Supplementary data

Cryo-EM reveals a conserved baseplate architecture and distinct accessory protein assemblies in mycobacteriophages Claus, Corndog and Mysterious

Sunil Kumar Tewary[1], Chun-Hsiung Wang[1], Ming-Ching Lin[1,2], Melvin C. Shen[1], Li-An Tsai[1,2], Jitendra Maharana[1], Ronelito J Perez[1], Chi-En Mou[3] Graham F. Hatfull[4], Todd L. Lowary[1,2] and Meng-Chiao Ho[1,2*]

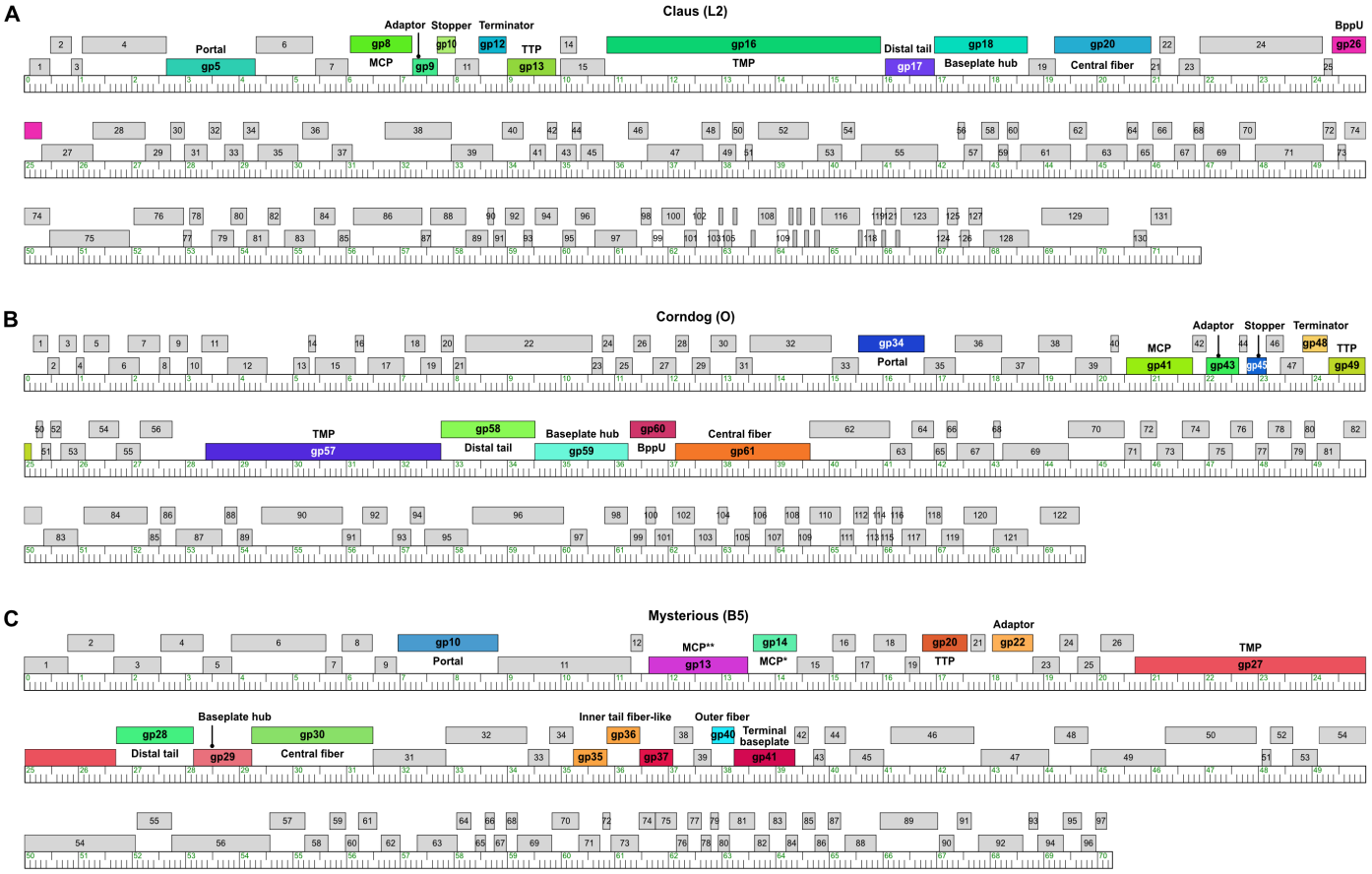
[1]Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan

[2]Institute of Biochemical Sciences, National Taiwan University, Taipei, Taiwan

[3]Department of Pharmacology and Toxicology, University of Toronto, Toronto, Canada

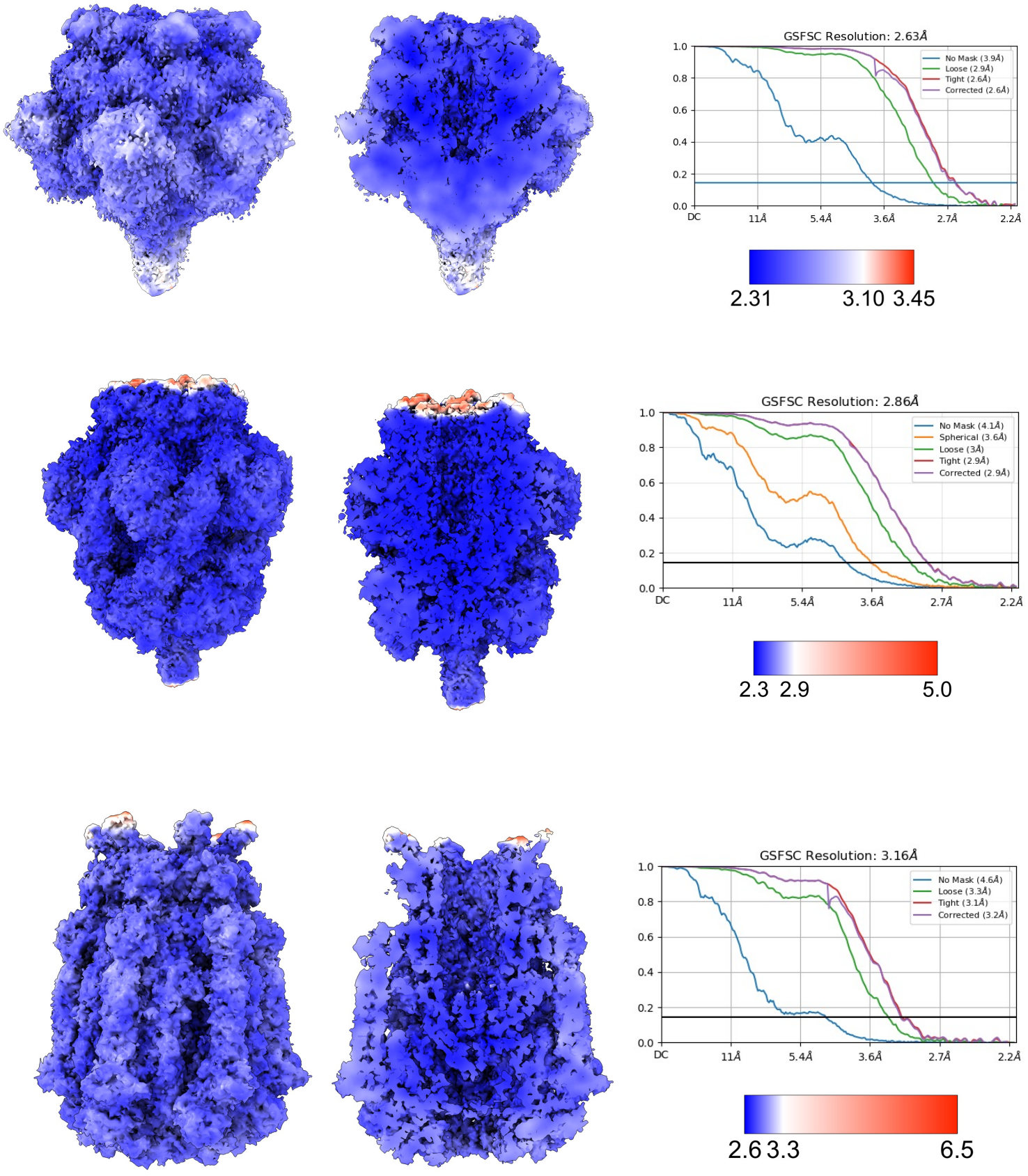
[4]Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania, USA.

