Electronic Supplementary Information

Icosahedral 60-meric porous structure of designed supramolecular protein nanoparticle TIP60

Junya Obata,[‡ab] Norifumi Kawakami,[‡c] Akihisa Tsutsumi,[‡d] Erika Nasu,[c] Kenji Miyamoto,[c] Masahide Kikkawa,[d] and Ryoichi Arai,*[ab]

[a] *Department of Biomolecular Innovation, Institute for Biomedical Sciences, Interdisciplinary Cluster for Cutting Edge Research, Shinshu University, Ueda, Nagano 386-8567, Japan. E-mail: rarai@shinshu-u.ac.jp*
[b] *Department of Applied Biology, Faculty of Textile Science and Technology, Shinshu University, Ueda, Nagano 386-8567, Japan.*
[c] *Department of Bioscience and Informatics, Faculty of Science and Technology, Keio University, Yokohama, Kanagawa 223-8522, Japan.*
[d] *Department of Cell Biology and Anatomy, Graduate School of Medicine, The University of Tokyo, Tokyo 113-0033, Japan.*

[‡] The first three authors contributed equally.

*Corresponding author: Ryoichi Arai
Email: rarai@shinshu-u.ac.jp

This PDF file includes:

Materials and Methods

Protein Expression and Purification

The TIP60 protein was expressed in *E. coli* BL21 Star (DE3) strain harboring the expression plasmid pET-Duet1-TIP60 with 2 L of LB broth (Lennox) (Nacalai Tesque) containing 50 μg/mL ampicillin sodium salt at 37 °C, as previously described.[1] Protein expression was induced by 0.05 mM IPTG at 0.5 of the optical density at 600 nm, and cells were further cultured for 16 h at 37 °C. The harvested cells were disrupted by sonication in 20 mM Tris buffer (pH 7.0) containing 300 mM NaCl. After centrifugation, the TIP60 protein was mainly present in the precipitate fraction. The precipitate fraction was resuspended in 20 mM Tris buffer (pH 8.0) containing 300 mM NaCl, 200 mM Arg-HCl, and 1 mM EDTA. The supernatant containing TIP60 was obtained by centrifugation two times. The TIP60 protein was purified using immobilized metal ion affinity chromatography (IMAC) with cOmplete His-Tag Purification Resin (Roche). The bound protein on the column was washed with 20 mM Tris buffer (pH 8.0) containing 300 mM NaCl, 200 mM Arg-HCl, and 1 mM EDTA, and eluted with 20 mM Tris buffer (pH 8.0) containing 300 mM NaCl, 200 mM Arg-HCl, 1 mM EDTA, and 250 mM imidazole. The eluted protein was dialyzed against 20 mM Tris buffer (pH 8.0) containing 1 mM EDTA and purified by anion exchange chromatography on a RESOURCE Q 6 mL (GE Healthcare) with 20 mM Tris buffer (pH 8.0) containing 1 mM EDTA by a linear gradient of NaCl from 0 to 1.5 M. The TIP60 protein was further purified by size exclusion column chromatography (SEC) on a Superose 6 10/300 GL column (GE Healthcare) with 20 mM Tris buffer (pH 8.0) containing 300 mM NaCl and 1 mM EDTA (Fig. S7).

Cryo-EM

An aliquot of 5 μL of the purified TIP60 sample (5 mg/mL) was applied to glow-discharged holey carbon EM grids (300 mesh). The grids were blotted for 6 s and then plunge-frozen in liquid ethane cooled by liquid nitrogen using a Vitorobot system (Thermo Fisher Scientific). Movies of frozen TIP60 particles embedded in vitreous ice were collected at liquid nitrogen temperature using a Talos Arctica transmission electron microscope (Thermo Fisher Scientific) operated at 200 kV and equipped with a Gatan K2 Summit direct electron detector at the Cryo-EM facility at the University of Tokyo (Tokyo, Japan). Movie stacks were acquired using 130,000× magnification with an accumulated dose of 50 electrons per Å^2 over 40 frames and a defocus range of 0.5–3.0 μm, which resulted in a superresolution pixel size of 1.03 Å per pixel.

Data Processing and 3D Reconstruction

Single-particle cryo-EM data processing and analysis were performed using RELION 3.1.[2] A total of 1,755 movie stacks were motion-corrected using RELION's own implementation.[2] CTF parameters were determined using the CTFFIND4 program.[3] The data processing workflow is summarized in Fig. S2. An initial model was calculated using a stochastic gradient descent algorithm using a subset of the calculated 2D classes. After 3D classification, refinements were performed with icosahedral symmetry (*I*1). A 3.3 Å resolution map was reconstructed using 115,596 particles. The final resolution was estimated using the gold-standard Fourier shell correlation (FSC) between two independently refined half maps (FSC = 0.143). Local resolution was estimated using RELION's own implementation, and a locally filtered and sharpened map was obtained.

Additionally, in further analysis of 3D classification using RELION, four subclasses were observed in 115,596 particles (Fig. S10). For two major subclasses 1 and 4, EM maps were reconstructed using 39,757 and 35,930 particles, respectively. The resolution of the both maps was estimated to be 3.5 Å (FSC = 0.143) (Fig. S11). Superposition of the EM density maps (Fig. S10E) suggested that these maps were almost the same and we did not find significant differences.

Model Building and Refinement

For model building, the EM map was further sharpened using auto-sharpen in Phenix.[4] The atomic structures of the Sm-like protein (PDB ID: 3BY7)[5] and the dimeric coiled-coil domain

of MyoX (PDB ID: 2LW9)[6] were fitted into the appropriate regions of the map corresponding to one protomer of TIP60 using UCSF Chimera[7]. The model of the linker region between the two domains was manually built using Coot.[8] Real-space refinement with secondary structure restraints using Phenix and model rebuilding using Coot were repeated multiple times to achieve an optimized fit between the coordinates and the map. The coordinates fit well with the densities from visual inspection. The icosahedral symmetry (*I*) in the EM map was determined using map_symmetry in Phenix to generate noncrystallographic symmetry (NCS) operators. Subsequently, the icosahedral NCS operators were applied to the TIP60 protomer to build the 60-meric structure of TIP60 using apply_ncs in Phenix. The coordinates were finally refined in real space against the EM map without post-processing and sharpening, using Phenix by applying NCS constraints and secondary structure restraints. The final coordinates were validated using MolProbity[9] in Phenix[10] and wwPDB EM validation report.[11] The cryo-EM model statistics are presented in Table S1. Figures were created using UCSF Chimera.[7] Superposition of the structures was performed using LSQKAB[12] in the CCP4 software suite.[13] The atomic coordinates and cryo-EM maps of TIP60 have been deposited in the Protein Data Bank (PDB ID code 7EQ9) and the Electron Microscopy Data Bank (EMDB ID codes EMD-31256), respectively.

In addition, the EM maps of subclasses 1 and 4 were reconstructed in further analysis of 3D classification as described above. Using the 3.5 Å resolution maps of the subclasses 1 and 4, atomic structure models of TIP60 were further refined. Superposition of the models (Fig. S10F–H) shows that the model structures are almost the same and we cannot find significant differences of conformation between the subclasses and the whole class (root mean square deviation (RMSD) of C_α atoms in all chains: 0.375 Å for the subclass 1 and the whole class, 0.516 Å for the subclass 4 and the whole class; RMSD of Cα atoms in chain A: 0.236 Å for the subclass 1 and the whole class, 0.246 Å for the subclass 4 and the whole class). Therefore, the final structure refined using the 3.3 Å resolution map from all particles (115,596) in the whole class was adopted.

