**Cell Reports, Volume *41***

## Supplemental information

# Toward structural-omics

# of the bovine retinal pigment epithelium

**Christopher E. Morgan, Zhemin Zhang, Masaru Miyagi, Marcin Golczak, and Edward W. Yu**

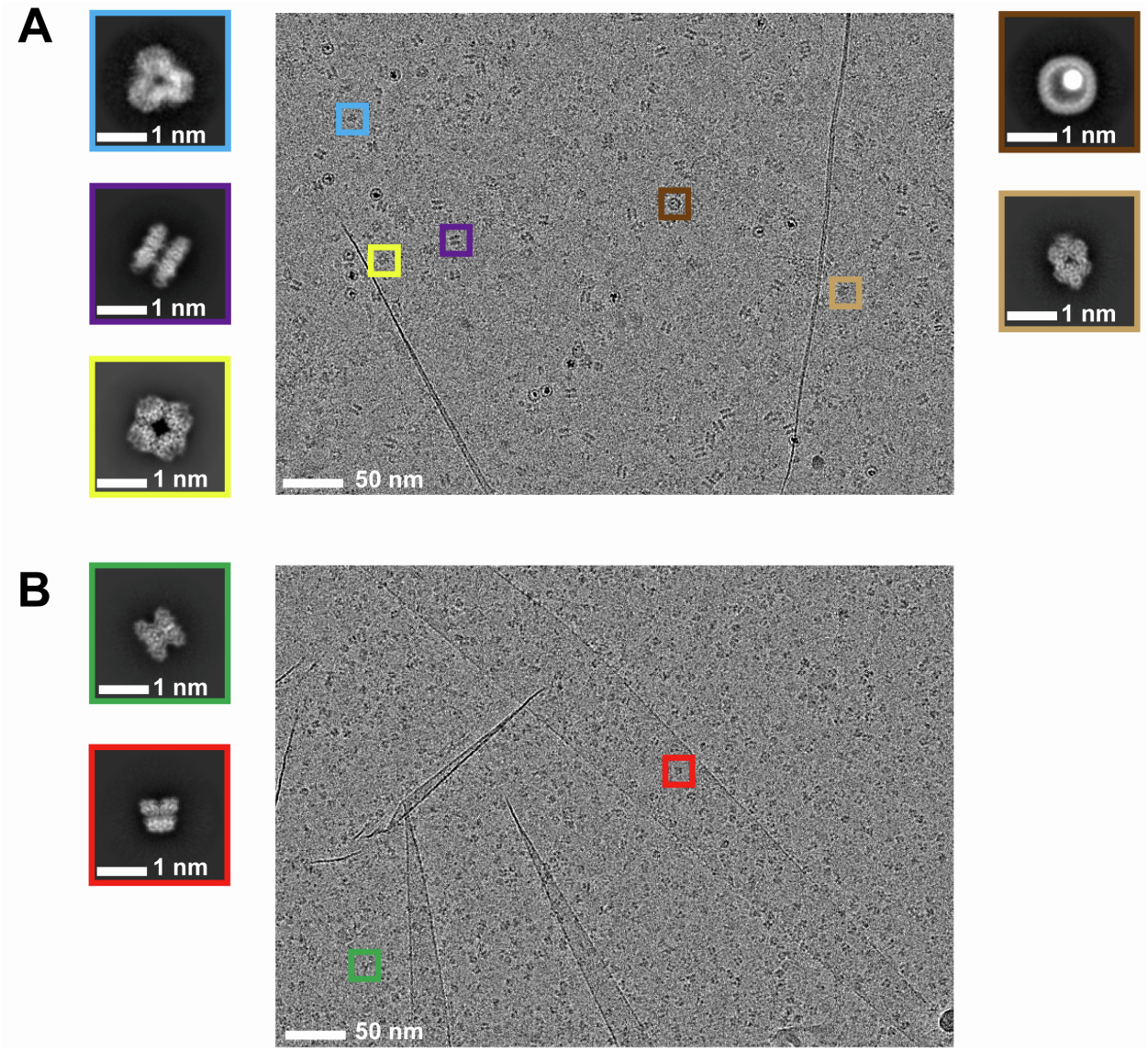


Figure S1. Micrographs of bovine RPE. Related to Figures 1-7. (A) Micrograph of the first peak of the RPE sample (300-650 kDa). The colored squares highlight single-particle images of different enzymes (cyan, DNPEP; purple, GS; yellow, MtCK; brown, FT; light brown, DYPSL2). (B) Micrograph of the second peak of the RPE sample (100-250 kDa). The colored squares highlight single-particle images of different enzymes (red, GAPDH; green, FPA). Two different cryo-EM grids were made for each peak of proteins for cryo-EM data collection (see Table S2).