[*]correspondence: joeho@as.edu.tw;tlowary@as.edu.tw

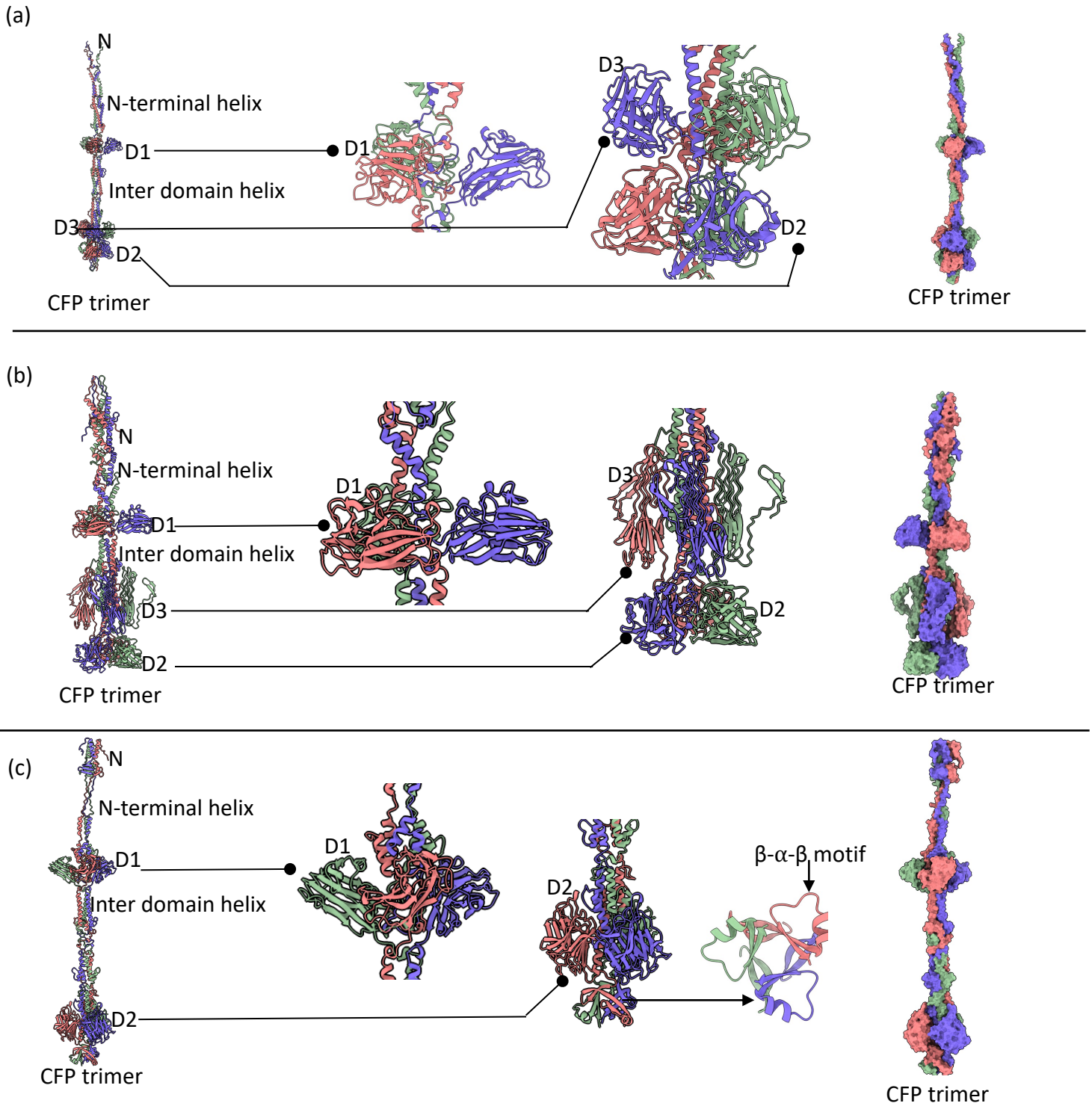


Supplementary Supplementary Figure 1. Comparative genomic organization of mycobacteriophages Claus, Corndog, and Mysterious.

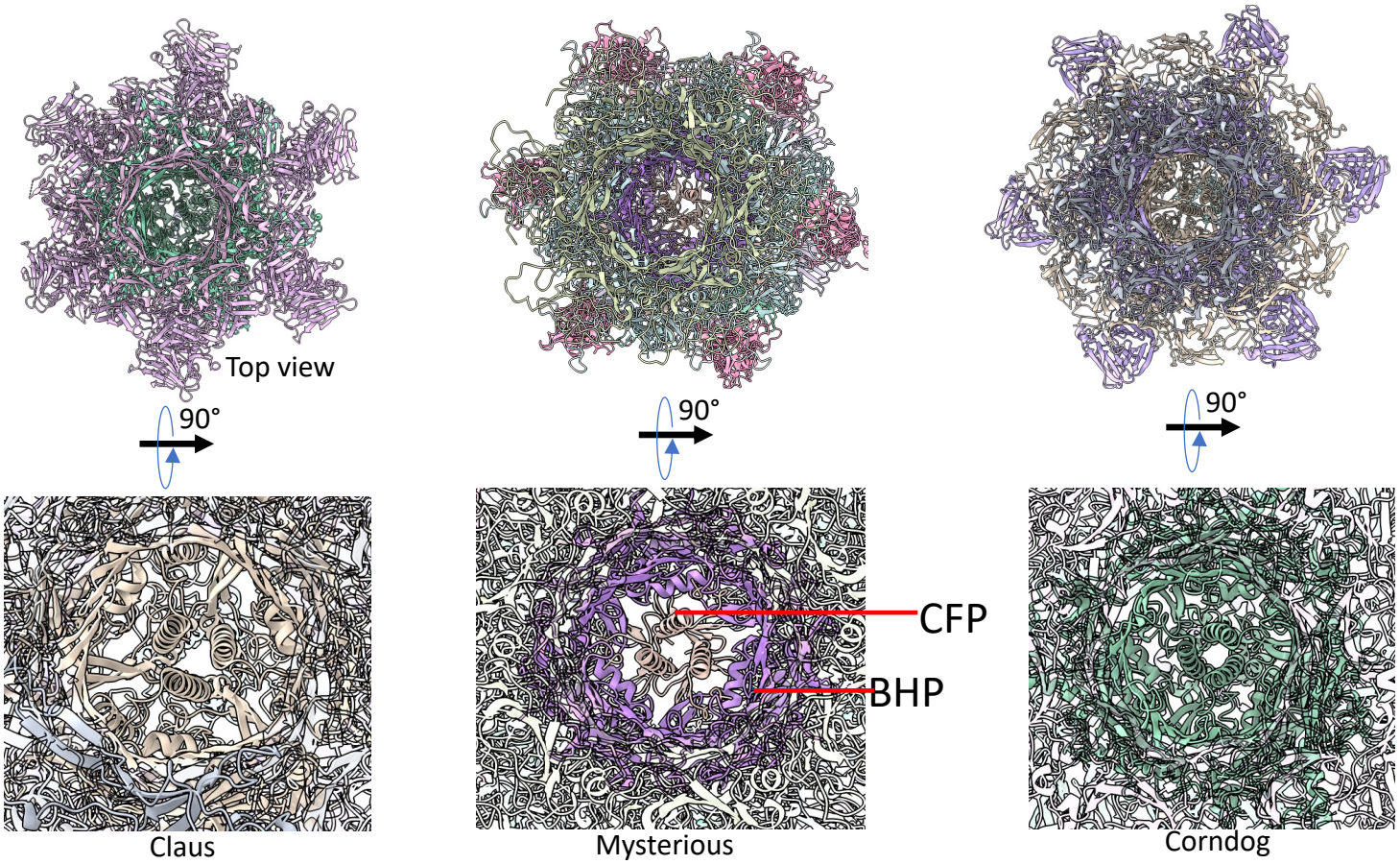
(A–C) Linear genome maps showing the annotated gene organization of (A) Claus (cluster L2), (B) Corndog (cluster L1), and (C) Mysterious (cluster B5). Genes are represented as boxes oriented in the direction of transcription, with colors highlighting conserved functional modules. Structural genes—including capsid, portal, protease, major capsid protein (MCP), scaffold, tape measure protein (TMP), baseplate, and tail components—are clustered and color-coded to emphasize synteny across the three phages. DNA metabolism and replication-associated genes are interspersed and shown in neutral tones, while regulatory and accessory genes are indicated separately. Conserved modules such as the head assembly, tail assembly, and baseplate regions are aligned to illustrate similarities and differences in genome architecture. Notable variations in gene content, order, and module boundaries are evident, particularly in the baseplate and tail fiber regions, reflecting diversification among clusters L2, L1, and B5. Scale bars indicate genome coordinates in kilobases.