Modification reactions of the interior mutant S50C by DTNB

We used 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB), which comprises two 2-nitro-5-thiobenzoate (TNB) moieties connected by a disulfide bond. DTNB reacts with free thiols by thiol−disulfide exchange, releasing one TNB. The other TNB conjugates with protein (Cys−TNB) and the TNB moiety can be released under reducing conditions by cleaving the disulfide bond. The DTNB modification and TNB release reactions were monitored by measuring the absorption maxima at 323 and 412 nm, attributable to DTNB (including Cys−TNB) and TNB, respectively. The TIP60 S50C mutant was dialyzed against buffer containing 20 mM Tris-HCl and 1 mM EDTA at pH 8.0. The protein (0.6 μM) was reacted with 360 μM DTNB dissolved in the buffer for 20 min at 30 °C. The reacted samples were ultrafiltrated using an Amicon Ultra-15 (MWCO 10 kDa, Merck) to remove excess DTNB and produced TNB. The samples with or without 0.1 M potassium citrate were monitored using a UV-1900 (Shimadzu). TNB released by cleavage of the Cys−TNB linkages was monitored with time by the absorption at 412 nm after the addition of 1 mM TCEP (final concentration) at 30 °C.

References

1. N. Kawakami, H. Kondo, Y. Matsuzawa, K. Hayasaka, E. Nasu, K. Sasahara, R. Arai and K. Miyamoto, *Angew. Chem. Int. Ed.*, 2018, **57**, 12400-12404.
2. J. Zivanov, T. Nakane, B. O. Forsberg, D. Kimanius, W. J. Hagen, E. Lindahl and S. H. Scheres, *eLife*, 2018, **7**, e42166.
3. A. Rohou and N. Grigorieff, *J. Struct. Biol.*, 2015, **192**, 216-221.
4. P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart, *Acta Crystallogr. D*, 2010, **66**, 213-221.

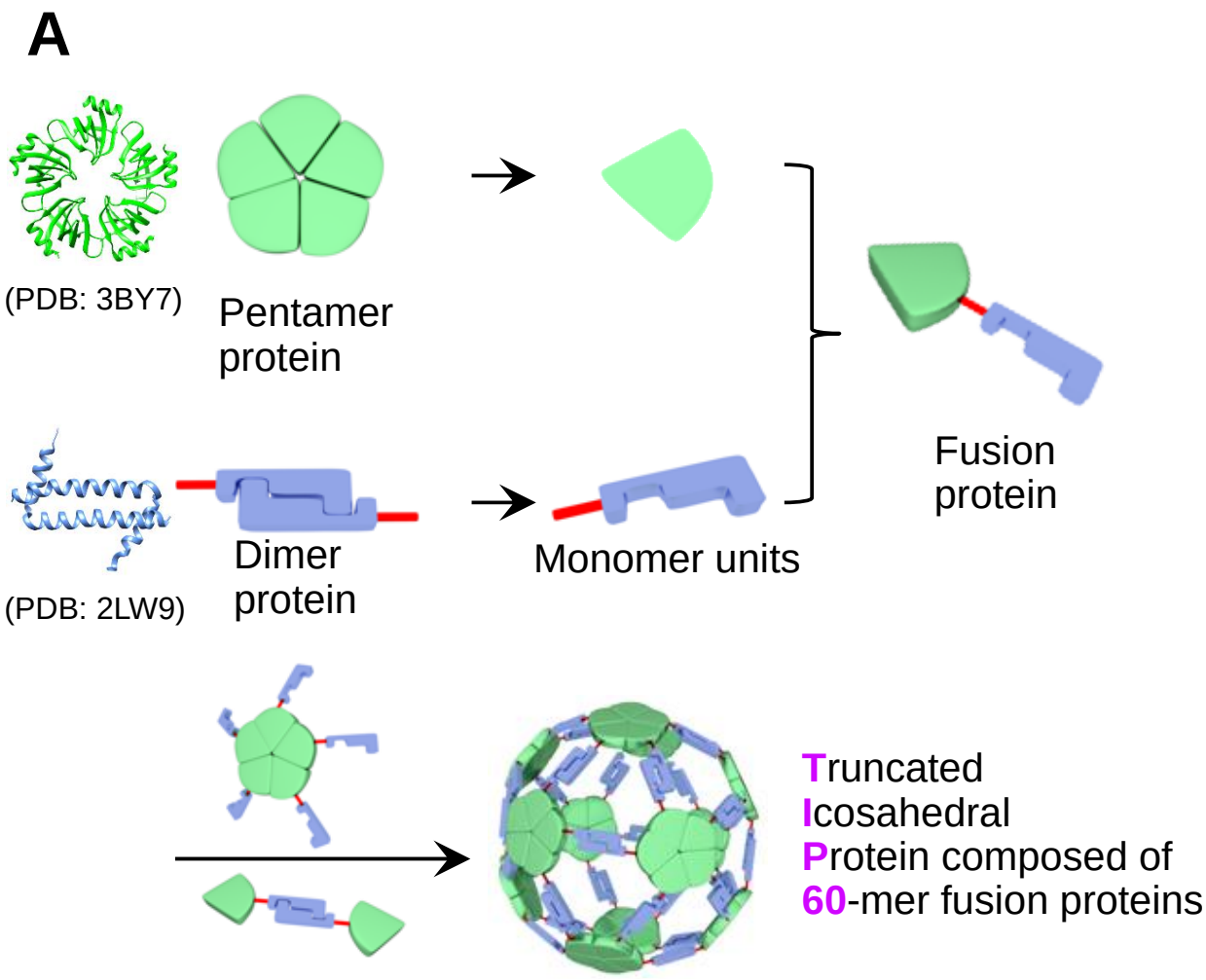
5. D. Das, P. Kozbial, H. L. Axelrod, M. D. Miller, D. McMullan, S. S. Krishna, P. Abdubek, C. Acosta, T. Astakhova, P. Burra, D. Carlton, C. Chen, H. J. Chiu, T. Clayton, M. C. Deller, L. Duan, Y. Elias, M. A. Elsliger, D. Ernst, C. Farr, J. Feuerhelm, A. Grzechnik, S. K. Grzechnik, J. Hale, G. W. Han, L. Jaroszewski, K. K. Jin, H. A. Johnson, H. E. Klock, M. W. Knuth, A. Kumar, D. Marciano, A. T. Morse, K. D. Murphy, E. Nigoghossian, A. Nopakun, L. Okach, S. Oommachen, J. Paulsen, C. Puckett, R. Reyes, C. L. Rife, N. Sefcovic, S. Sudek, H. Tien, C. Trame, C. V. Trout, H. van den Bedem, D. Weekes, A. White, Q. Xu, K. O. Hodgson, J. Wooley, A. M. Deacon, A. Godzik, S. A. Lesley and I. A. Wilson, *Proteins*, 2009, **75**, 296-307.
6. Q. Lu, F. Ye, Z. Wei, Z. Wen and M. Zhang, *Proc. Natl. Acad. Sci. U.S.A.*, 2012, **109**, 17388-17393.
7. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605-1612.
8. P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, *Acta Crystallogr. D*, 2010, **66**, 486-501.
9. V. B. Chen, W. B. Arendall, 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson and D. C. Richardson, *Acta Crystallogr. D*, 2010, **66**, 12-21.
10. P. V. Afonine, B. P. Klaholz, N. W. Moriarty, B. K. Poon, O. V. Sobolev, T. C. Terwilliger, P. D. Adams and A. Urzhumtsev, *Acta Crystallogr. D*, 2018, **74**, 814-840.
11. S. Gore, E. Sanz Garcia, P. M. S. Hendrickx, A. Gutmanas, J. D. Westbrook, H. Yang, Z. Feng, K. Baskaran, J. M. Berrisford, B. P. Hudson, Y. Ikegawa, N. Kobayashi, C. L. Lawson, S. Mading, L. Mak, A. Mukhopadhyay, T. J. Oldfield, A. Patwardhan, E. Peisach, G. Sahni, M. R. Sekharan, S. Sen, C. Shao, O. S. Smart, E. L. Ulrich, R. Yamashita, M. Quesada, J. Y. Young, H. Nakamura, J. L. Markley, H. M. Berman, S. K. Burley, S. Velankar and G. J. Kleywegt, *Structure*, 2017, **25**, 1916-1927.
12. W. Kabsch, *Acta Crystallogr. A*, 1976, **32**, 922-923.
13. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson, *Acta Crystallogr. D*, 2011, **67**, 235-242.
14. X. Robert and P. Gouet, *Nucleic Acids Res.*, 2014, **42**, W320-324.