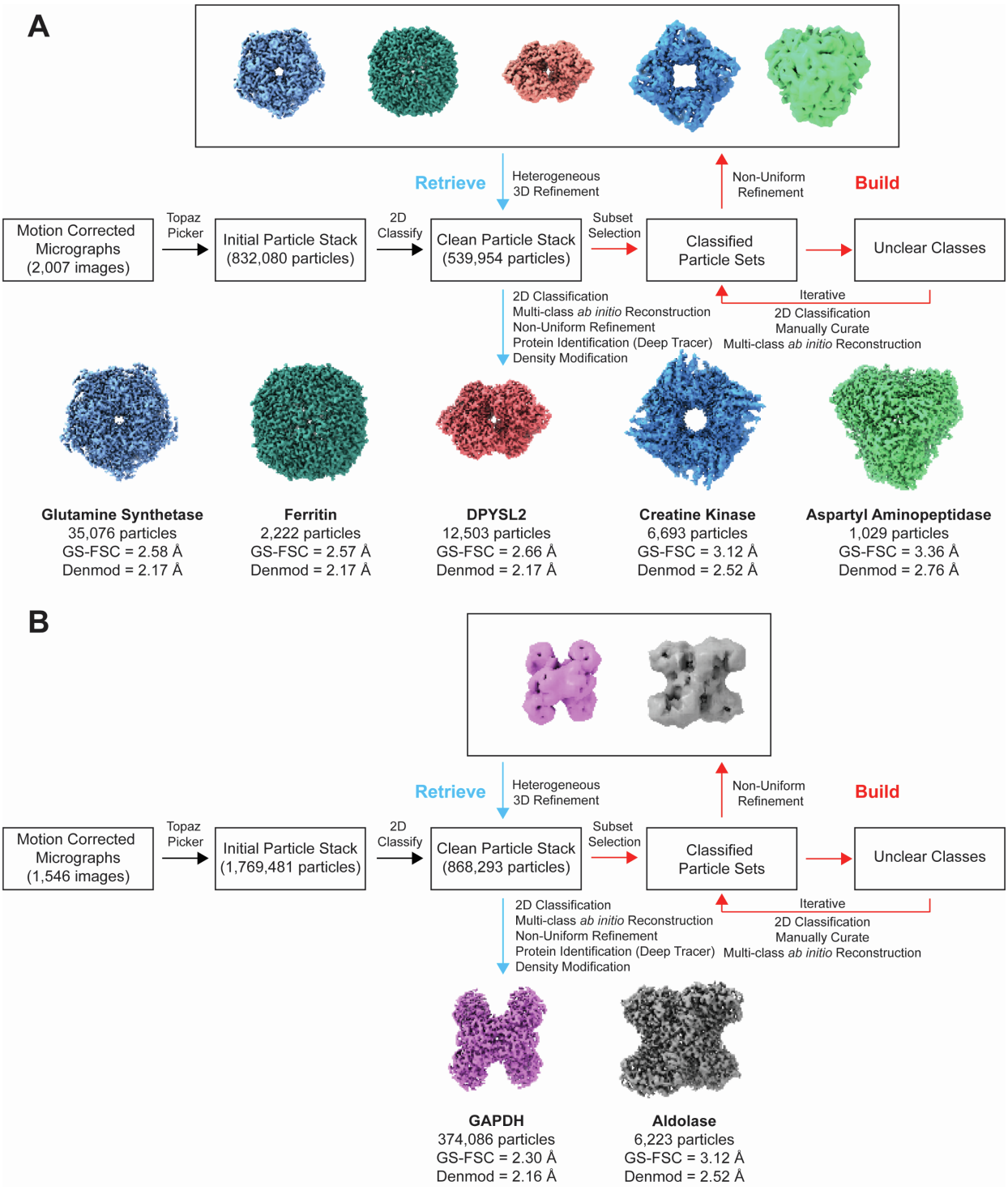


Figure S2. Build-and-Retrieve workflow. Related to Figures 1-7. (A) Workflow for RPE proteins in 300-650 kDa range. Cryo-EM processing begins as standard workflow, as motion-corrected micrographs are picked, particles undergo 2D classification and initial models are iteratively built. These low-resolution initial models are then used to retrieve particles from the cleaned dataset, resulting in 5 high-resolution maps from the 300-650 kDa RPE proteins: GS, FT, DPYSL2, MtCK and DNPEP. (B) Similar workflow for the 100-250 kDa RPE proteins resulted in two high-resolution maps: GAPDH and FPA.

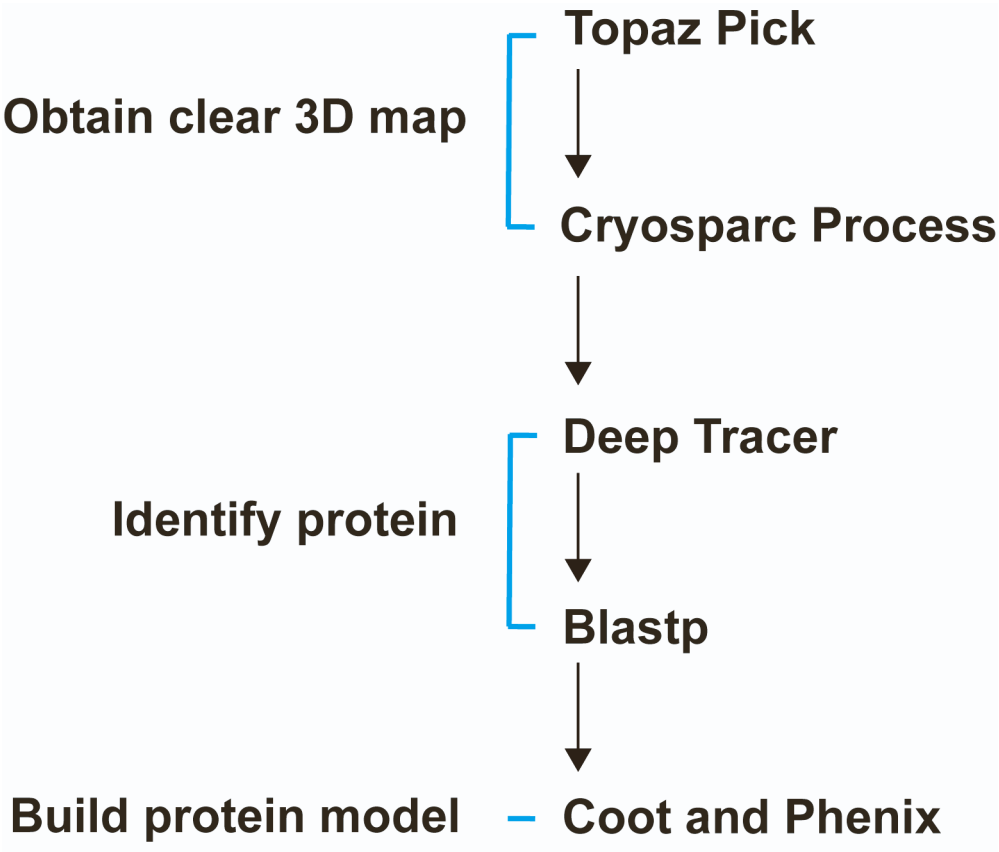


Figure S3. Schematic overview of the workflow of the BaR protocol. Related to Figures 1-7. The software used in each step of the BaR procedure are included.