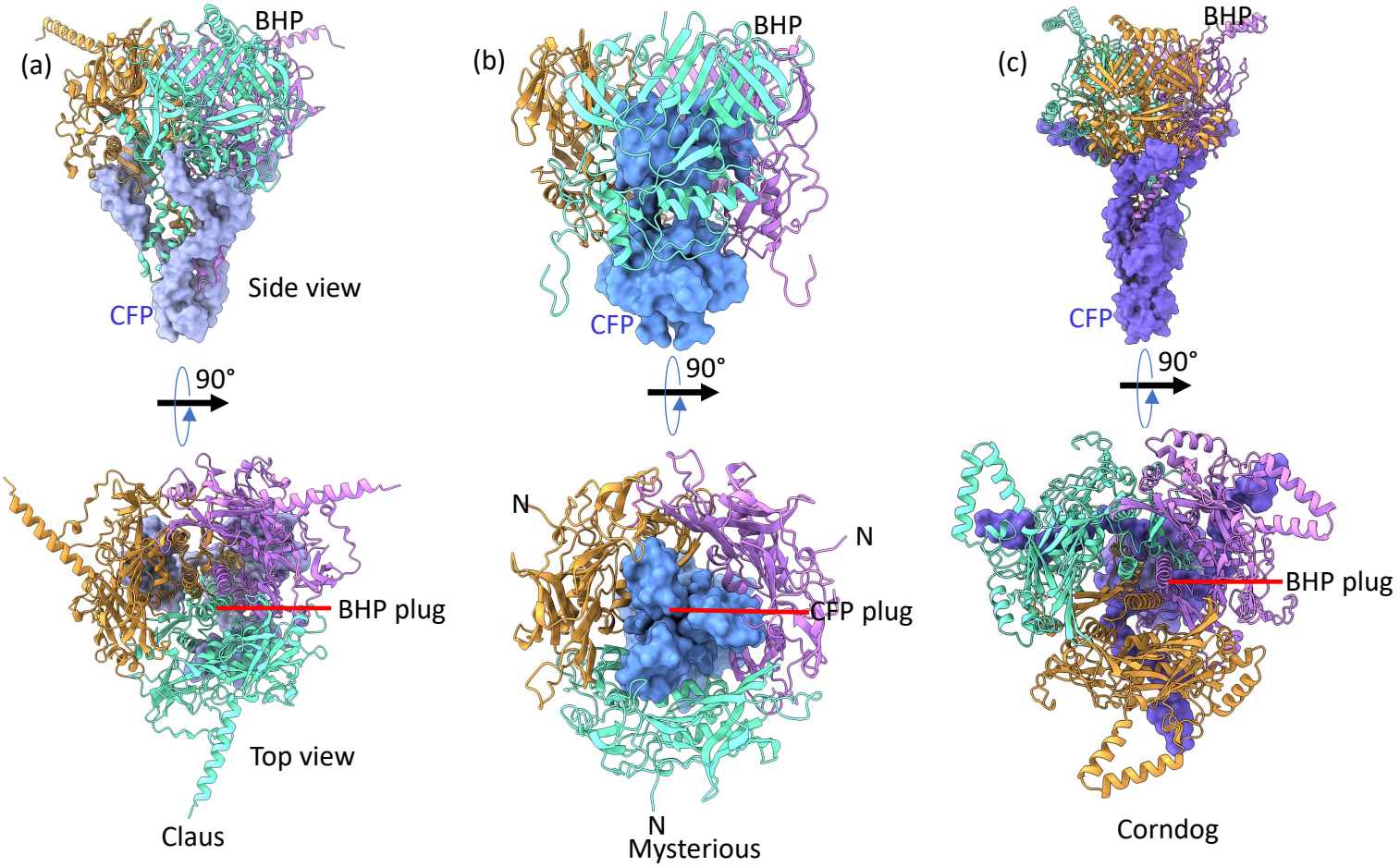
Supplementary Figure 2. Claus, Corndog and Mysterious map resolution. (a) Claus-local map-resolution (b) Corndog-local map-resolution (c) Mysterious-local map-resolution



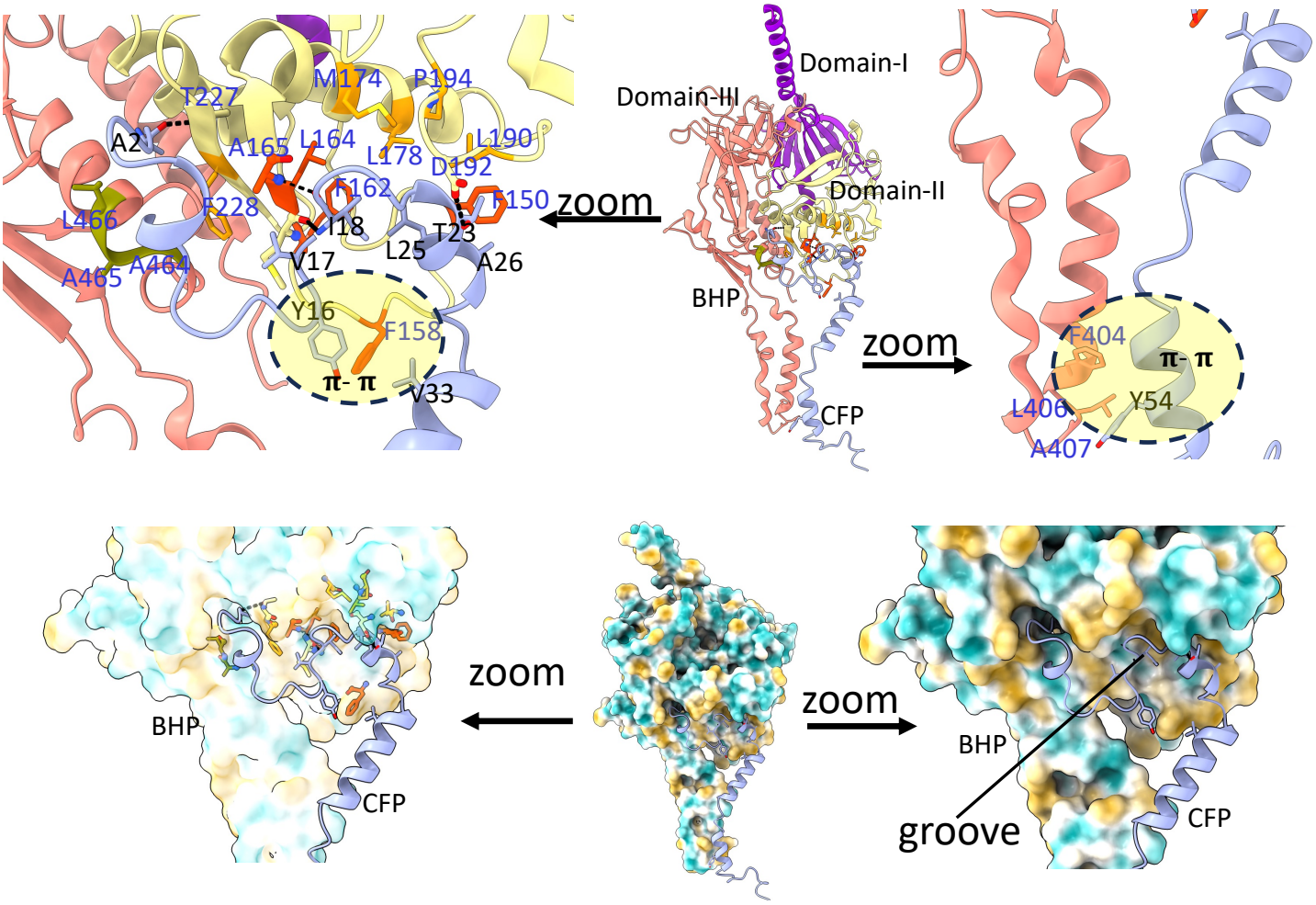
Supplementary Figure 3. Alpha-fold predicted central fiber protein (CFP) an accessory protein complex in Claus, Corndog and Mysterious. **(a)** The central fiber protein of Claus is predicted as a trimer displayed in both ribbon and surface representations. The CFP comprises three domains (D1, D2, and D3), with D1 separated from D2/D3 by a long interdomain α-helix. Domains D2 and D3 are arranged in tandem as a dimer of dimer to form the distal tip that form the receptor binding protein (RBP) domain. **(b)** Corndog CFP has also three domains D1, D2 and D3 but unlike Claus D3-trimer occupy a central position while the D2 present at the tip of the CFP. **(c)** Unlike Claus and Corndog CFP of Mysterious is a two domain (D1-D2) structure relatively longer than both of the phages. D2 domain is a trimer at the tip of the CFP with a small extended β-α-β motif at the nose of the tip. All the domains are shown as zoom view, three chains are green, red and blue colored.



