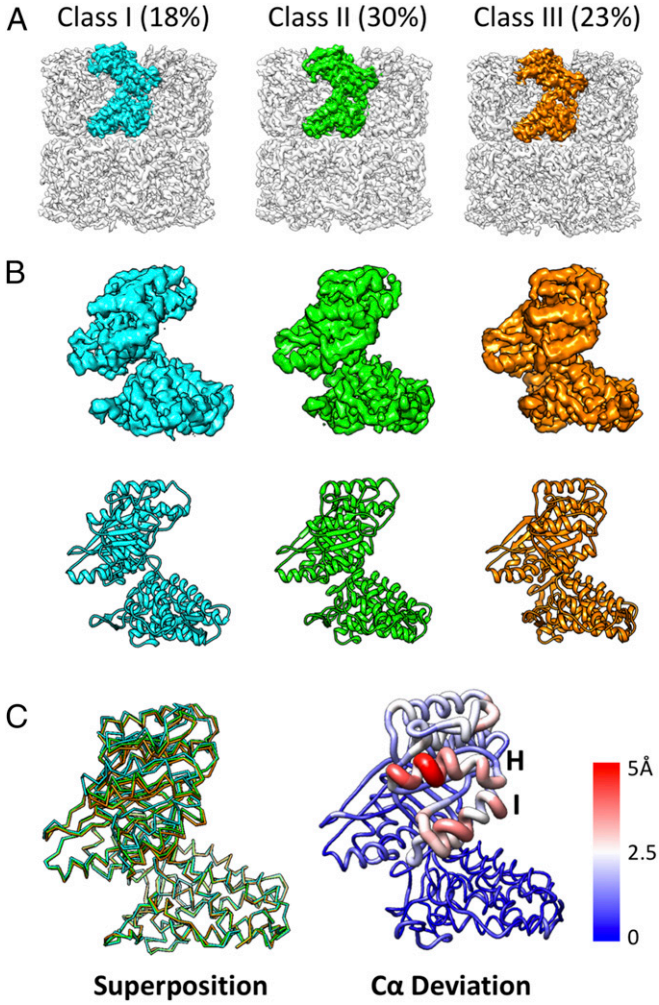

Fig. 3. Conformational variants of GroEL subunits. (*A*) Focused classification and localized reconstruction resulted in three major converged classes of subunits with different populations. (*B*) Segmented subunit maps (EMD-8750) and respective model (PDB ID code 5W0S), rotated with respect to the view shown in *A*. (*C*) Superposition and C-alpha deviation distribution of the three classes derived from the cryo-EM map.

equatorial domain. The apical domain binds to unfolded substrate, and a large conformational change is triggered by ATP binding and hydrolysis at the upper part of the equatorial domain (26).

Three Conformational Variants Match with Three GroEL Crystal Structures. To further assess the validity of the three GroEL subunits derived from the focused classification, we compared the models with the 1XCK crystal structure, which was refined using relaxed noncrystallographic symmetry (NCS) on each subunit. This resulted in 14 discrete conformations (16). The greatest conformational variations occurred in the apical domain, with a maximum C-alpha deviation of ∼5 Å. When aligning the models based on the equatorial domains, we noted a rotational spread of 12.3° among the apical domains (Fig. S5). Our three dominant cryo-EM conformations (Fig. 3*C*) showed similar rotational heterogeneity in their apical domains. We found a close match for each dominant cryo-EM subunit with a corresponding PDB crystal structure (Fig. 4 and Table S3): 1XCK chain J matched the D7-imposed cryo-EM model, 1XCK chain E matched class I, 1XCK chain G matched class II, and 1XCK chain I matched class III. These conformational similarities suggest that the observed cryo-EM conformations are not artifacts generated by the focused classification scheme.