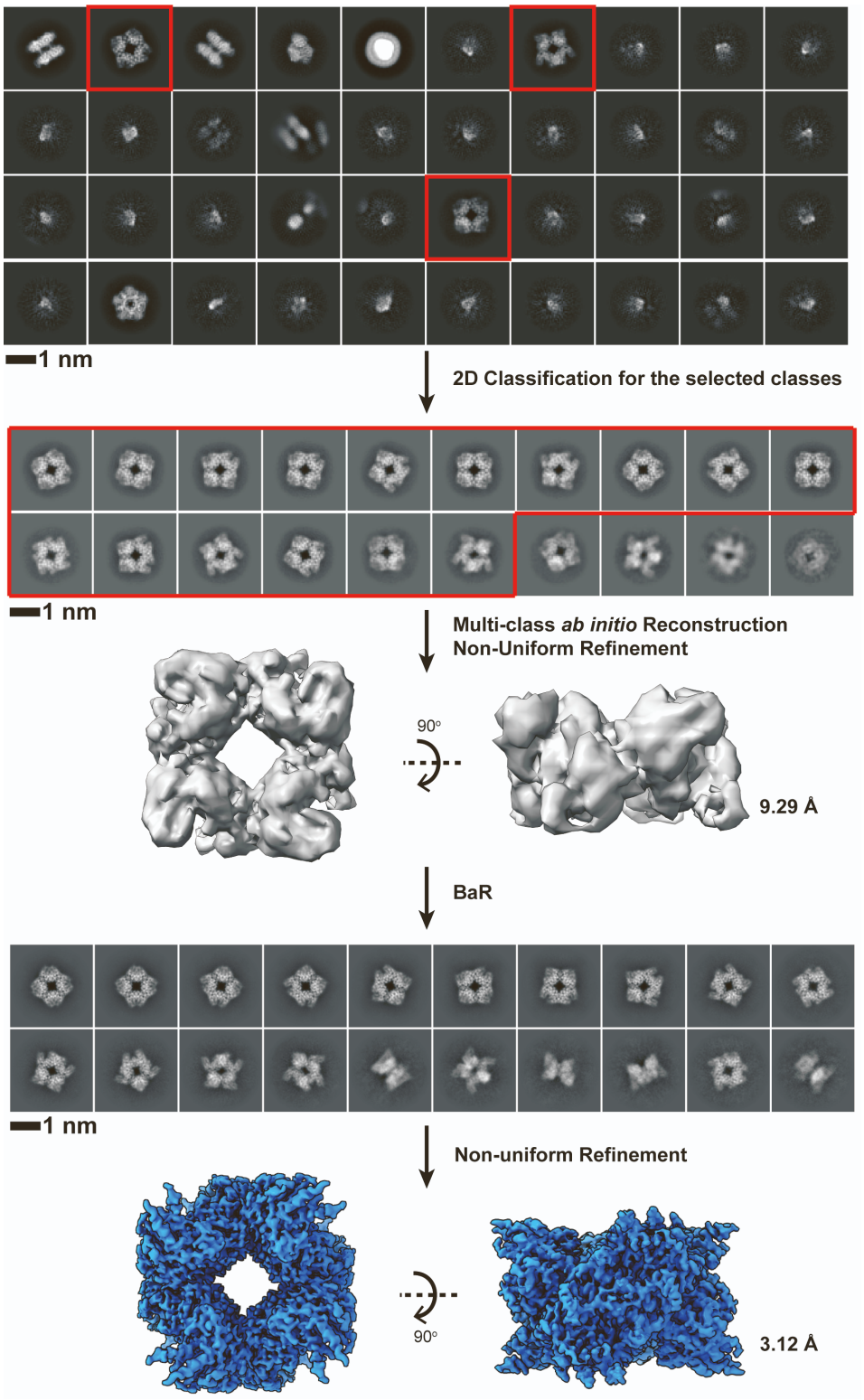


Figure S4. Cryo-EM maps of uMtCK before and after BaR. Related to Figures 1-7. This figure indicates that the BaR methodology can significantly improve the quality of the cryo-EM map for protein identification and structural determination. Two different cryo-EM grids were made for each peak of proteins for cryo-EM data collection (see Table S2).

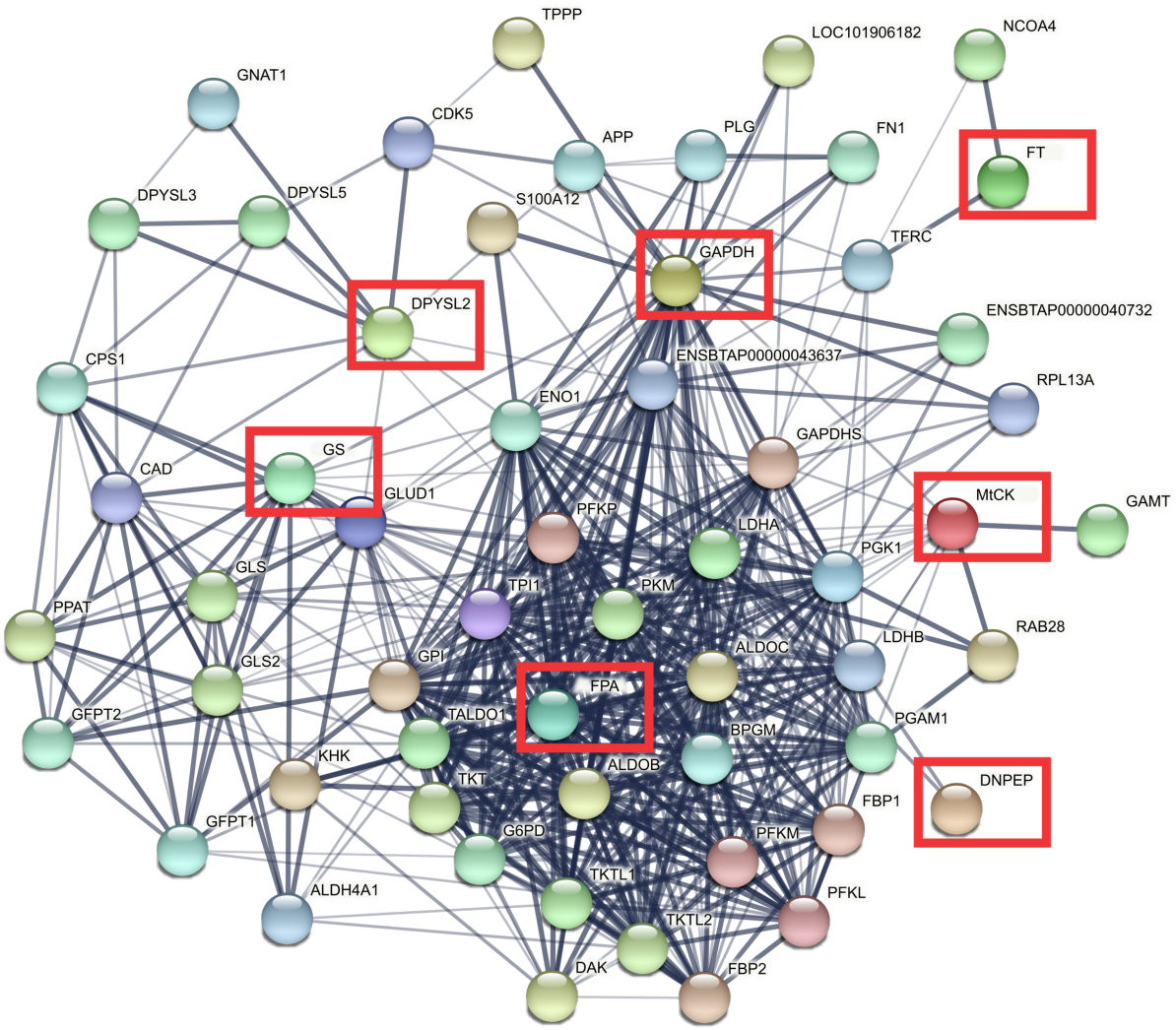


Figure S5. Interaction network of RPE. Related to Figures 1-7. This interaction network is created using the STRING database. Line thickness depicts interaction confidence and view was expanded to show 50 interactions. Results show all proteins identified from BaR interact through a complex network of proteins. The seven enzymes, GAPDH, FT, DNPEP, GS, DPYSL2, FPA and MtCK, are highlighted by red squares.