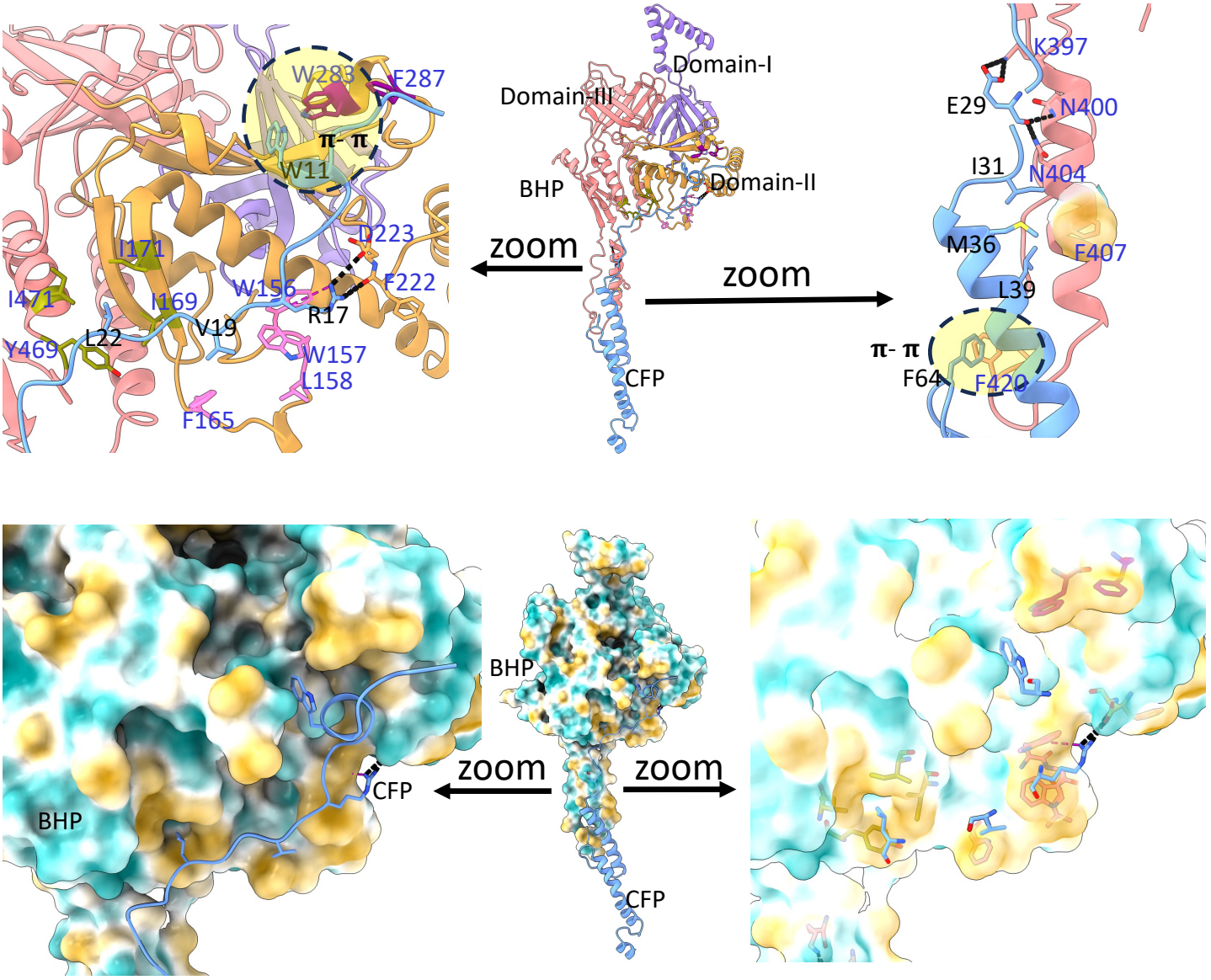
Supplementary Figure 4. CFP architecture in the baseplate of Claus, Mysterious and Corndog remain conserved. (a) Claus top view of the complete baseplate (left, upper panel) and zoomed view of the BHP and CFP (left, bottom panel). Rest of the surrounding oligomers are made transparent for clarity. (b) Mysterious baseplate top view showing the BHP and CFP (c) Corndog baseplate top view showing the BHP and CFP.



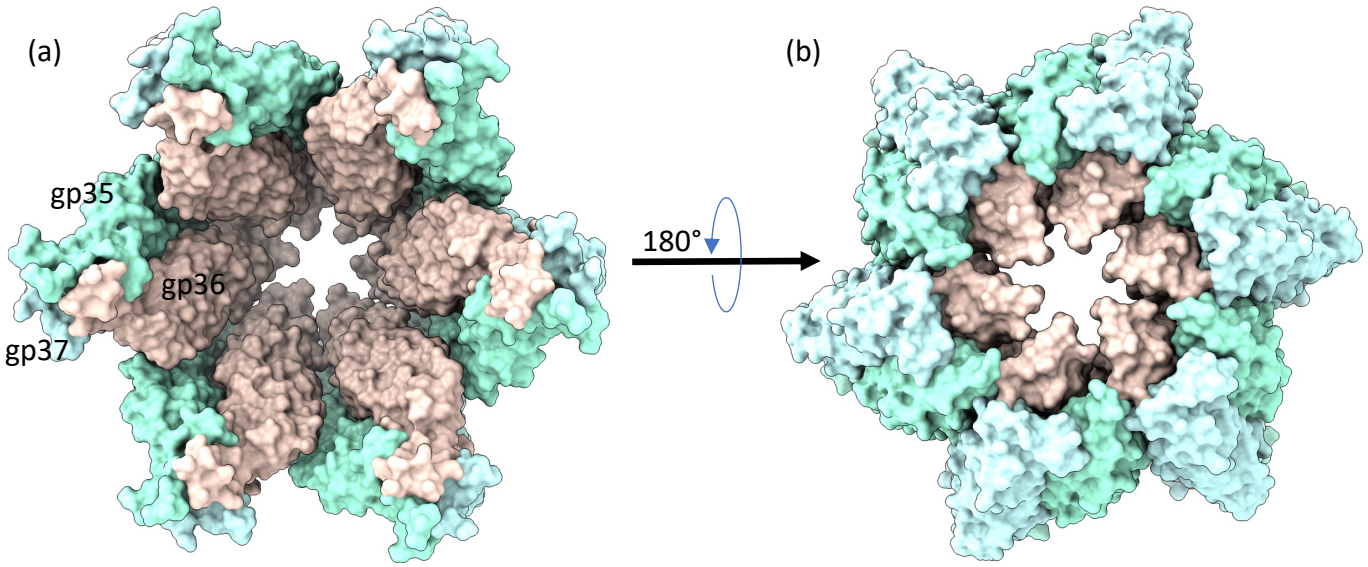
Supplementary Figure 5. BHP closure of the Claus, Mysterious and Corndog comparison. (a) Claus BHP wrap by CFP (Side view, left top panel**)** and top view showing BHP internal three helical bundle plug. (b) Mysterious BHP internal plug is outsourced to CFP (blue surface). (c) Corndog internal plug is same a claus.