Conformational Distribution in Single GroEL Oligomers. After establishing the existence of three primary conformational variants among the subunits, we traced back the locations of each subunit in the GroEL oligomers (Fig. 5 *A* and *B*). The number of subunits in each conformational class varied from particle to particle, and their relative locations in the GroEL oligomer differed as well (Fig. 5 *A* and *C* and Fig. S6). Almost all of the individual GroEL oligomers comprised a mixture of three, rather than one or two, dominant class conformations in varying ratios (Fig. S6). Furthermore, the majority of oligomers contained 10–12 subunits that fell into one or more of the three conformation classes (Fig. S6). It should be noted that a majority of GroEL oligomers also contained subunits that were not classified into the three dominant classes. In addition, each conformation class was nearly distributed evenly among all subunit locations. These observations indicate that the motion of the apical domain on each apo-GroEL subunit is not globally correlated as an oligomer (14 subunits) or at the level of cis/trans rings (7 subunits). Interestingly, only 687 (1.8%) of the GroEL oligomers in our dataset were composed solely of the three major conformation classes and thus could be used to build atomic models (Fig. 5*C*). Conversely, no oligomer was composed entirely of subunits other than the three major conformation classes.

To explore any possible conformational correlations among the subunits in GroEL, we applied normalized mutual information (NMI) analysis (27), measuring the relative correlation between pairs of conformers. We calculated NMI for each GroEL oligomer containing more than nine subunits (77%) in one of the three major conformation classes (Fig. S6). Fig. 5*D* shows the calculated NMI, revealing the overall correlation in each subunit pair. Correlation was significantly higher between adjacent pairs of subunits than between nonadjacent pairs of subunits (one-sided *t* test, $P < 4.0 \times 10^{-6}$), suggesting that a propensity of adjacent subunits belong to the same conformational class. Such a correlation between adjacent pairs may imply the existence of intersubunit domain–domain interactions, which are possibly influenced by ATP and substrate binding (16, 28, 29).

Discussion

GroEL has two back-to-back seven-subunit rings and undergoes large conformational changes driven by ATP binding and hydrolysis. Our cryo-EM analysis used 37,367 particle images in which the subunits assumed three dominant conformations. These conformations were found to match with those in a published GroEL crystal structure, 1XCK (16). Our observation confirms that the conformers observed in the crystal structure were not caused by a crystallization artifact, given that the cryo-EM structure is free from crystal packing constraints. Interestingly, these conformations are distributed at random in the subunits of each oligomer in ∼71% of our data. However, 29% of the subunits in our dataset were not sufficiently resolved to allow us to establish atomic models associated with them; it is

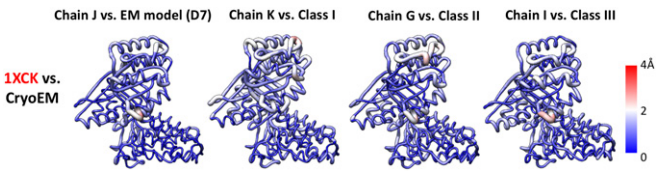

Fig. 4. The three major conformational variants match conformations in the GroEL crystal structure. Distances between C-alphas matched between the cryo-EM models and selected crystal subunit structures from 1XCK.