Table S1. Cryo-EM model statistics.

Model		
Composition (#)		
Chains	60	
Atoms	62520	(Hydrogens: 0)
Residues	Protein: 7620	Nucleotide: 0
Water	0	
Ligands	0	
Bonds (RMSD)		
Length (Å) (# > 4σ)	0.002 (0)	
Angles (°) (# > 4σ)	0.472 (0)	
MolProbity score	1.36	
Clash score	3.70	
Ramachandran plot (%)		
Outliers	0.00	
Allowed	3.25	
Favored	96.75	
Rotamer outliers (%)	0	
Cβ outliers (%)	0	
Peptide plane (%)		
Cis proline/general	0.0/0.0	
Twisted proline/general	0.0/0.0	
CaBLAM outliers (%)	1.68	
ADP (B-factors)		
Iso/Aniso (#)	62520/0	
min/max/mean		
Protein	138.88/334.80/205.47	
Occupancy		
Mean	1	
occ = 1 (%)	100	
0 < occ < 1 (%)	0	
occ > 1 (%)	0	

Data		
Box		
Lengths (Å)	214.75, 214.75, 214.75	
Angles (°)	90.00, 90.00, 90.00	
Supplied Resolution (Å)	3.3	
Resolution Estimates (Å)	Masked	Unmasked
d FSC (half maps; 0.143)	3.4	3.5
d 99 (full/half1/half2)	4.9/4.8/4.8	4.9/4.9/4.9
d model	3.2	3.2
d FSC model (0/0.143/0.5)	3.2/3.3/3.6	3.2/3.4/3.8
Map min/max/mean	-0.01/0.06/0.00	

Model vs. Data	
CC (mask)	0.88
CC (box)	0.87
CC (peaks)	0.84
CC (volume)	0.87



B

MGSSHHHHHHSQDPKNIKIMRLVTGEDIIGNIS
ESQGLITIKKAFVIIPMQATPGKPVQLVLSPWQ
PYTDDKEIVIDDSKVITITSPKDDIIKSYESHT
RVLENKQVEEILRLEKEIEDLQRMKEQQELSLT
EASLQKLQERRDQELRRLEEE

Fig. S1. Design and construction of TIP60.
(*A*) Design strategy of TIP60. Ribbon representation of a pentameric Sm-like protein (PDB ID: 3BY7)[5] and a dimeric MyoX-coil domain (PDB ID: 2LW9)[6] are shown on the left side. [Reprinted with permission from (N. Kawakami *et al.*, *Angew. Chem. Int. Ed.*, 2018, **57**, 12400-12404). Copyright (2018) John Wiley and Sons.]
(*B*) The full-length amino acid sequence of a TIP60 protein. An N-terminal His_6-tag, a Sm-like domain, a linker, and a MyoX-coil domain are shown in black, green, red, and blue, respectively. (Theoretical pI / Mw: 5.51 / 17766.22)

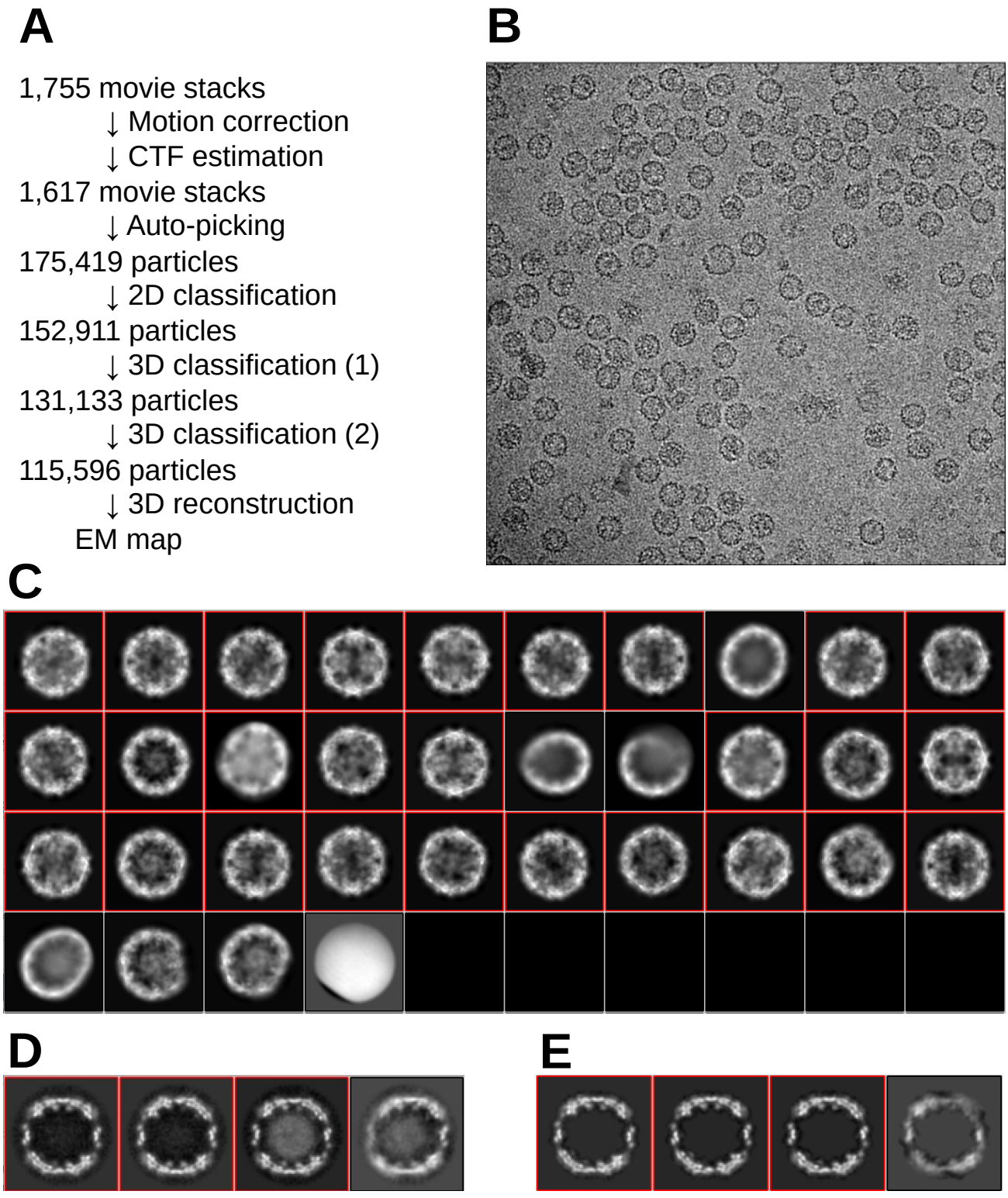
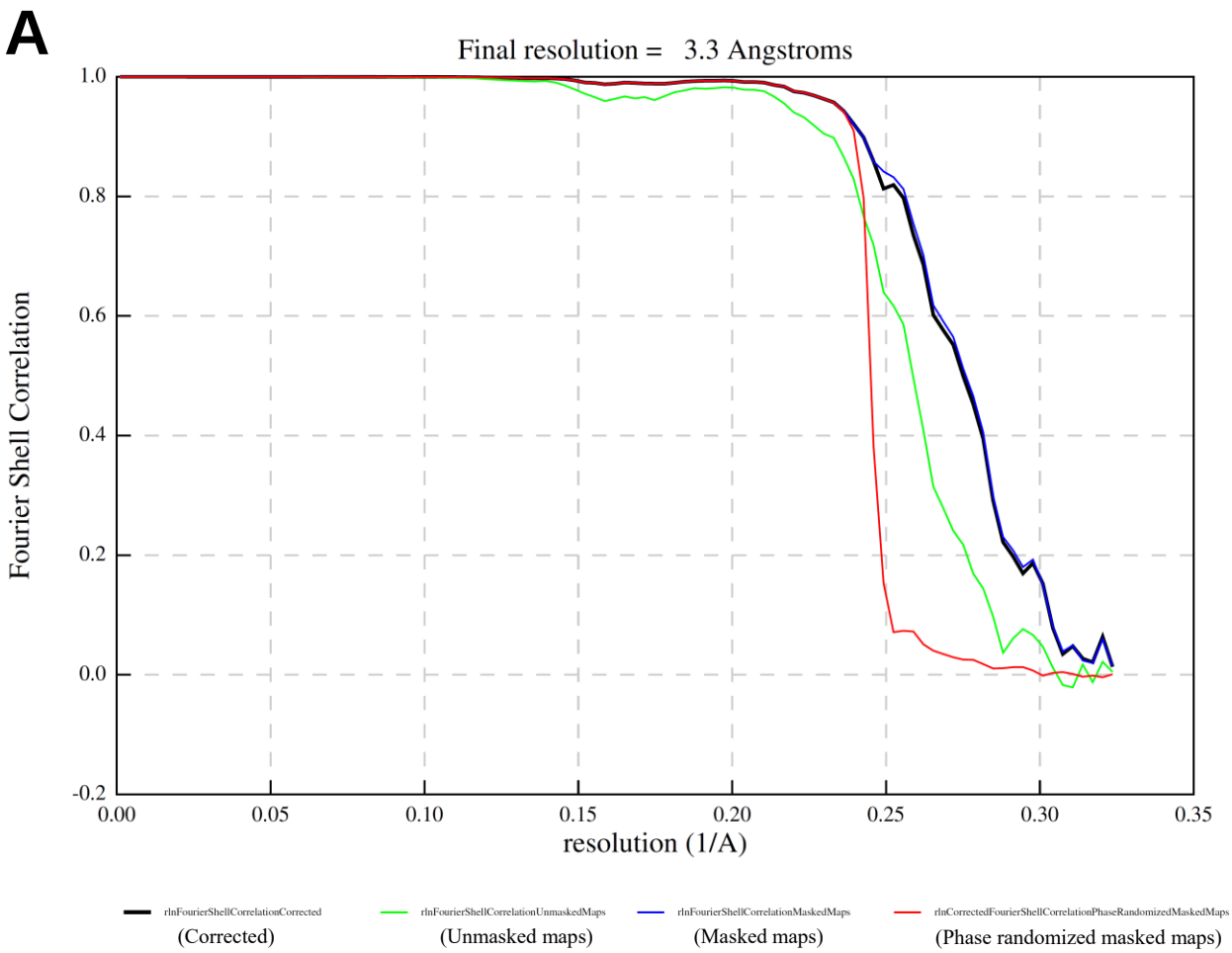