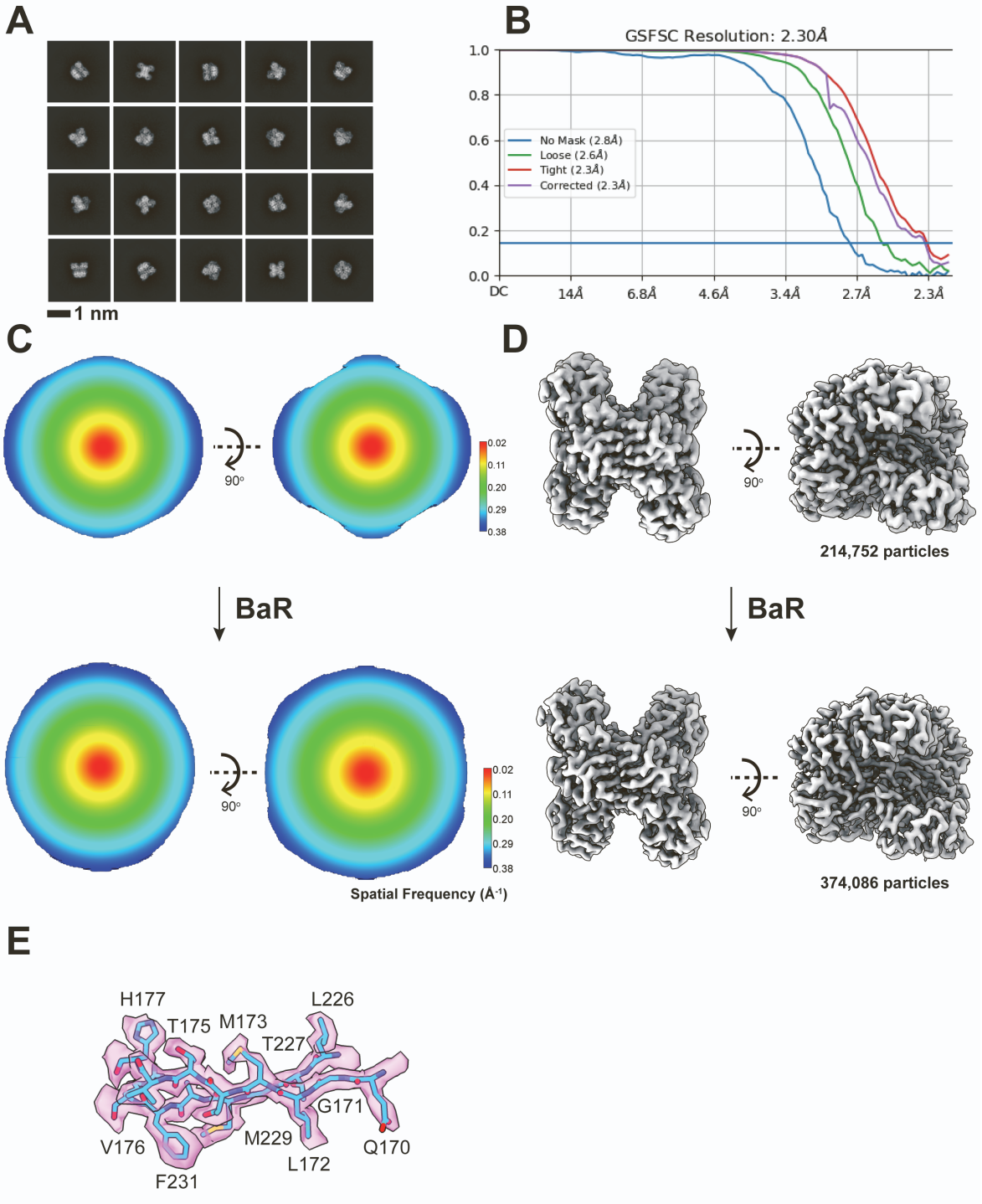


Figure S6. Cryo-EM structural determination of bovine GAPDH. Related to Figure 1. (A) Representative 2D classes. (B) GS-FSC curve showing final resolution of 2.30 Å. (C) Angular distribution calculated in 3D FSC for particle projection before and after BaR. The 3D FSCs are displayed at a threshold of 0.143, colored according to the spatial frequency. (D) Cryo-EM maps of bovine GAPDH before and after BaR. Particles contribute to the initial and final maps were labeled. (E) Representative fitting of local structure into cryo-EM density. Two different cryo-EM grids were made for each peak of proteins for cryo-EM data collection (see Table S2).

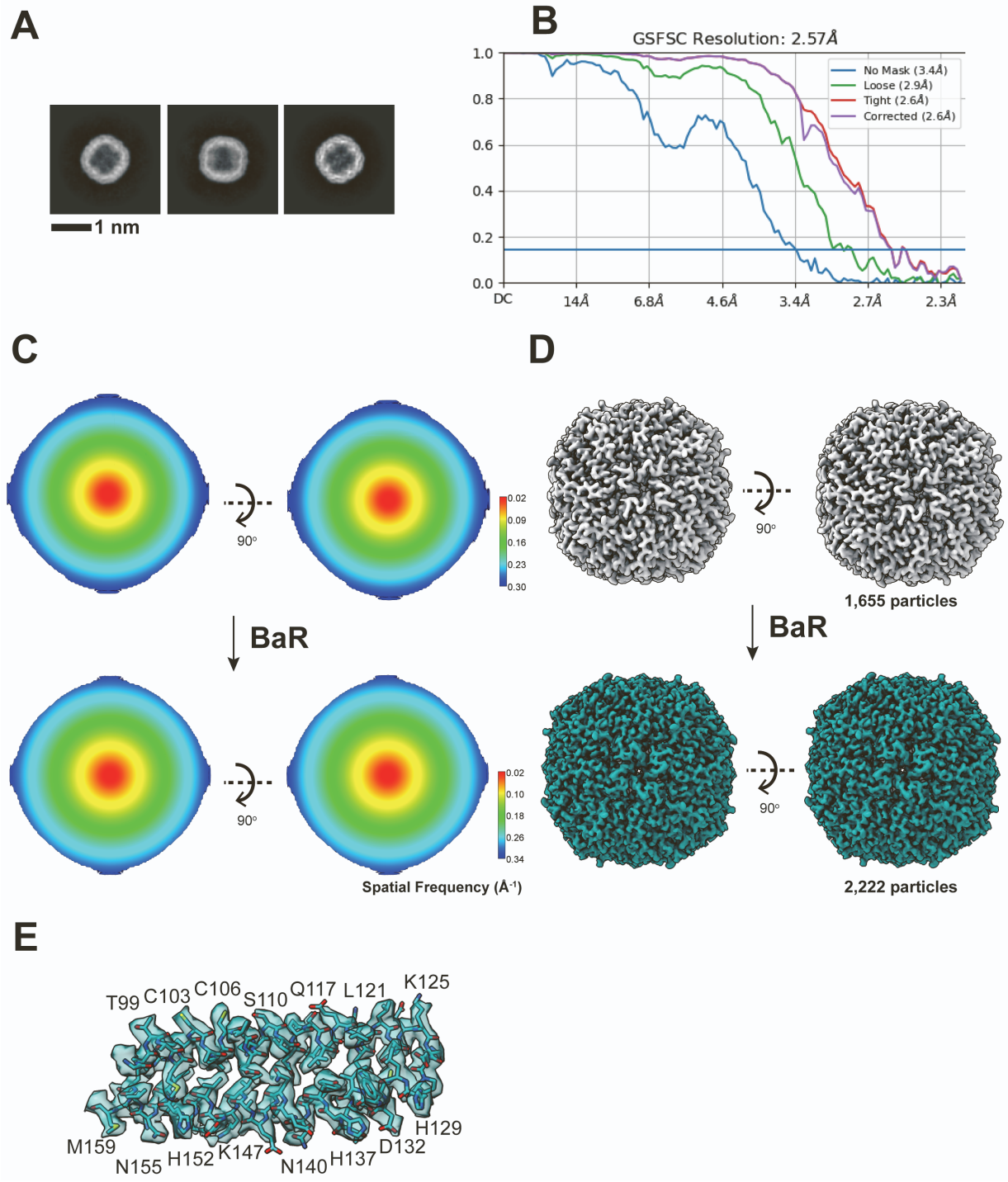


Figure S7. Cryo-EM structural determination of bovine FT. Related to Figure 2. (A) Representative 2D classes. (B) GS-FSC curve showing final resolution of 2.57 Å. (C) Angular distribution calculated in 3D FSC for particle projection before and after BaR. The 3D FSCs are displayed at a threshold of 0.143, colored according to the spatial frequency. (D) Cryo-EM maps of bovine FT before and after BaR. Particles contribute to the initial and final maps were labeled. (E) Representative fitting of local structure into cryo-EM density. Two different cryo-EM grids were made for each peak of proteins for cryo-EM data collection (see Table S2).