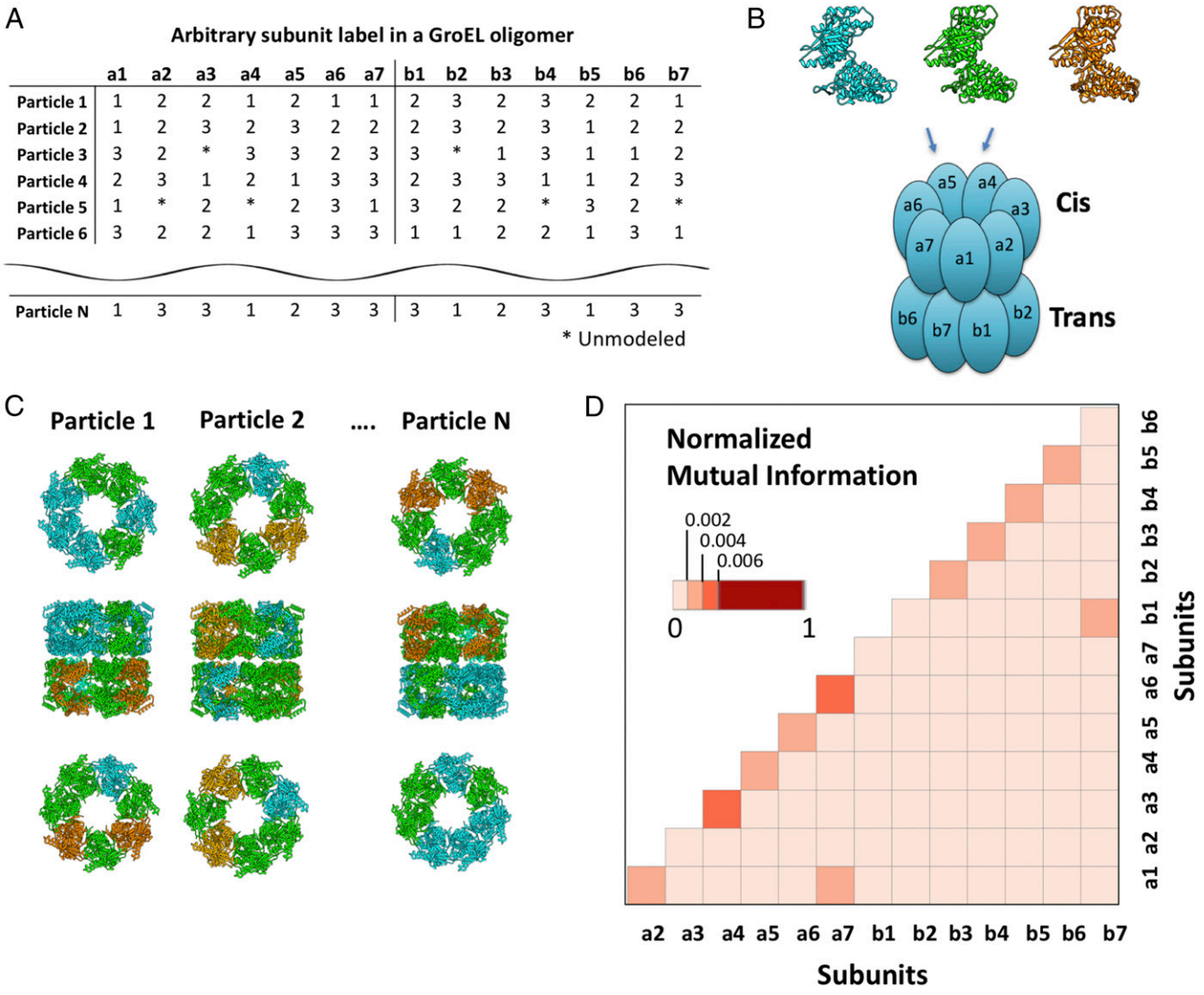

Fig. 5. Structural correlation between adjacent subunits among conformational variants in a single apo-GroEL oligomer. (*A*) Subunit conformations for each GroEL particle. (*B*) Assembly of subunit conformations to the respective subunit location of each GroEL complex. (*C*) Models for one GroEL particle at a time. (*D*) Correlation based on normalized mutual information between pairs of subunits.

possible that these additional conformations could be identified using a larger image dataset.

The primary difference among the three conformers is confined to the apical domains of apo-GroEL (Fig. 3 and Movie S1). Although these motions are limited, they are potentially related to functionality in substrate binding and ATP hydrolysis (15, 28). One intriguing, as-yet unanswered question concerns the possible connections between the apical domain conformational flexibility of the GroEL oligomer and its complex allosteric behavior and/or ability to capture a wide diversity of substrate proteins. A closer inspection of our subunit conformers shows that a salt bridge, D83-K327, formed between the apical and equatorial domains seems to be related to or to respond to the apical domain flexibility; its distance varied significantly from 4 Å to 6 Å among the three dominant cryo-EM conformations (Fig. S7). This distance variability in D83-K327 is similar to that among the matched crystal structures (16). Interestingly, these residues are located in the close proximity of the entry path of ATP to its binding pocket (24). Thus, D83-K327 could be a chemical switch or sensor linking the apical domain flexibility and the ATP-binding pocket. Evidence that binding of the substrate protein to the apical domain is sensed by the nucleotide binding pocket in the equatorial domain comes from the observation that substrate binding accelerates the release of ADP by ~100-fold (30).