Fig. S2. Single-particle cryo-EM analysis of TIP60.
(*A*) Image-processing procedure of cryo-EM analysis of TIP60.
(*B*) Representative motion-corrected cryo-electron micrograph of TIP60.
(*C*) 2D class averages of TIP60 particles. The class averages selected for 3D classification are shown in red frames.
(*D*) 3D class averages (1) of TIP60 particles. The class averages selected for 3D classification (2) are shown in red frames.
(*E*) 3D class averages (2) of TIP60 particles. The class averages selected for 3D reconstruction are shown in red frames.



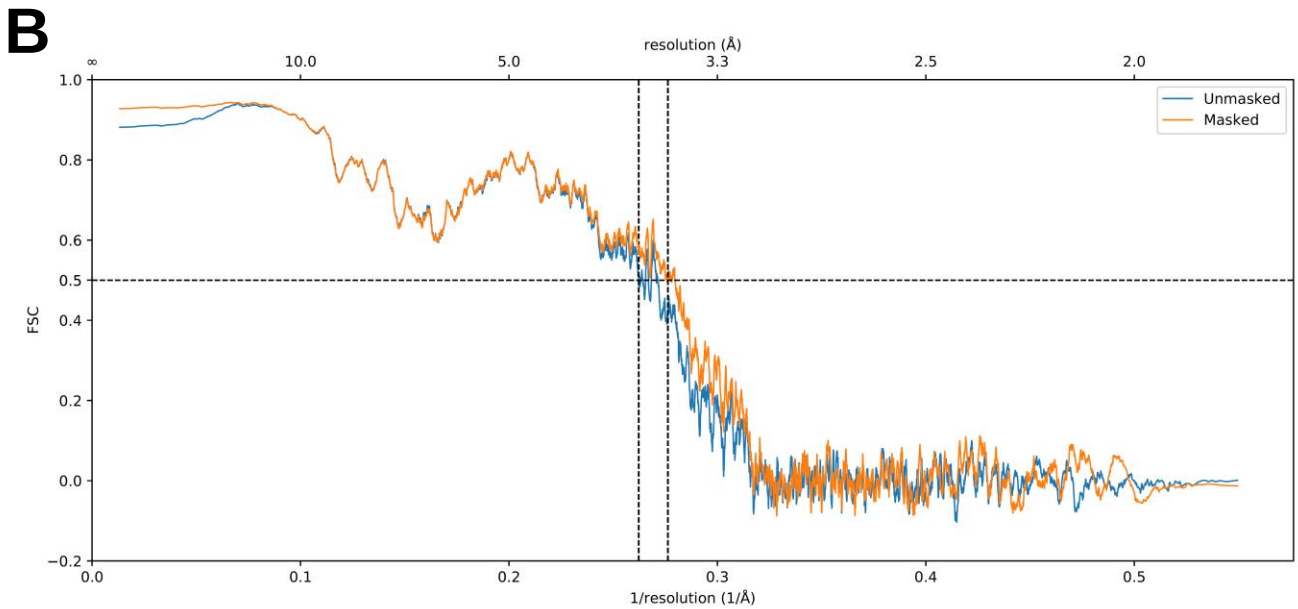


Fig. S3. Fourier shell correlation (FSC) curves.
(*A*) FSC curves of unmasked and masked reconstructions.
(*B*) Model-to-map FSC curves.