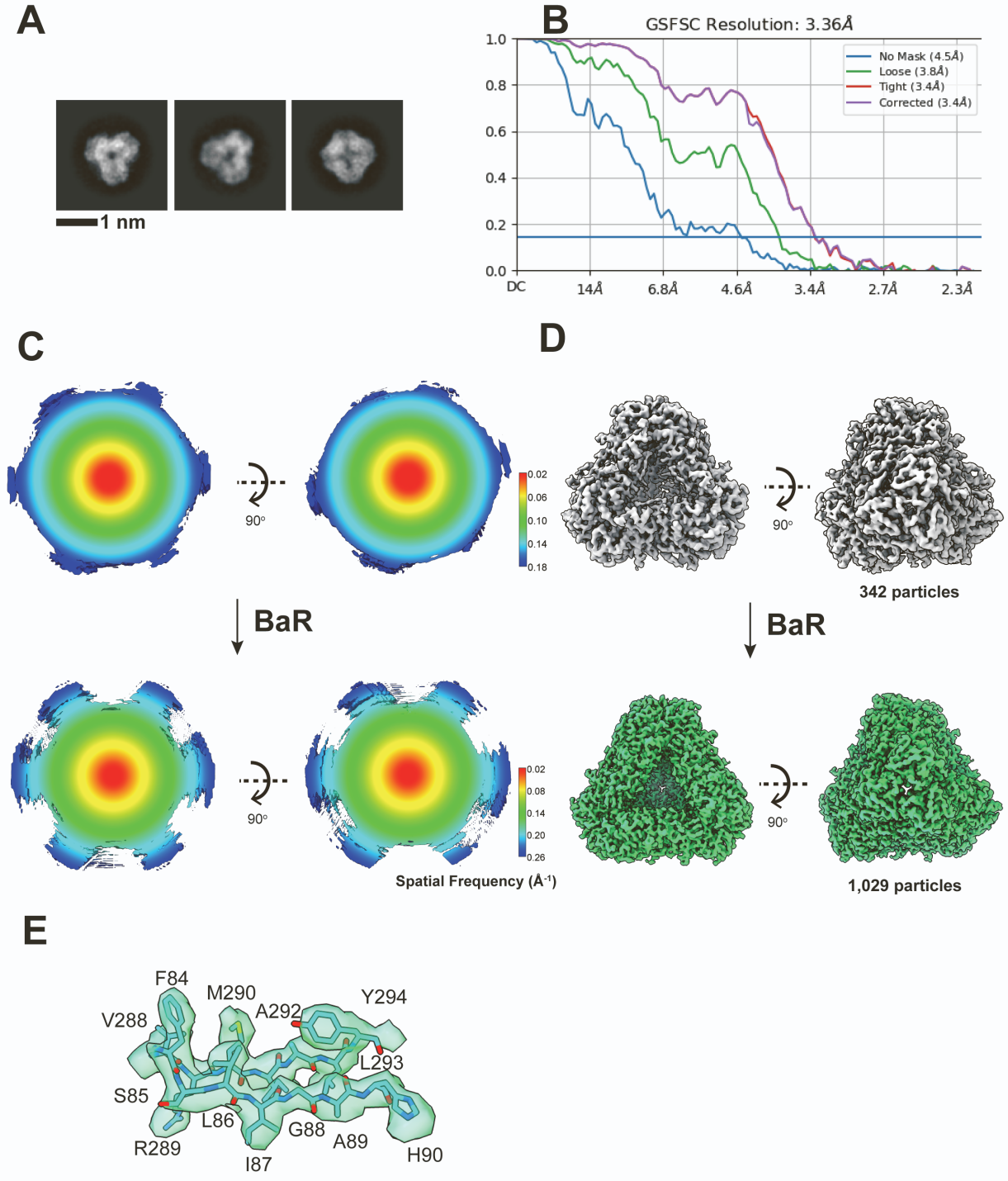


Figure S8. Cryo-EM structural determination of bovine DNPEP. Related to Figure 3. (A) Representative 2D classes. (B) GS-FSC curve showing final resolution of 3.36 Å. (C) Angular distribution calculated in 3D FSC for particle projection before and after BaR. The 3D FSCs are displayed at a threshold of 0.143, colored according to the spatial frequency. (D) Cryo-EM maps of bovine DNPEP before and after BaR. Particles contribute to the initial and final maps were labeled. (E) Representative fitting of local structure into cryo-EM density. Two different cryo-EM grids were made for each peak of proteins for cryo-EM data collection (see Table S2).

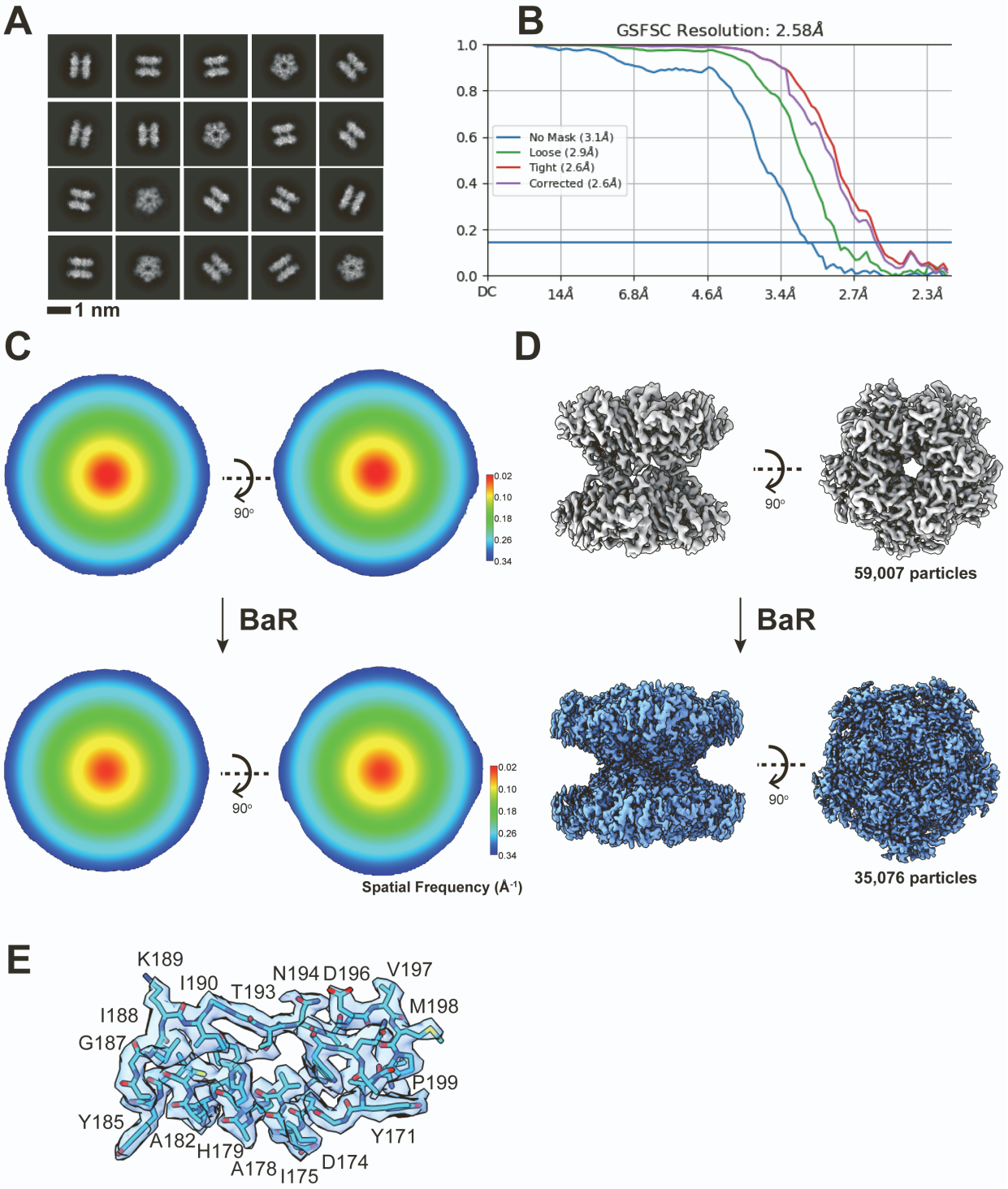


Figure S9. Cryo-EM structural determination of bovine GS. Related to Figure 4. (A) Representative 2D classes. (B) GS-FSC curve showing final resolution of 2.58 Å. (C) Angular distribution calculated in 3D FSC for particle projection before and after BaR. The 3D FSCs are displayed at a threshold of 0.143, colored according to the spatial frequency. (D) Cryo-EM maps of bovine GS before and after BaR. Particles contribute to the initial and final maps were labeled. (E) Representative fitting of local structure into cryo-EM density. Two different cryo-EM grids were made for each peak of proteins for cryo-EM data collection (see Table S2).