Structural fluctuations in proteins can occur simultaneously and at different protein structure levels, including residues, domains, and subunits of macromolecules. Even when the proteins are at chemical equilibrium, slight variations in atom positions exist (31, 32). A robust structural strategy for resolving independent fluctuations, as well as the potential correlations among fluctuations within a single molecular complex, has not yet been demonstrated experimentally. Here we have shown that cryo-EM can be used to decipher the protein dynamics of identical subunits in a molecular machine. Using our protocol, we generated full or partial atomic models for each GroEL oligomer, one particle at a time. This approach may have a wider application for studying oligomeric macromolecules and uncovering different conformational hotspots. Moreover, our method makes it possible to explore the conformational networks within a single protein oligomer, thereby providing the opportunity to study structure-function relationships at a greater level of detail.

Materials and Methods

GroEL Protein Purification. We expressed and purified GroEL as described previously (33). For cryo-EM specimen preparation and imaging, we applied 3-μL aliquots of GroEL (0.1 mg/mL) to glow-discharged holey-carbon Quantifoil R1.2/1.3 grids, blotted them for 2–3 s, and then plunge-froze them in liquid ethane using a Leica EM GP plunge freezer. We transferred the grids into cartridges and loaded them into a JEM 3200FSC (300 KeV) electron microscope with an in-column Ω filter (25 eV energy slit). We recorded images at 0.7–2.5 μm underfocus on a K2 Summit direct electron detector (Gatan) in superresolution mode at nominal 30K magnification, corresponding to a sampling of 1.23 Å/pixel (superresolution sampling of 0.62 Å/pixel).

Image Processing. We initially binned each movie stack by two and then corrected for drift and radiation damage using DE_process_frames.py (Direct Election Ltd.; Fig. S1) (20, 34). The first three frames were ignored during movie processing. We used EMAN2 to automatically select 42,000 particle images from 260 micrographs (35). Contrast transfer function parameters were estimated internally based on the boxed particles (e2ctf.py). We then performed 2D reference-free averaging with ∼2,000 particle images using the default parameters in EMAN2 and generated an initial 3D map based on the 2D class averages, with D7 symmetry imposed. We converted a set of contrast transfer function-corrected particles to a RELION-compatible format using E2refinetorelion3d.py. We performed all further refinements using RELION 1.4 (36), starting with maps that were low-pass filtered to 50 Å from the initial model generated by EMAN2. At first, we used RELION to perform 3D classification ($K = 3$) with C1 symmetry to exclude non-GroEL-like particles. Then we excluded ∼6,000 low-correlated particles and subsequently refined the 37,367 remaining particles using the "auto_refine" command, which resulted in a 3.8-Å map with imposed D7 symmetry. Using the determined particle orientation, we constructed the final map using only frames 3–20 (∼20 $e^-/Å^2$ cumulative exposure). The final resolution of the map was 3.5 Å based on the gold-standard Fourier shell correlation (FSC) at 0.143.

Focused Classification and Localized Reconstruction of Individual Subunits. The overall workflow for the focused classification and localized reconstruction of GroEL subunits is described in Fig. 2 (10, 25). We converted the 3.5-Å map-derived GroEL model into a density map using the e2pdb2mrc.py program of EMAN2. We applied a mask around the GroEL complex using standard automasking in RELION and also generated a mask around a single subunit (Fig. 2*A*). The subtraction of a single-subunit mask from the mask of the entire GroEL complex yielded a mask with one subunit deleted from the entire GroEL. This one subunit-deleted mask was multiplied to the 3.5-Å GroEL map, and the resulting masked map (i.e., a modified GroEL 3D map with only 13 subunits; Fig. 2*B*) then was used for further processing steps.

Using the "relion_particle_symmetry_expand" command (37), we computed the rotational matrix for each of the 14 subunits in each original GroEL particle image, based on the D7 symmetry and the already well-refined orientation parameters for the particles used to compute the 3.5-Å map (Fig. 1*C*). These operations resulted in 14-fold more particle images (hereby called symmetry-expanded images), with each of the 14 subunits aligned to the same orientation. Different projection images from the modified map with only 13 subunits were generated according to the particle orientations images and were subtracted from the respective symmetry-expanded images. This resulted in $14 \times 37{,}367$ subparticle images containing only a signal of a single subunit (called a subparticle hereinafter). These subparticles were subjected to 3D focused classification while keeping all of the orientations fixed at the values determined in the refinement of the original 3.5-Å map. In the classification, 10 classes of 3D maps were computed, each of which showed unique conformations of a single GroEL subunit (Fig. 2*D*). On examination of these maps, three of them showed well-resolved density, such as α-helices and β-strands. The other maps showed less well-resolved features, and we discarded the particles in those classes because they could not be used to build reliable atomic models. For the final map reconstruction, we grouped the symmetry-expanded images in each of the three major classes and performed another round of 3D autorefinement runs without refining orientations (Fig. 3*A*) (36). The resolution of these three maps was 3.7 Å based on a gold standard FSC at 0.143 from two independent datasets.