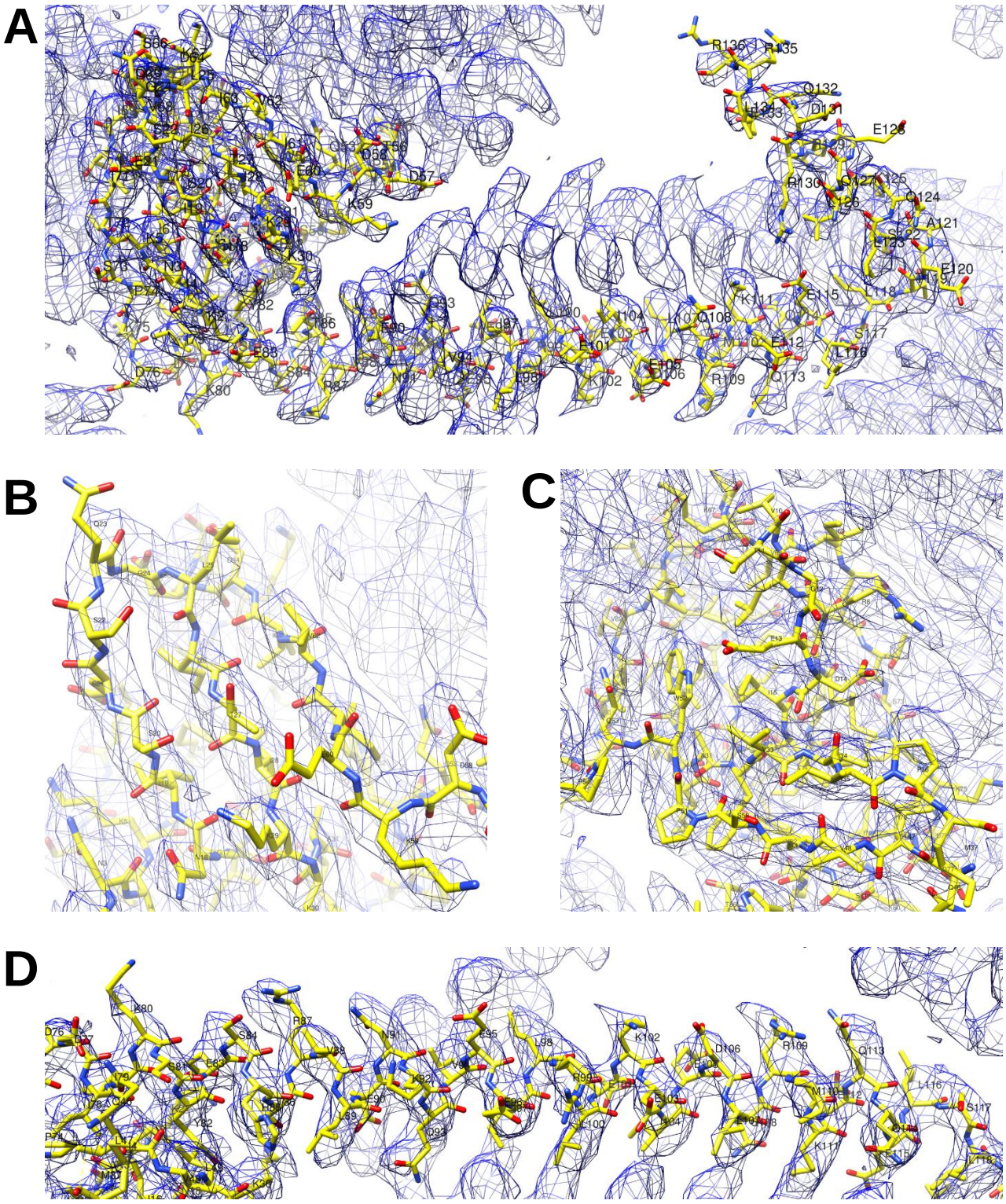


Fig. S4. Refined models of TIP60 fit into the sharpened EM map.
(*A*) A view of a TIP60 protomer (chain A). (*B*) A beta sheet region of the Sm-like domain. (*C*) A core region of the Sm-like domain. (*D*) A long α-helix structure of the C-terminal region of the Sm-like domain, a linker region, and a part of the MyoX-coil domain.

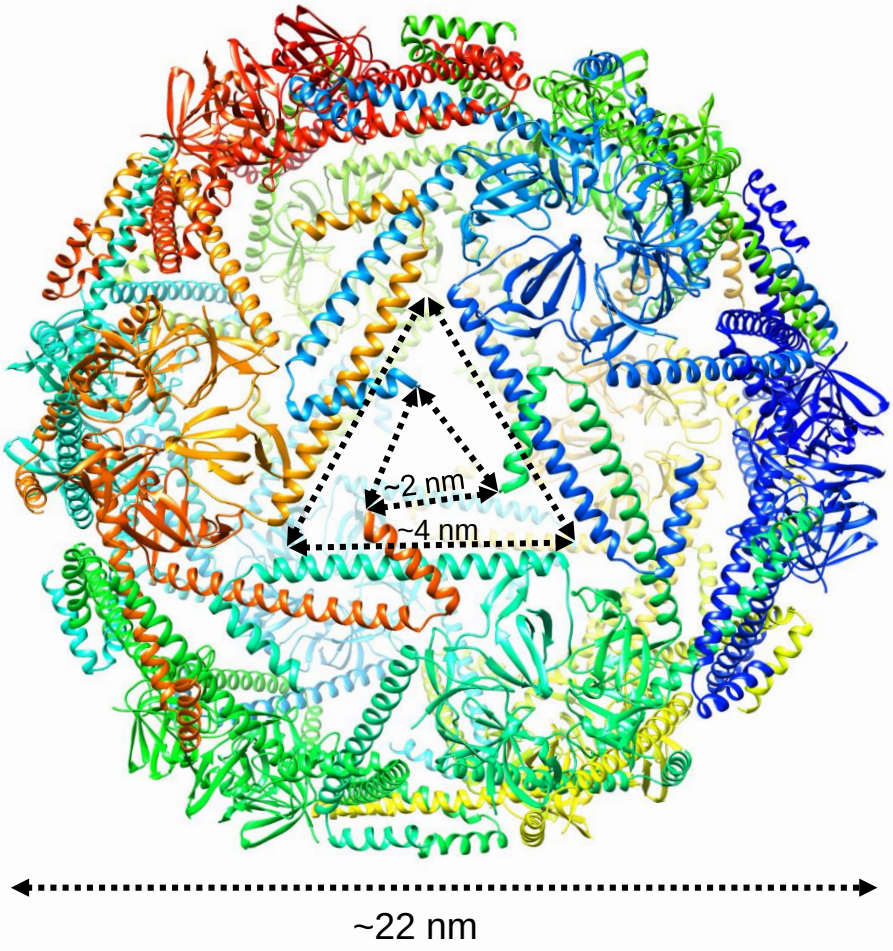


Fig. S5. Regular-triangle-like pores composed of long α-helical edges on the surface of a TIP60 nanoparticle.
The length of each edge of a long α-helix is ~4 nm and the distance between the visible C-terminal residues (R136) of the TIP60 structure is ~2 nm. The overall structure of TIP60 is a 60-meric icosahedral hollow sphere with twenty regular-triangle-like pores (outer diameter: ~22 nm; inner diameter: ~15 nm).

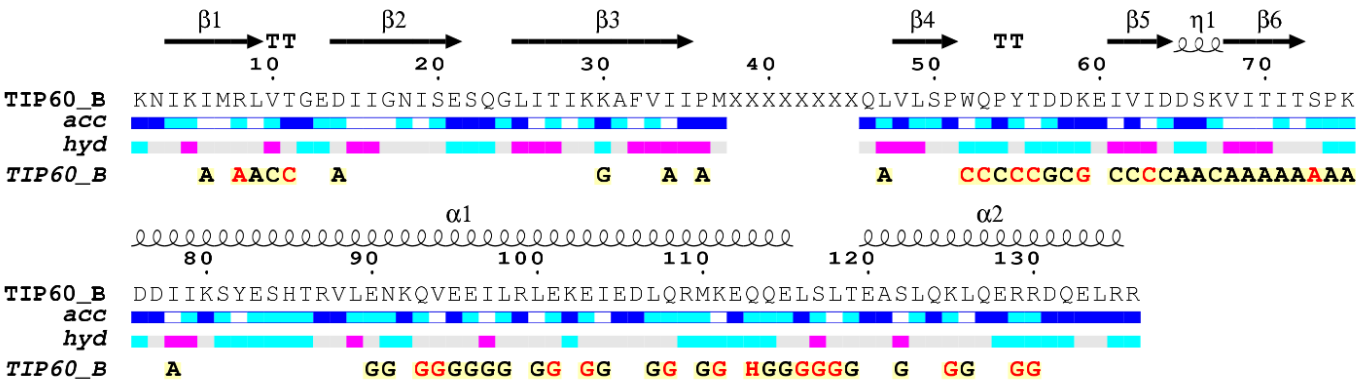


Fig. S6. Amino acid sequence of TIP60 with secondary structure elements presented on top (alpha-helices with squiggles, beta-strands with arrows and turns with TT letters) depicted using ENDscript 2.[14]
Solvent accessibility is rendered by a first bar below the sequence (blue is accessible, cyan is intermediate, and white is buried) and hydropathy by a second bar below (pink is hydrophobic, white is neutral, and cyan is hydrophilic). Bottom letters and symbols depict protein subunit–subunit contacts from a TIP60 protomer (subunit B). A red letter shows a contact < 3.2 Å, and a black letter shows a contact between 3.2 and 5 Å. “X” represents unstructured region.