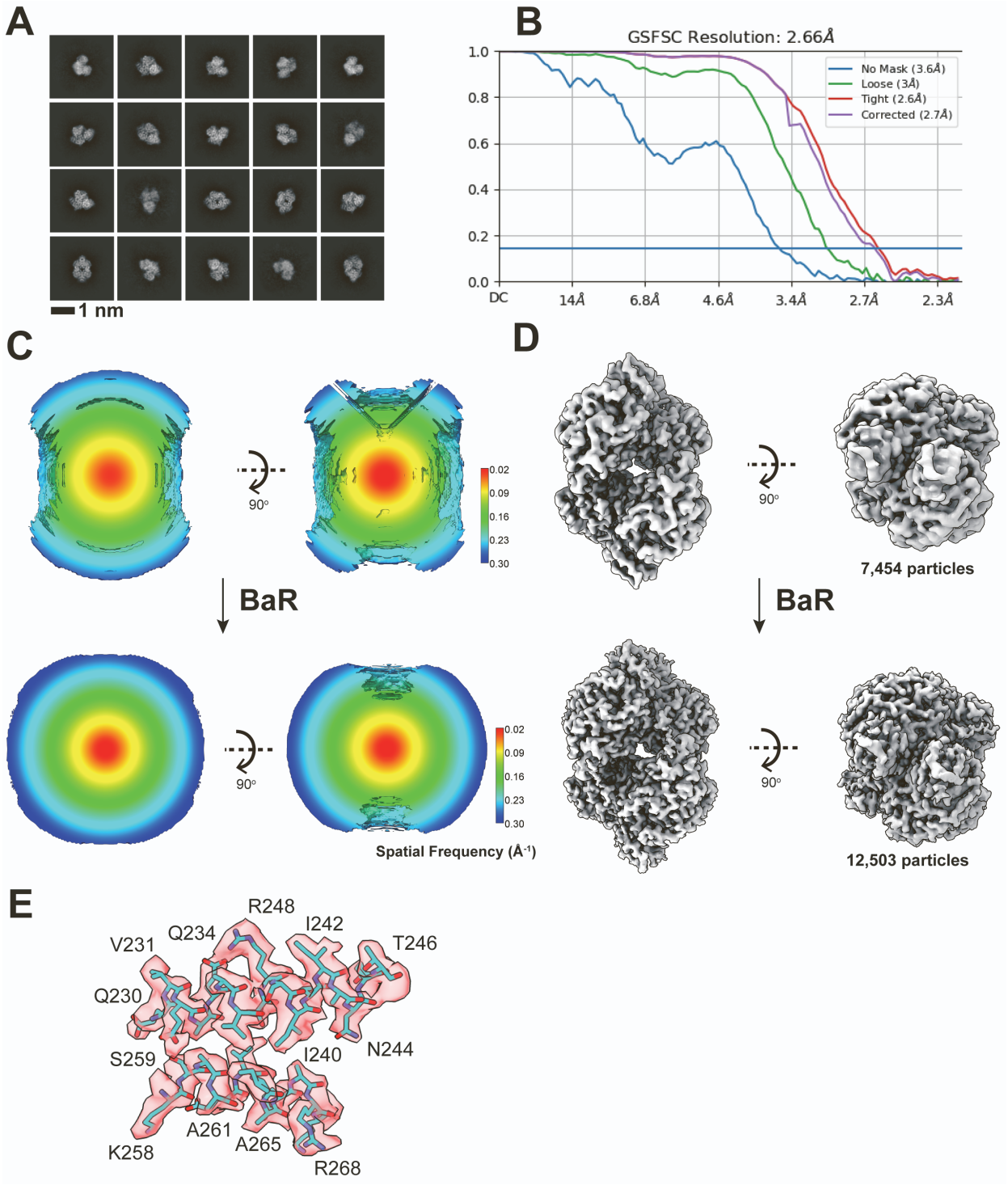


Figure S10. Cryo-EM structural determination of bovine DPYSL2. Related to Figure 5. (A) Representative 2D classes. (B) GS-FSC curve showing final resolution of 2.66 Å. (C) Angular distribution calculated in 3D FSC for particle projection before and after BaR. The 3D FSCs are displayed at a threshold of 0.143, colored according to the spatial frequency. (D) Cryo-EM maps of bovine DPYSL2 before and after BaR. Particles contribute to the initial and final maps were labeled. (E) Representative fitting of local structure into cryo-EM density. Two different cryo-EM grids were made for each peak of proteins for cryo-EM data collection (see Table S2).