As a control, we took advantage of a simulated dataset from a GroEL model (PDB ID code 1SS8) that was generated with crystallographic symmetry displaying a single conformation in all of the subunits (10, 15, 25). That dataset consisted of the same number of particles with identical Euler angles and similar signal-to-noise ratio levels compared with our experimental data but was expected to have no heterogeneity at the apical domain, being projected from a single rigid model. We processed the control data the same way that we processed our experimental GroEL data and achieved a 3.8-Å map with imposed D7 symmetry. The control map had no noticeable variation in local resolution between the apical and equatorial domains, and its resolution was more isotropic than that of the experimental GroEL map. We next performed localized classification, splitting the subunits into 10 classes, and found a single dominant population of subunits (Fig. S4). We refined the models based on the D7-imposed map and subtracted single subunits from the same initial model used for the experimental GroEL data. All of the models were very close to 1SS8 and the D7-imposed cryo-EM experimental map. These results further indicate that our results truly reflect the naturally existing conformational variation among single GroEL subunits. In addition, as a positive control, we placed our three cryo-EM subunit conformations at random in the 14-subunit location in the oligomer at a ratio of 4:5:5 and derived a synthetic model of a GroEL complex. Using this model, we generated two simulated particle datasets with/without noise (10), then processed each simulated dataset as previously described (Fig. S4*C*). We found three major populations in both datasets, corresponding to the three cryo-EM subunit conformations used as input.

Model Building. To derive our molecular model, we used Chimera's Fit in Map tool to perform rigid-body fitting of a previously published model (1SS8) (15) into the cryo-EM density map (38). We then used a 20-Å color zone in Chimera to segment the density around a single subunit. To optimize the model, we used Phenix.real_space_refine with default parameters (20), which quickly adjusted the fit-to-density for the subunit. We then rigid-body fit the resulting model into each of the other 13 subunits of the density map. To improve the protein–protein interface, we optimized the complex using Phenix.real_space_refine with the default parameters and the complete 3.5-Å density map as a constraint. We then used Coot (39) to manually adjust loop regions that did not converge into the density. We also adjusted Ramachandran outliers and amino acids with distortions in their bond lengths and angles. A final round of model optimization was performed with Phenix.real_space_refine using new parameters: global minimization, morphing, and atomic displacement. Molprobity (40) statistics were assessed, and the resulting model was complete.

To assess fit-to-density, we derived cross-correlations at the amino acid level and by means of a map/model FSC (Fig. S2). To perform this assessment, we generated a weighted map, derived solely from an atomic model that accounted for both ADP of all atoms and weak/negative density of all charged oxygen atoms, and compared it with the experimental map (41). The weighted map provides a better approximation of the experimental map by simulating map variability as opposed to treating all atoms equally. The correlations for both the FSC and the per-residue assessment showed improvements when properly weighted, further demonstrating that our model provides a good approximation of the experimental data. Finally, we computed an EM-ringer score (42).

Computational Analysis to Capture Correlation Among the Subunits. We used mutual information (MI), which measures the conformational correlation between a pair of subunits (43). The MI is defined by the entropy measure, which represents the uncertainty of information contents for random variables. The entropy of a single subunit, X, was defined as

$$H(X) = -\sum_i p_i \log_2 p_i,$$

where p_i is the marginal probability of class i for subunit X and $i \in \{1,2,3\}$. The joint entropy of two subunits X and Y was defined as

$$H(X, Y) = -\sum_i \sum_j p_{ij} \log_2 p_{ij},$$

where p_{ij} is the joint probability of class i and class j for subunits X and Y. From those definitions, the MI of two subunits X and Y was defined as

$$MI(X, Y) = H(X) + H(Y) - H(X, Y).$$

If there is no correlation between X and Y, then $MI(X, Y) = 0$. A higher MI value implies a stronger correlation between the pair of variables.