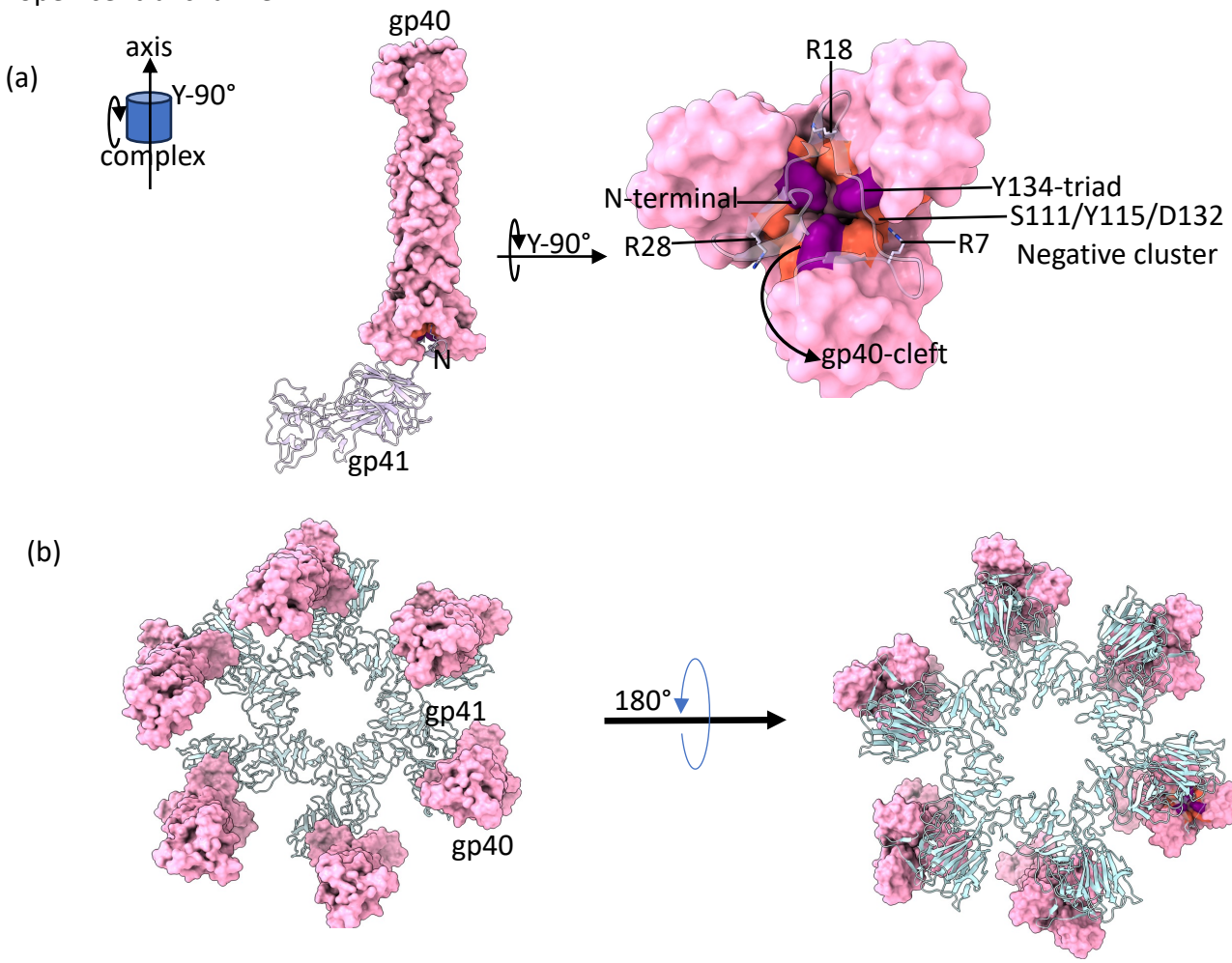
Supplementary Figure 6. BHP–CFP interaction interface in the Claus mycobacteriophage baseplate. Overall architecture of the Claus baseplate highlighting the interaction between the baseplate hub protein (BHP; gp18) and the connector/CFP (gp20). Top panels show zoomed views of the gp18–gp20 interface, revealing discrete contact regions and key interacting residues. Aromatic residues form prominent π–π stacking interactions at the core of the interface, while surrounding hydrogen-bonding and hydrophobic contacts stabilize the assembly. Three interaction patches on gp18 are highlighted: Patch-I red, (Phe150, Phe158, Phe162, Met163, Leu164, and Ala165), Patch-II orange, (Met174, Leu178, Leu190, Pro194, and Phe228), and Patch-III olive, (Ala464, Ala465, and Leu466). Bottom panels show electrostatic surface representations of BHP and CFP, illustrating a complementary groove on BHP that accommodates CFP, promoting tight packing and interface specificity. Together, hydrogen-bond anchors and π–π stacking interactions generate a stable and well-organized gp18–gp20 interface.



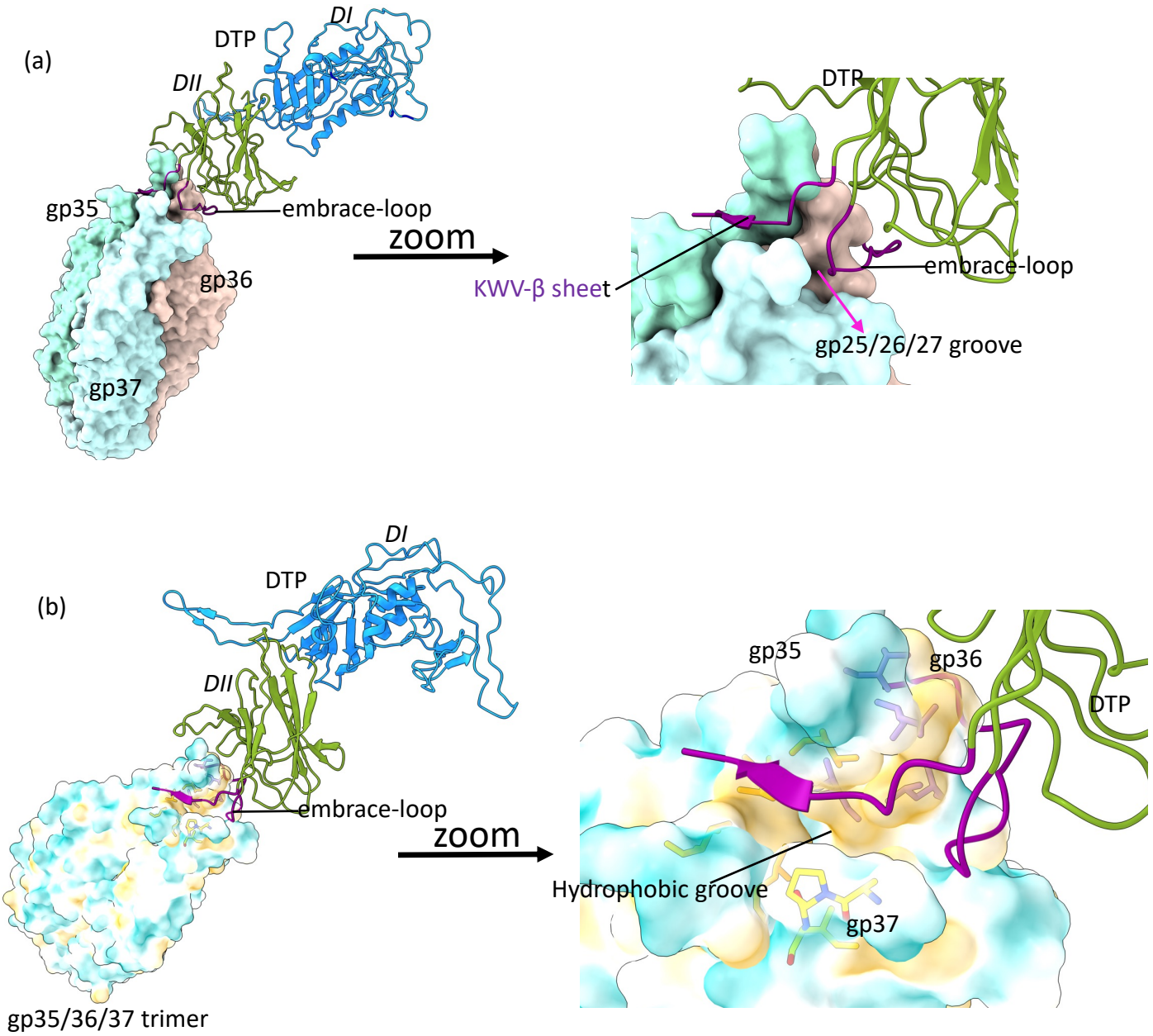
Supplementary Figure 7. Corndog BHP–CFP (gp59–gp61) interaction interface. Structural overview of the BHP–CFP complex highlighting the interaction network that stabilizes the gp59–gp61 interface. Zoomed views show three hydrophobic patches on gp59: Patch I pink, (W156, W157, L158, F165), Patch II brown, (W283, F287), and Patch III green, (I169, I171, I471, Y469). Aromatic residues from these patches form π–π and cation–π interactions, while additional salt bridges and hydrogen bonds, including contacts involving the CFP N-terminal region, further reinforce binding. Surface representations illustrate complementary hydrophobic grooves on BHP that accommodate CFP, together forming a cooperative interaction network that locks the complex in place.