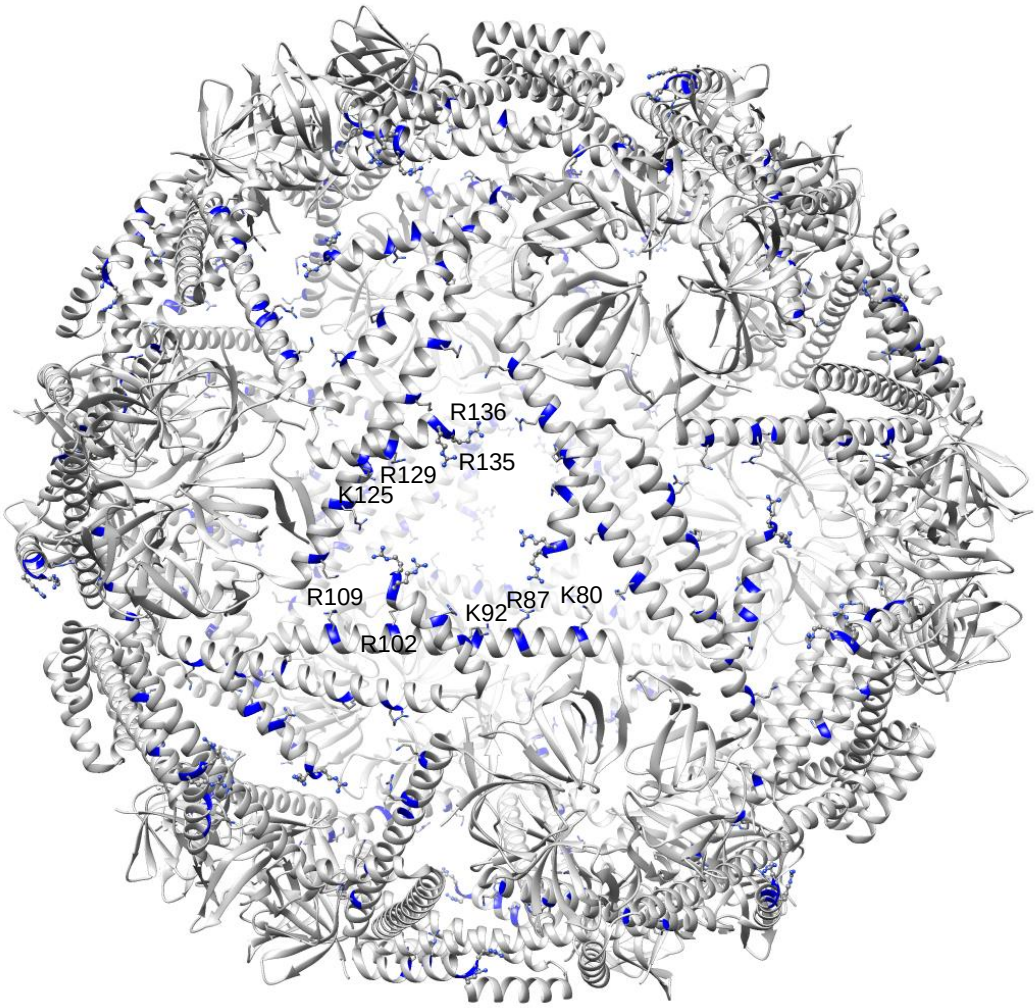


Fig. S7. Basic residues (K80, R87, K92, K102, R109, K125, R129, R135, and R136) near the entrance of the pores of TIP60. The basic residues are shown as ball and stick models and blue ribbon representation.

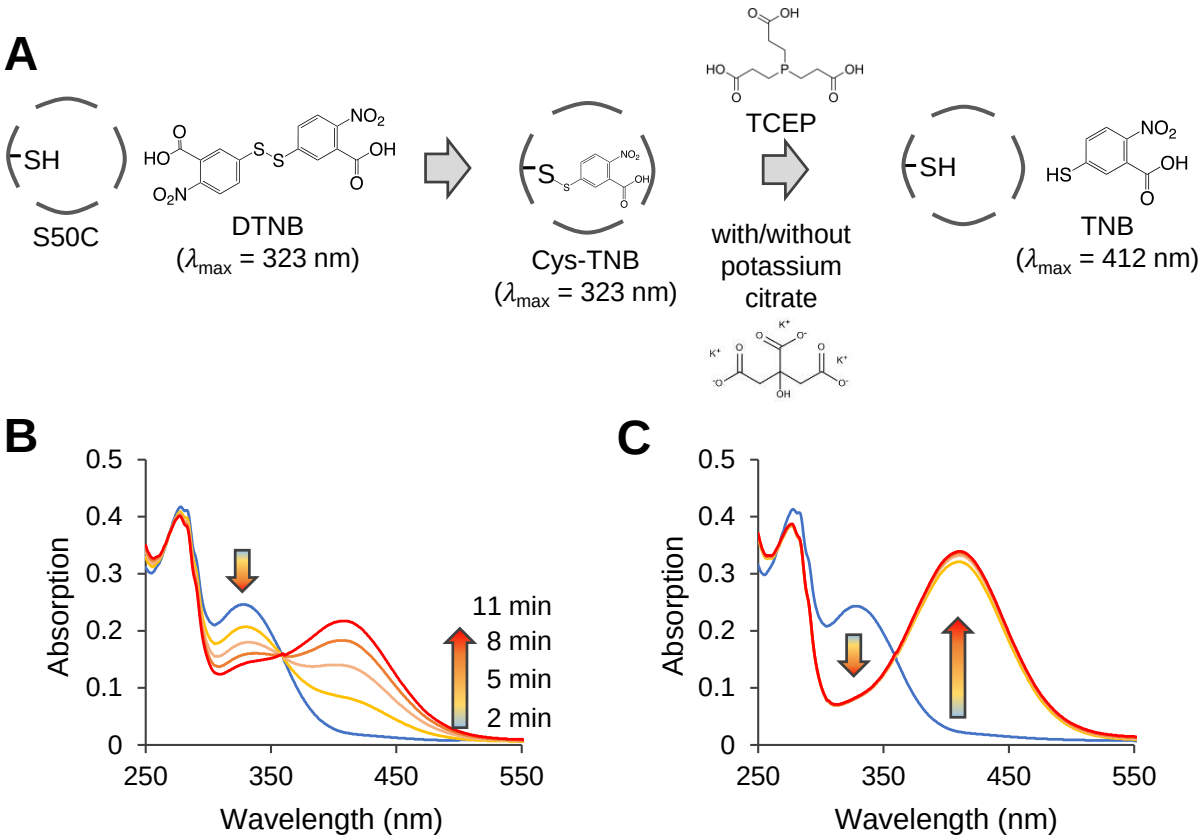


Fig. S8. (*A*) Scheme of modification of the S50C mutant of TIP60 by DTNB and release of TNB.
The reaction of DTNB modified S50C and TCEP without potassium citrate (*B*) or with 0.1 M potassium citrate (*C*) . Blue lines show absorption of the samples before the reaction with TCEP. Yellow, light orange, orange and, red lines represent the samples 2, 5, 8, 11 minutes after the addition of TCEP, respectively.