Instead of using the MI measure, we used the normalized mutual information (NMI), which is a modified MI measure (27). The NMI converts the maximal value of MI into 1. The NMI was defined by the following equation:

$$NMI(X,Y) = \frac{MI(X,Y)}{min(H(X),H(Y))}.$$

To compare the NMI distribution of the data with a random distribution, we generated 100 replicate shuffled datasets based on the original data. In the replicated data, each particle in the original data was rotated at random.

ACKNOWLEDGMENTS. We thank Drs. Michael F. Schmid, Matthew L. Baker, and Steve J. Ludtke (Baylor College of Medicine) for their valuable advice on image processing and model building, and Dr. Sjors H. Scheres (Medical Research Council, Cambridge, UK) for advice on focused classification analysis. We are grateful for the use of the computing resources at the Extreme Science and Engineering Discovery Environment allocation MCB150009 (ACI-1053575) and the Computational and Integrative Biomedical Research Center at Baylor College of Medicine. This work was supported by the National Institutes of Health (Grants P41GM103832 and R01 GM079429) and the Robert Welch Foundation (Grant Q1242). C.F.H. was supported by a predoctoral fellowship under the National Library of Medicine Training Program in Biomedical Informatics (Grant T15LM007093) awarded to the Keck Center of the Gulf Coast Consortia.