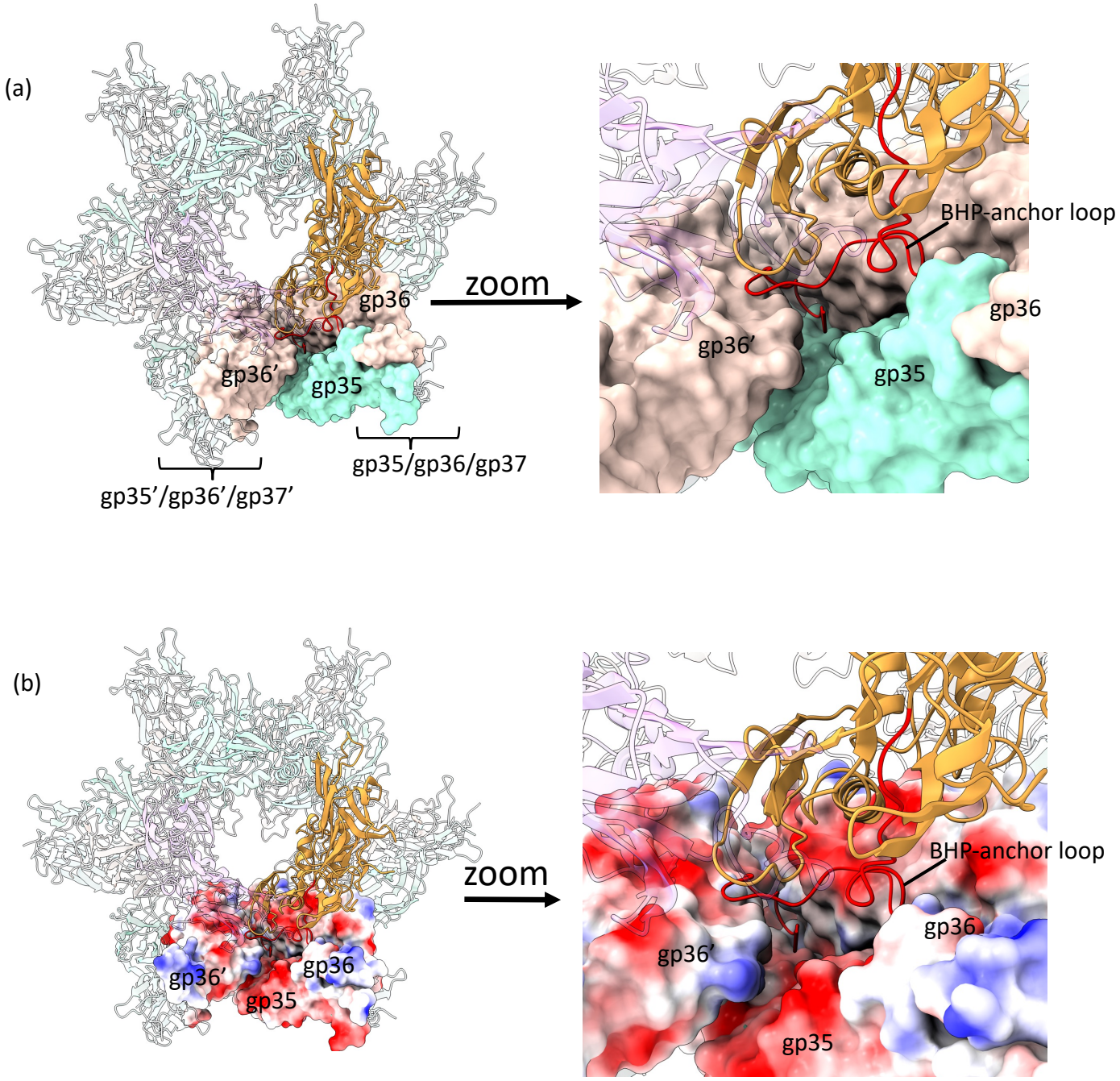
Supplementary Figure 8. Mysterious gp35, gp36 and gp37 heterotrimer assembly. (a) Top view of the hexmeric trimer showing open central channel. (b) Bottom view of the hexamric heterotrimer showing open central channel.



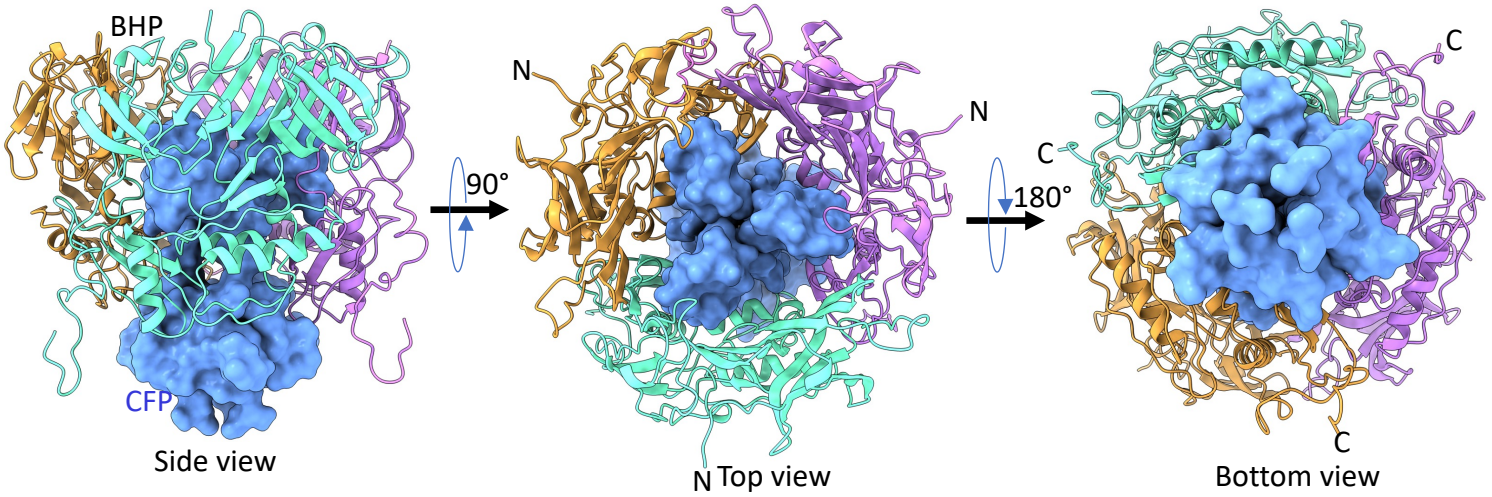
Supplementary Figure 9. Mysterious gp40–gp41 complex interactions. (a) Gp41 N-terminal residues R7, R18, and R28 (purple) insert into the gp40 groove, stabilized by the gp40 S111/Y115/D132 negative charge cluster (pink) and Y134 triad (light purple) within the cleft. (b) *Mysterious* gp40 and gp41 complex hexagonal tube structure complex looking form the top and bottom view