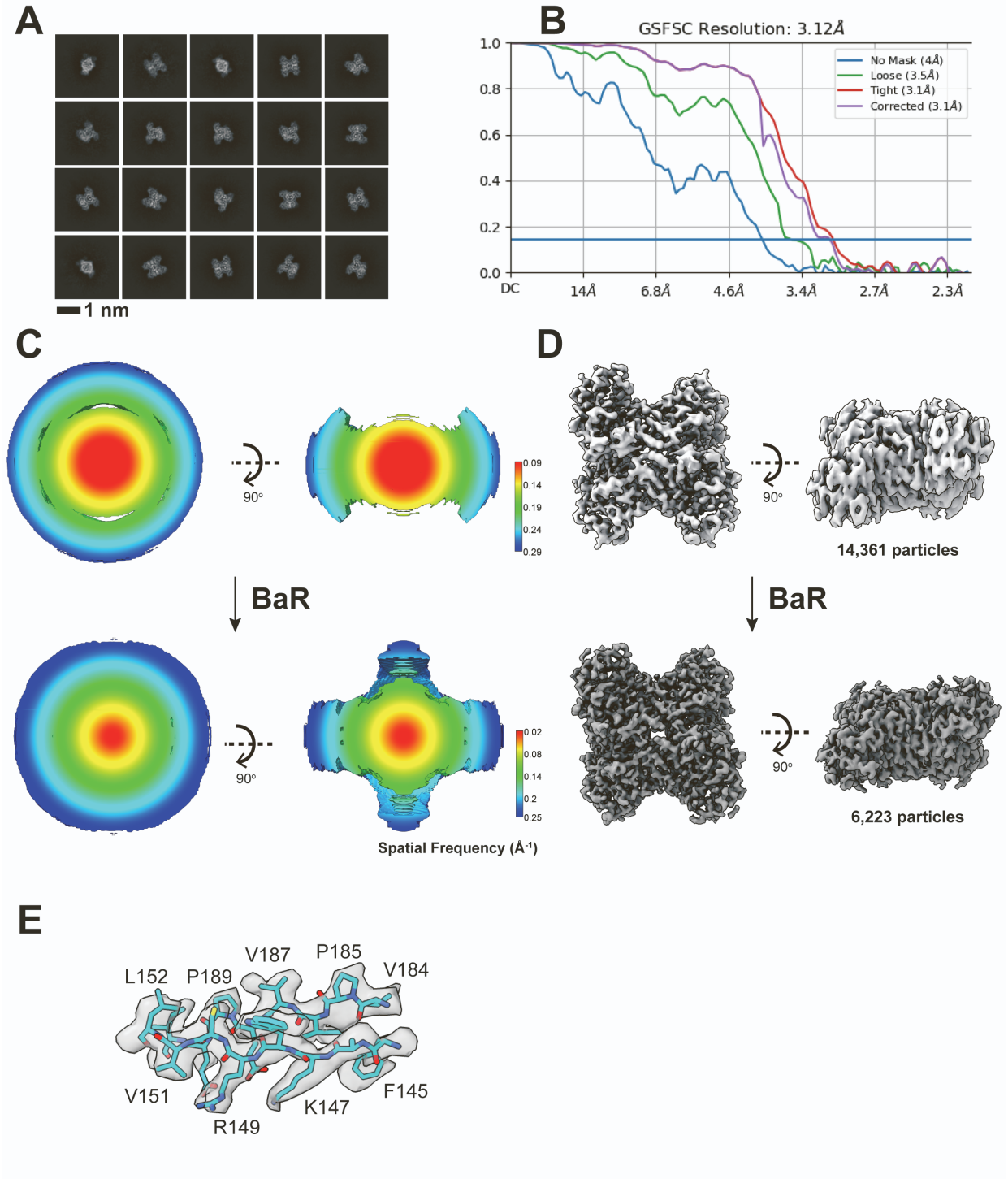


Figure S11. Cryo-EM structural determination of bovine FPA. Related to Figure 6. (A) Representative 2D classes. (B) GS-FSC curve showing final resolution of 3.12 Å. (C) Angular distribution calculated in 3D FSC for particle projection before and after BaR. The 3D FSCs are displayed at a threshold of 0.143, colored according to the spatial frequency. (D) Cryo-EM maps of bovine FPA before and after BaR. Particles contribute to the initial and final maps were labeled. (E) Representative fitting of local structure into cryo-EM density. Two different cryo-EM grids were made for each peak of proteins for cryo-EM data collection (see Table S2).

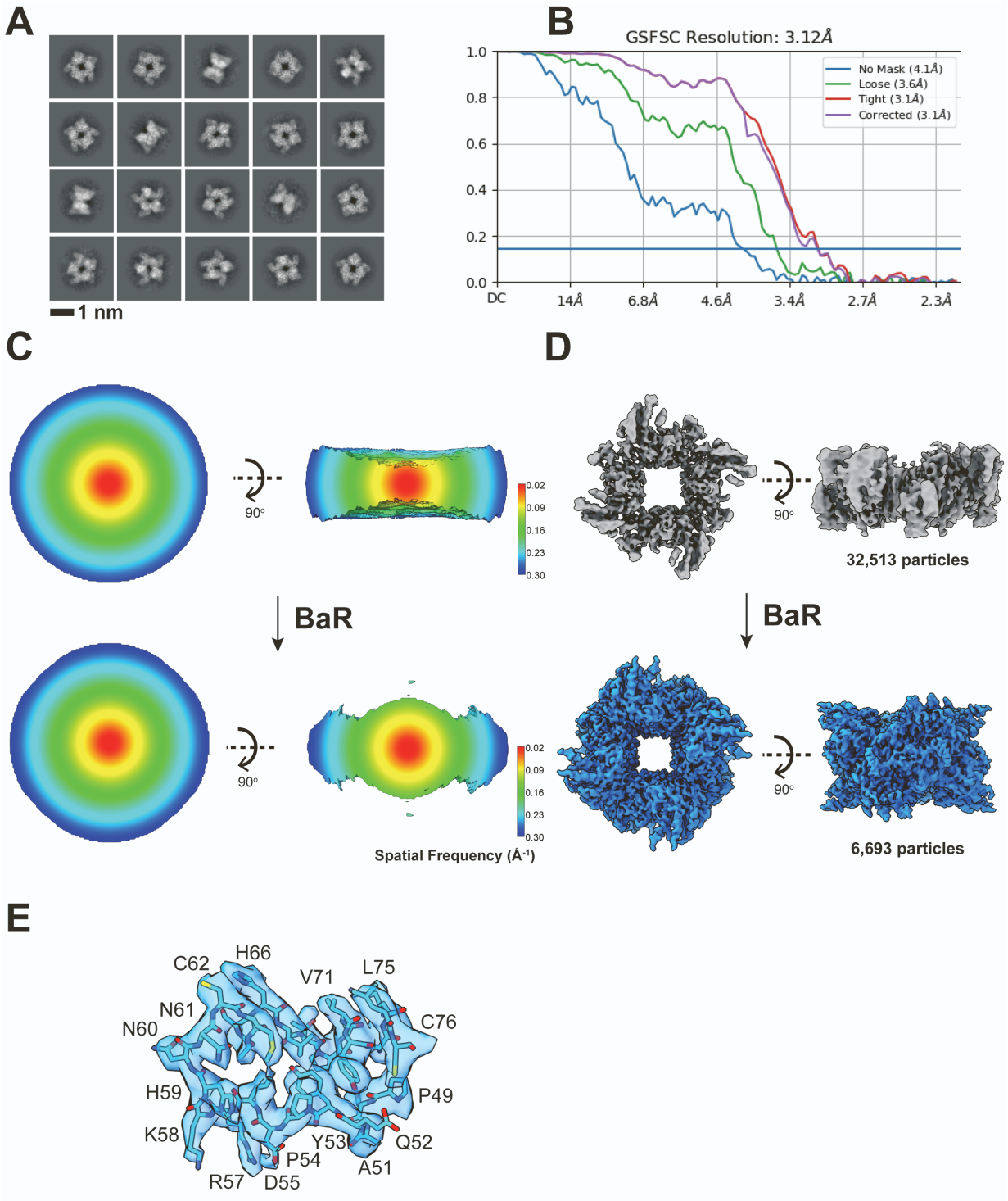


Figure S12. Cryo-EM structural determination of bovine uMtCK. Related to Figure 7. (A) Representative 2D classes. (B) GS-FSC curve showing final resolution of 3.12 Å. (C) Angular distribution calculated in 3D FSC for particle projection before and after BaR. The 3D FSCs are displayed at a threshold of 0.143, colored according to the spatial frequency. (D) Cryo-EM maps of bovine uMtCK before and after BaR. Particles contribute to the initial and final maps were labeled. (E) Representative fitting of local structure into cryo-EM density. Two different cryo-EM grids were made for each peak of proteins for cryo-EM data collection (see Table S2).