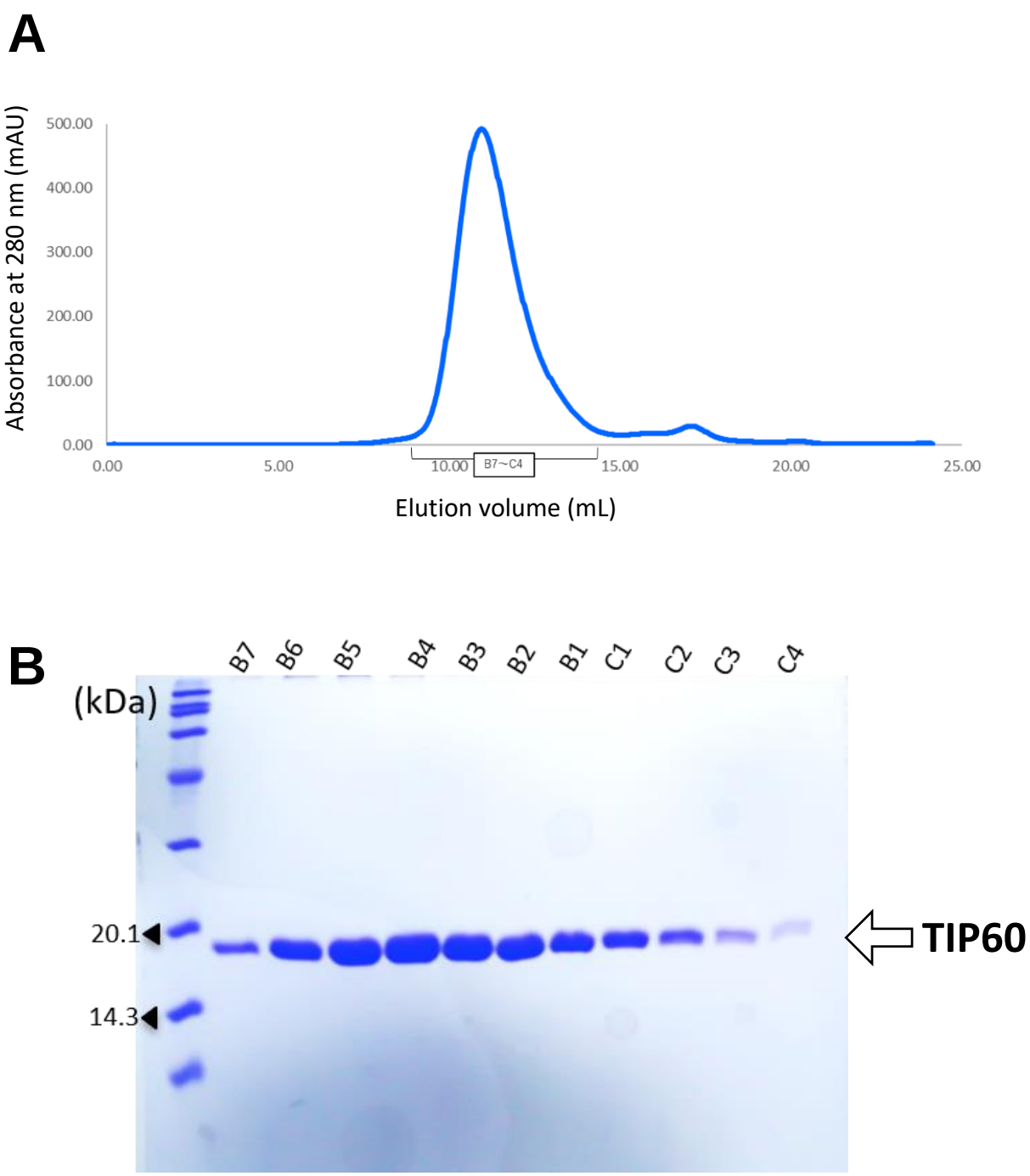


Fig. S9. Final purification step of the TIP60 protein by size exclusion chromatography (SEC).
(*A*) SEC chromatogram of TIP60 performed using Superose 6 10/300 GL and ÄKTAexplorer 10S (GE healthcare).
(*B*) Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis of the eluted fractions of the SEC experiment. Proteins in the gels were stained with Coomassie brilliant blue R-250.