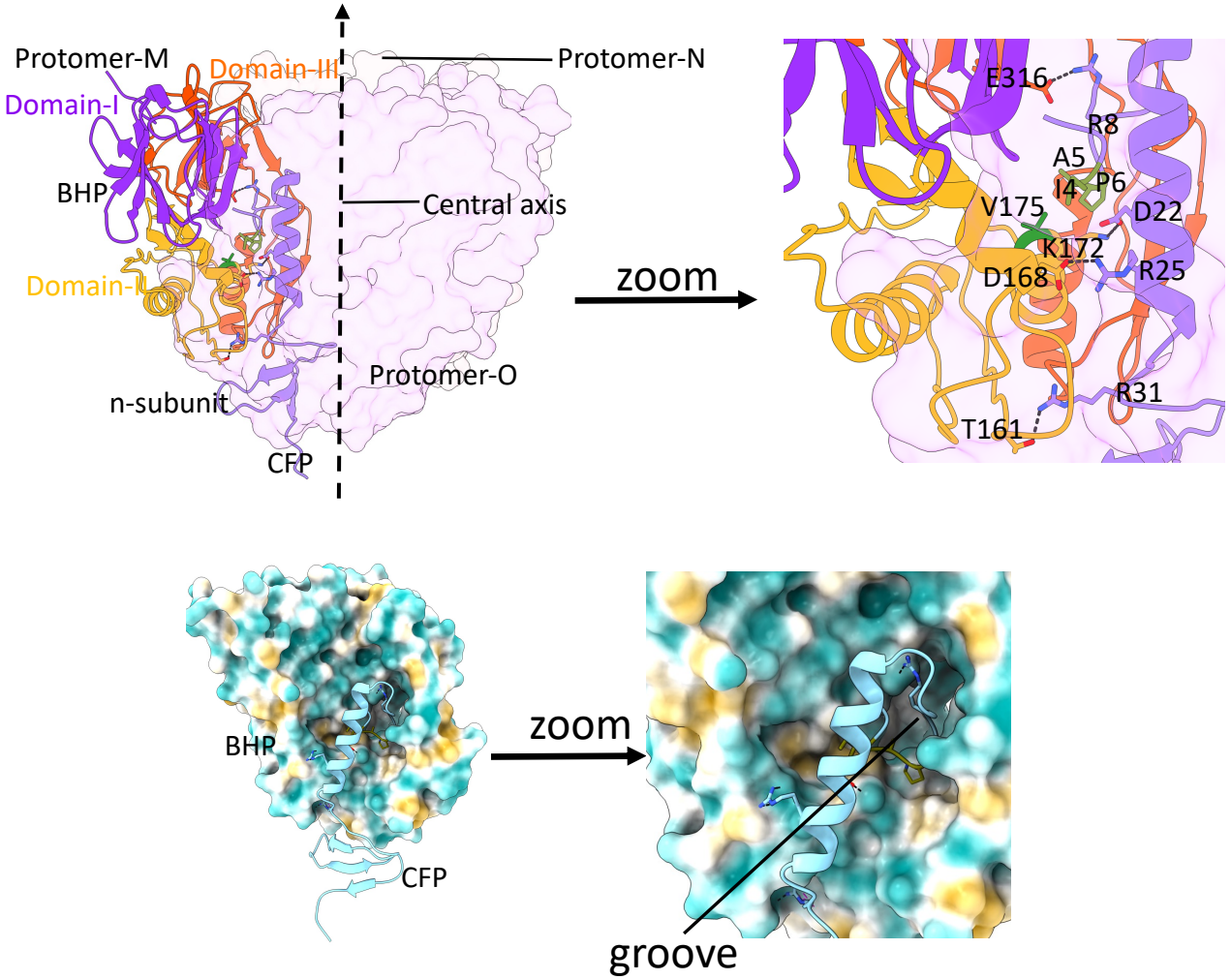
Supplementary Figure 10. Mysterious gp35/36/37 and gp41 complex interactions. (a) DTP embrace loop (293–345; TQPIKWVHGADCEQPADCAAMPILFSDTCTIETIDVITSPPPNCGGCLPVGLL; blue residues not modelled because electron density is missing) inserts into the N-terminal groove of the gp35/36/37 trimer, while DTP residues K297–W298–V299 form a short β-strand that pairs in parallel with an N-terminal β-strand of gp35, providing directional stabilization. (b) The same figure in a is shown here but the gp35/36/37 trimer hydrophobic surface forming the key residues are shown as stick, while the gp28 embrace loop shown as purple ribbon.



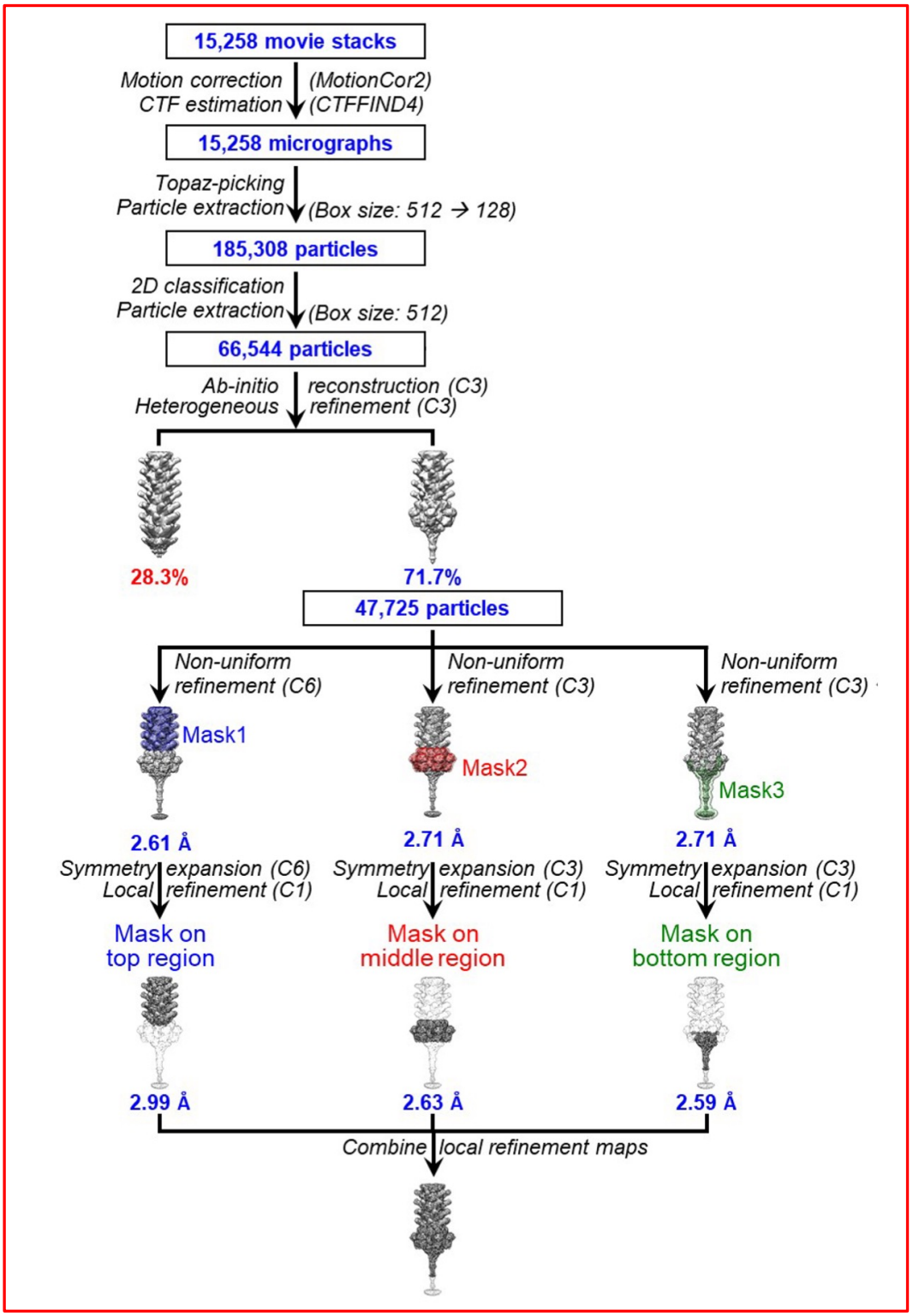
Supplementary Figure 11. C-terminal anchor loop of BHP stabilizes inter-trimer interfaces in the baseplate. (a)The C-terminal region of the BHP (red) extends outward to form an anchor loop that latches onto the interface between neighboring gp35/36/37 trimers (dashed), positioning it outside the central cavity. This loop adopts a U-shaped geometry at the end, descending along the outer surface of the complex and engaging predominantly with gp36, with more limited contacts to gp35. Acting as a clamping shaft, the BHP C-terminal anchor secures adjacent trimeric units and reinforces inter-trimer connectivity. (b) The interactions are electrostatic in nature.