**Table S1. Proteomic analysis. Related to Figures 1-7.**

| **A. RPE proteins (Peak 1: 300-650 kDa).** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Rank** | **Log2 intensity** | **Log2 iBAQ[a]** | **Unique peptides** | **Mol. weight [kDa]** | **Score** | **MS/MS count** | **Protein IDs** |
| **1** | **34.55** | **30.64** | **31** | **42.03** | **323.31** | **231** | **GLNA_BOVIN Glutamine synthetase** |
| **2** | **33.71** | **29.42** | **28** | **46.89** | **323.31** | **217** | **KCRU_BOVIN Creatine kinase U-type, mitochondrial** |
| **3** | **32.55** | **27.68** | **35** | **62.27** | **323.31** | **123** | **DPYL2_BOVIN Dihydropyrimidinase-related protein 2** |
| 4 | 32.51 | 26.79 | 48 | 139.70 | 323.31 | 149 | RET3_BOVIN Retinol-binding protein 3 |
| 5 | 30.76 | 26.24 | 3 | 50.28 | 323.31 | 232 | TBA1D_BOVIN Tubulin alpha-1D |
| 6 | 29.42 | 26.14 | 11 | 19.79 | 242.96 | 42 | CRYAA_BOVIN Alpha-crystallin A |
| 7 | 30.98 | 25.94 | 42 | 53.72 | 323.31 | 149 | VIME_BOVIN Vimentin |
| 8 | 31.01 | 25.88 | 33 | 71.24 | 323.31 | 196 | HSP7C_BOVIN Heat shock cognate 71 kDa protein |
| 9 | 30.84 | 25.82 | 35 | 84.73 | 323.31 | 129 | HS90A_BOVIN Heat shock protein HSP 90-alpha |
| 10 | 30.00 | 25.63 | 26 | 55.42 | 323.31 | 73 | IMDH1_BOVIN Inosine-5-monophosphate dehydrogenase 1 |
| . . | | | | | | | |
| **20** | **28.56** | **24.44** | **14** | **51.82** | **323.31** | **30** | **DNPEP_BOVIN Aspartyl aminopeptidase** |
| . . | | | | | | | |
| **29** | **27.32** | **23.89** | **6** | **21.05** | **86.178** | **11** | **FRIH_BOVIN Ferritin heavy chain** |
| **B. RPE proteins (Peak 2: 100-250 kDa).** | | | | | | | |
| **Rank** | **Log2 intensity** | **Log2 iBAQ[a]** | **Unique peptides** | **Mol. weight [kDa]** | **Score** | **MS/MS count** | **Protein IDs** |
| **1** | **36.09** | **32.61** | **20** | **35.86** | **323.31** | **481** | **G3P_BOVIN Glyceraldehyde-3-phosphate dehydrogenase** |
| 2 | 31.97 | 28.54 | 25 | 42.71 | 323.31 | 133 | KCRB_BOVIN Creatine kinase B-type |
| 3 | 32.25 | 28.39 | 22 | 36.59 | 323.31 | 67 | LDHA_BOVIN L-lactate dehydrogenase A |
| 4 | 31.17 | 27.48 | 18 | 36.72 | 323.31 | 71 | LDHB_BOVIN L-lactate dehydrogenase B |
| 5 | 31.79 | 27.20 | 33 | 71.24 | 323.31 | 196 | HSP7C_BOVIN Heat shock cognate 71 kDa protein |
| 6 | 30.93 | 27.18 | 19 | 29.17 | 323.31 | 58 | 1433E_BOVIN 14-3-3 protein epsilon |
| 7 | 31.42 | 27.02 | 28 | 67.90 | 323.31 | 70 | TKT_BOVIN Transketolase |
| 8 | 28.32 | 26.99 | 5 | 8.543 | 41.343 | 14 | GBG1_BOVIN Guanine nucleotide-binding protein G(T) subunit gamma-T1 |
| 9 | 30.86 | 26.95 | 3 | 50.28 | 323.31 | 232 | TBA1D_BOVIN Tubulin alpha-1D |
| 10 | 30.24 | 26.75 | 15 | 27.74 | 323.31 | 45 | 1433Z_BOVIN 14-3-3 protein zeta/delta |
| . . | | | | | | | |
| **12** | **30.10** | **26.12** | **3** | **39.54** | **23.96** | **5** | **ALDOB_BOVIN Fructose-bisphosphate aldolase B** |

[a]The iBAQ value is obtained by dividing protein intensities by the number of theoretically observable tryptic peptides between 6 and 30 amino acids, and is on average highly correlated with protein abundance.

**Table S2. RPE cryo-EM data collection and refinement statistics. Related to Figures 1-7.**

| **Data collection** | **Peak 1 (300-650 kDa)** | | | | | **Peak 2 (100-250 kDa)** | |
|---|---|---|---|---|---|---|---|
| Magnification | 81,000 | | 81,000 | | | 81,000 | 81,000 |
| Voltage (kV) | 300 | | 300 | | | 300 | 300 |
| Electron Microscope | Krios-GIF-K3 | | Krios-GIF-K3 | | | Krios-GIF-K3 | Krios-GIF-K3 |
| Defocus (um) | -1.0 to -2.5 | | -1.0 to -2.5 | | | -1.0 to -2.5 | -1.0 to -2.5 |
| Energy filter width (eV) | 20 | | 20 | | | 20 | 20 |
| Pixel size (Å) | 1.07 (0.535) | | 1.07 (0.535) | | | 1.07 (0.535) | 1.07 (0.535) |
| Total dose ($e^-$/ Å$^2$) | 37 | | 36 | | | 36 | 36 |
| Number of frames | 35 | | 37 | | | 35 | 37 |
| Number of micrographs | 1,497 | | 510 | | | 970 | 576 |
| Initial particle images (no.) | 832,080 | | | | | 1,769,481 | |
| **Refinement** | **GS** | **FT** | **DPYSL2** | **MtCK** | **DNPEP** | **GAPDH** | **FPA** |
| Total Particles (no.) | 35,076 | 2,222 | 12,503 | 6,693 | 1,029 | 374,086 | 6,223 |
| GS-FSC Resolution (0.143, Å)[a] | 2.58 | 2.57 | 2.66 | 3.12 | 3.36 | 2.30 | 3.12 |
| Denmod Resolution (Å)[b] | 2.17 | 2.17 | 2.17 | 2.52 | 2.76 | 2.16 | 2.52 |
| Model composition | | | | | | | |
| Chains | 10 | 24 | 4 | 8 | 12 | 4 | 4 |
| Protein residues | 3,430 | 4,152 | 1,972 | 2,904 | 5,064 | 1,328 | 1,376 |
| r.m.s.d. | | | | | | | |
| Bond lengths (Å) | 0.005 | 0.002 | 0.004 | 0.002 | 0.005 | 0.004 | 0.004 |
| Bond angles (°) | 0.590 | 0.437 | 0.980 | 0.504 | 0.665 | 0.550 | 0.542 |
| **Validation** | **GS** | **FT** | **DPYSL2** | **MtCK** | **DNPEP** | **GAPDH** | **FPA** |
| MolProbity score | 1.81 | 1.21 | 1.55 | 1.73 | 2.13 | 1.62 | 1.80 |
| Clash score | 5.60 | 4.34 | 6.74 | 7.78 | 10.51 | 4.48 | 8.84 |
| Ramachandran plot | | | | | | | |
| Favored (%) | 96.11 | 98.83 | 98.17 | 99.16 | 96.89 | 96.74 | 97.66 |
| Allowed (%) | 3.89 | 1.17 | 1.83 | 0.84 | 3.11 | 3.26 | 2.34 |
| Disallowed (%) | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| CC Mask | 0.82 | 0.84 | 0.84 | 0.75 | 0.77 | 0.83 | 0.71 |

[a]Gold-Standard Fourier-Shell Correlation
[b]Post-Processed with Resolve Cryo-EM from PHENIX