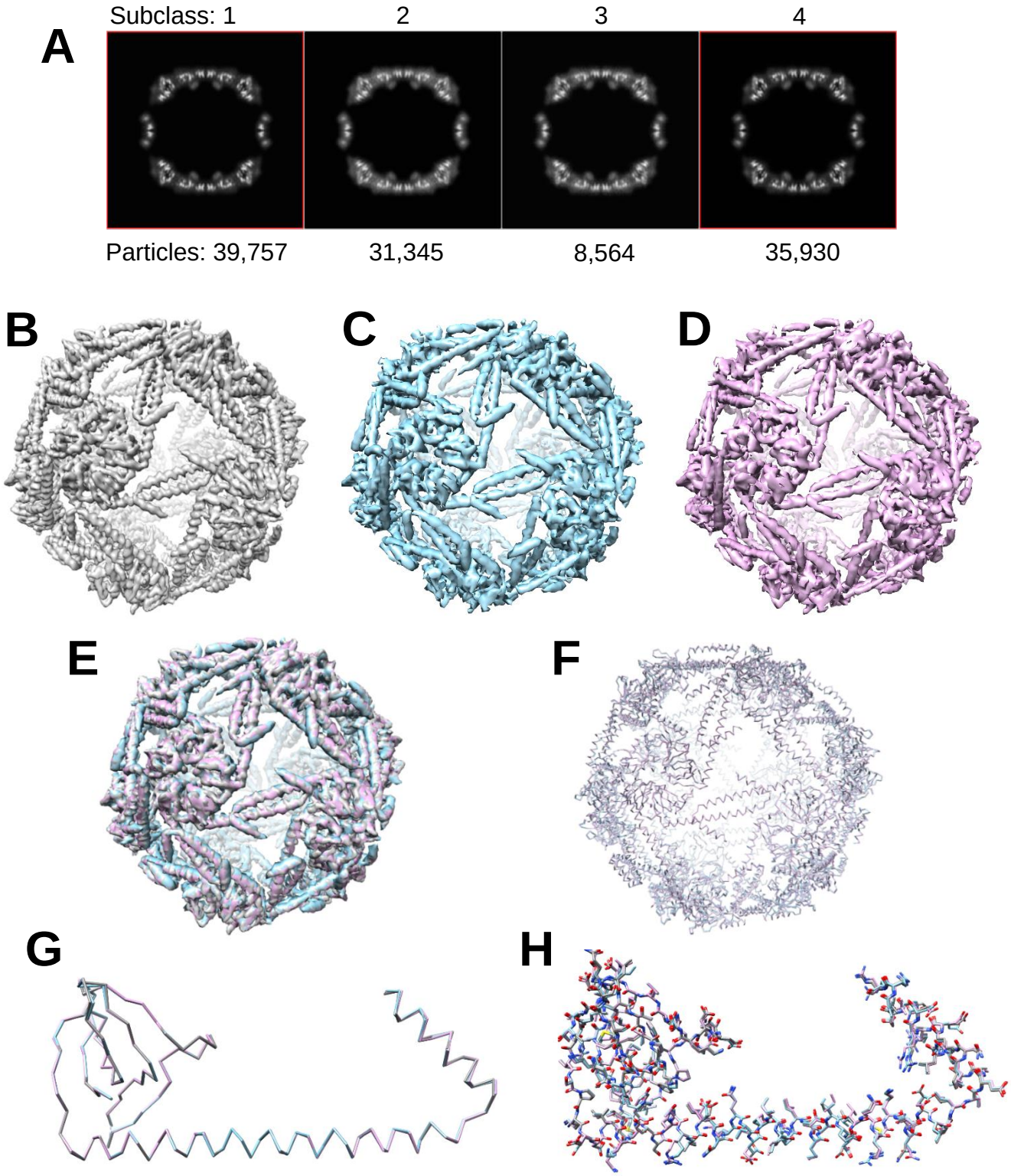
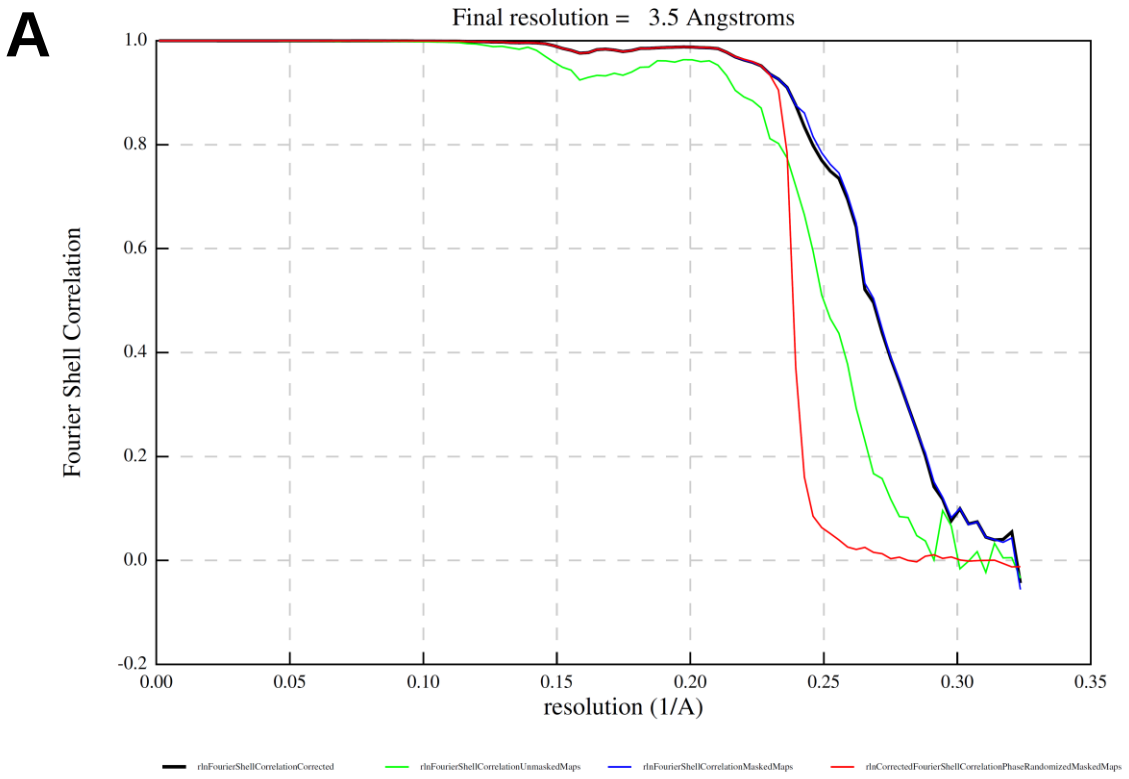


Fig. S10. Cryo-EM structure analyses of TIP60 in subclasses.
(*A*) 3D subclass averages of TIP60 particles.
(*B*) EM density map (3.3 Å resolution) reconstructed from whole class particles (115,596).
(*C*) EM density map (3.5 Å resolution) reconstructed from subclass 1 particles (39,757).
(*D*) EM density map (3.5 Å resolution) reconstructed from subclass 4 particles (35,930).
(*E*) Superposition of the EM density maps of B (gray), C (cyan), and D (magenta).
(*F*) Superposition of TIP60 backbone structures refined using each map, B (gray), C (cyan), or D (magenta). (RMSD of C_α atoms: 0.375 Å for the subclass 1 and the whole class, 0.516 Å for the subclass 4 and the whole class.)
(*G*, *H*) Superposition of backbone (*G*) and stick (*H*) model structures of a protomer of TIP60 refined using each map, B (gray), C (cyan), or D (magenta). (RMSD of C_α atoms in a protomer (chain A): 0.236 Å for the subclass 1 and the whole class, 0.246 Å for the subclass 4 and the whole class)



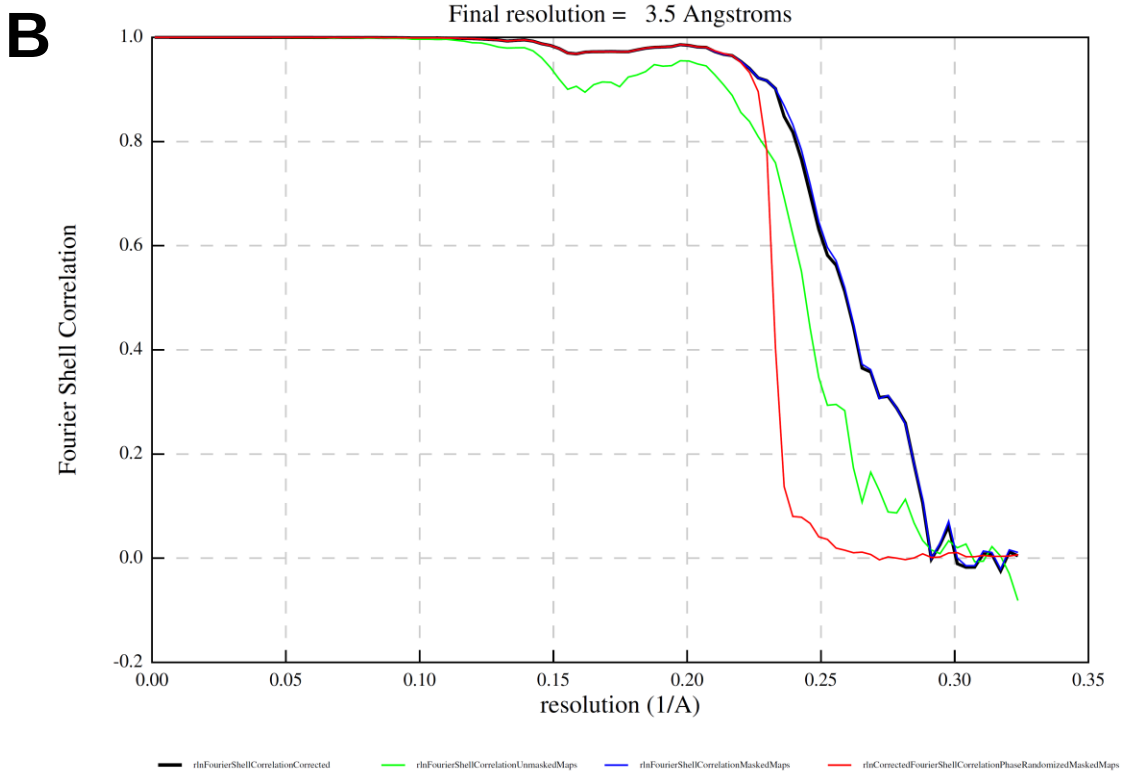


Fig. S11. Fourier shell correlation (FSC) curves of TIP60 subclasses.
(*A*) FSC curves of unmasked and masked reconstructions for subclass 1.
(*B*) FSC curves of unmasked and masked reconstructions for subclass 4.