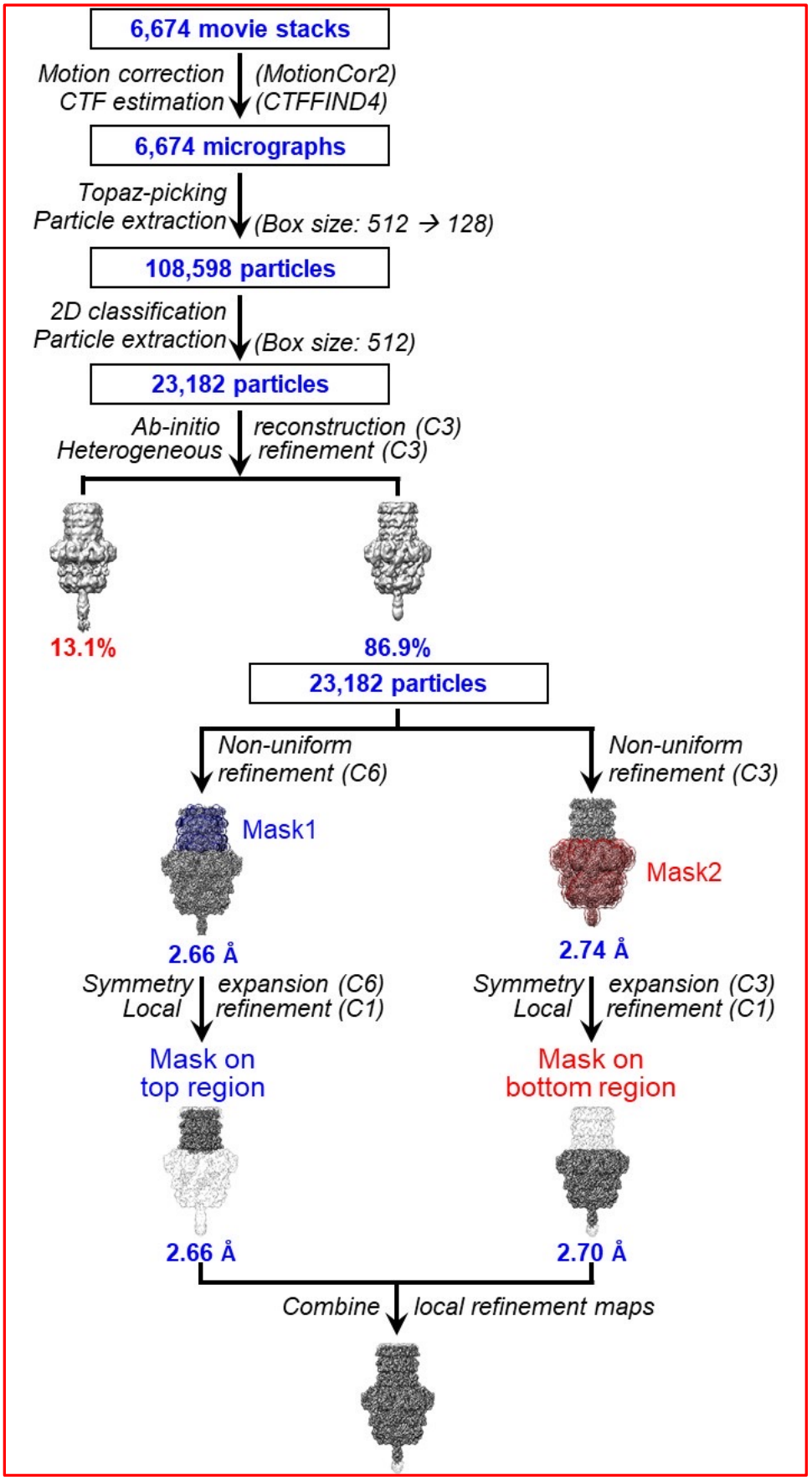
Supplementary Figure 12. Mysterious CFP (gp30) filling the central cavity of BHP (gp29) homotrimer. Side, top and bottom view of the complex showing CFP filling the central cavity of BHP trimer.



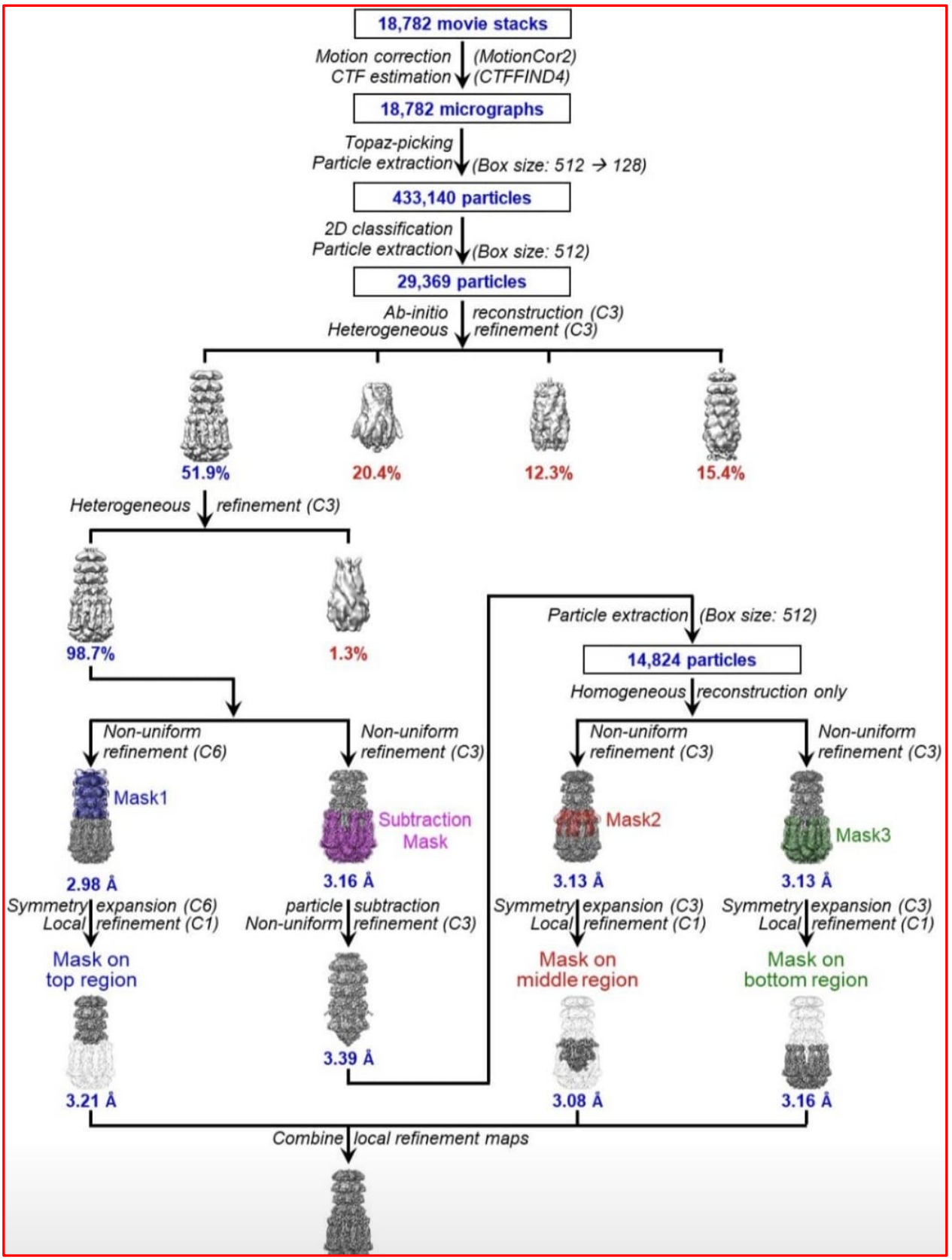
Supplementary Figure 13. Mysterious BHP-CFP interactions. Top left Overall structure of the BHP–CFP complex showing the domain organization of BHP (Domains I–III) and one subunit of the bound CFP. BHP two protomers are shown as transparent surface while the interacting protomer as ribbon. (Top right) Zoomed-in view of the BHP–CFP interface highlighting key interacting residues. Salt bridges and hydrogen bonds are shown as dashed lines, illustrating the electrostatic network that stabilizes the interface. Hydrophobic patch of CFP (olive) and BHP (green). (Bottom left) Surface representation of BHP with CFP shown as a ribbon, revealing the surface groove that accommodates CFP. (Bottom right) Magnified view of the groove illustrating the deep insertion of the CFP helix and the shape-complementary interface formed by combined electrostatic and hydrophobic interactions. Collectively, salt bridges, hydrogen bonds, and hydrophobic contacts cooperatively lock CFP within the BHP groove.



Supplementary Figure 14. Claus baseplate 3-D map reconstruction. Cryo-EM data processing and refinement workflow for the Claus baseplate complex. A total of 15,258 micrographs were processed in cryoSPARC, yielding 185,308 initially classified particles and 47,725 particles contributing to the final high-resolution reconstruction through symmetry-guided non-uniform and local refinements, with masked regions resolved to 2.59–2.99 Å. Final map resolutions were determined using gold-standard FSC at the 0.143 criterion.



Supplementary Figure 15. Corndog baseplate 3-D map reconstruction. Cryo-EM data processing and reconstruction workflow for the Corndog baseplate complex. From 6,674 micrographs, iterative particle picking, classification, symmetry-guided refinement, and local C1 refinements yielded a composite reconstruction with region-specific resolutions of 2.66–2.70 Å, estimated using the gold-standard FSC 0.143 criterion.



Supplementary Figure 16. Mysterious baseplate 3-D map reconstruction. Cryo-EM data processing and reconstruction workflow for the Mysterious baseplate complex. From 18,782 micrographs, extensive particle classification and symmetry-guided refinement produced region-specific reconstructions of the top, middle, and bottom regions at 3.08–3.39 Å resolution, based on gold-standard FSC at the 0.143 criterion.