1. Libich DS, Fawzi NL, Ying J, Clore GM (2013) Probing the transient dark state of substrate binding to GroEL by relaxation-based solution NMR. *Proc Natl Acad Sci USA* 110:11361–11366.
2. Fiaux J, Bertelsen EB, Horwich AL, Wüthrich K (2002) NMR analysis of a 900K GroEL-GroES complex. *Nature* 418:207–211.
3. Cheng Y, Grigorieff N, Penczek PA, Walz T (2015) A primer to single-particle cryo-electron microscopy. *Cell* 161:438–449.
4. Liao M, Cao E, Julius D, Cheng Y (2013) Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* 504:107–112.
5. Ludtke SJ, et al. (2008) De novo backbone trace of GroEL from single particle electron cryomicroscopy. *Structure* 16:441–448.
6. Zhang J, et al. (2010) Mechanism of folding chamber closure in a group II chaperonin. *Nature* 463:379–383.
7. Ludtke SJ (2016) Single-particle refinement and variability analysis in EMAN2.1. *Methods Enzymol* 579:159–189.
8. Mazhab-Jafari MT, Rubinstein JL (2016) Cryo-EM studies of the structure and dynamics of vacuolar-type ATPases. *Sci Adv* 2:e1600725.
9. Wang Z, et al. (2017) An allosteric transport mechanism for the AcrAB-TolC multidrug efflux pump. *eLife* 6:e24905.
10. Bai X-C, Rajendra E, Yang G, Shi Y, Scheres SHW (2015) Sampling the conformational space of the catalytic subunit of human γ-secretase. *eLife* 4:e11182.
11. Zhou M, et al. (2015) Atomic structure of the apoptosome: Mechanism of cytochrome c- and dATP-mediated activation of Apaf-1. *Genes Dev* 29:2349–2361.
12. Nguyen TH, et al. (2016) Cryo-EM structure of the yeast U4/U6.U5 tri-snRNP at 3.7 Å resolution. *Nature* 530:298–302.
13. Karplus M, Kuriyan J (2005) Molecular dynamics and protein function. *Proc Natl Acad Sci USA* 102:6679–6685.
14. Hayer-Hartl M, Bracher A, Ulrich Hartl F (2016) The GroEL–GroES chaperonin machine: A nano-cage for protein folding. *Trends Biochem Sci* 41:62–76.
15. Chaudhry C, Horwich AL, Brunger AT, Adams PD (2004) Exploring the structural dynamics of the *E. coli* chaperonin GroEL using translation-libration-screw crystallographic refinement of intermediate states. *J Mol Biol* 342:229–245.
16. Bartolucci C, Lamba D, Grazulis S, Manakova E, Heumann H (2005) Crystal structure of wild-type chaperonin GroEL. *J Mol Biol* 354:940–951.
17. Ludtke SJ, Chen D-H, Song J-L, Chuang DT, Chiu W (2004) Seeing GroEL at 6 Å resolution by single particle electron cryomicroscopy. *Structure* 12:1129–1136.
18. Joseph AP, et al. (2016) Refinement of atomic models in high-resolution EM reconstructions using Flex-EM and local assessment. *Methods* 100:42–49.
19. Clare DK, et al. (2012) ATP-triggered conformational changes delineate substrate-binding and -folding mechanics of the GroEL chaperonin. *Cell* 149:113–123.
20. Wang Z, et al. (2014) Cryo-EM atomic model of brome mosaic virus derived from direct electron detection images and a real-space model optimization protocol. *Biophys J* 106:600a.
21. Braig K, et al. (1994) The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature* 371:578–586.
22. Braig K, Adams PD, Brünger AT (1995) Conformational variability in the refined structure of the chaperonin GroEL at 2.8 Å resolution. *Nat Struct Biol* 2:1083–1094.
23. Kucukelbir A, Sigworth FJ, Tagare HD (2014) Quantifying the local resolution of cryo-EM density maps. *Nat Methods* 11:63–65.
24. Fei X, Yang D, LaRonde-LeBlanc N, Lorimer GH (2013) Crystal structure of a GroEL-ADP complex in the relaxed allosteric state at 2.7 Å resolution. *Proc Natl Acad Sci USA* 110:E2958–E2966.
25. Ilca SL, et al. (2015) Localized reconstruction of subunits from electron cryomicroscopy images of macromolecular complexes. *Nat Commun* 6:8843.
26. Suzuki Y, Yura K (2016) Conformational shift in the closed state of GroEL induced by ATP-binding triggers a transition to the open state. *Biophys Physicobiol* 13:127–134.
27. Estévez PA, Tesmer M, Perez CA, Zurada JM (2009) Normalized mutual information feature selection. *IEEE Trans Neural Netw* 20:189–201.
28. Wang J, Chen L (2003) Domain motions in GroEL upon binding of an oligopeptide. *J Mol Biol* 334:489–499.
29. Wang J, Boisvert DC (2003) Structural basis for GroEL-assisted protein folding from the crystal structure of (GroEL-KMgATP)14 at 2.0 Å resolution. *J Mol Biol* 327:843–855.
30. Ye X, Lorimer GH (2013) Substrate protein switches GroE chaperonins from asymmetric to symmetric cycling by catalyzing nucleotide exchange. *Proc Natl Acad Sci USA* 110:E4289–E4297.
31. Kim SY, Miller EJ, Frydman J, Moerner WE (2010) Action of the chaperonin GroEL/ES on a non-native substrate observed with single-molecule FRET. *J Mol Biol* 401: 553–563.
32. Jiang Y, et al. (2011) Sensing cooperativity in ATP hydrolysis for single multisubunit enzymes in solution. *Proc Natl Acad Sci USA* 108:16962–16967.
33. Rye HS, et al. (1999) GroEL-GroES cycling: ATP and nonnative polypeptide direct alternation of folding-active rings. *Cell* 97:325–338.
34. Bammes BE, Rochat RH, Jakana J, Chen D-H, Chiu W (2012) Direct electron detection yields cryo-EM reconstructions at resolutions beyond 3/4 Nyquist frequency. *J Struct Biol* 177:589–601.
35. Ludtke SJ, Baldwin PR, Chiu W (1999) EMAN: Semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* 128:82–97.
36. Scheres SHW (2012) RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* 180:519–530.
37. Kimanius D, Forsberg BO, Scheres SH, Lindahl E (2016) Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *eLife* 5:e18722.
38. Pettersen EF, et al. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612.
39. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132.
40. Davis IW, Murray LW, Richardson JS, Richardson DC (2004) MOLPROBITY: Structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res* 32:W615–W619.
41. Hryc CF, et al. (2017) Accurate model annotation of a near-atomic resolution cryo-EM map. *Proc Natl Acad Sci USA*.
42. Barad BA, et al. (2015) EMRinger: Side chain-directed model and map validation for 3D cryo-electron microscopy. *Nat Methods* 12:943–946.
43. Steuer R, Kurths J, Daub CO, Weise J, Selbig J (2002) The mutual information: Detecting and evaluating dependencies between variables. *Bioinformatics* 18: S231–S240.