

Supplementary Information for

Structural atlas of the intact jumbo phage phiKZ

Hao Xiao^{1,2,3#}, Zeng Peng^{1,2#}, Junquan Zhou^{1,2#}, Yuan Chen^{1,2}, Yuning Peng^{1,2}, Yixiong Tang^{1,2}, Tao Li⁴, Wenyan Chen^{1,2}, Sheng-You Huang⁴, Lingpeng Cheng^{1,2}, Hongrong Liu^{1,2,3*}

¹Institute of Interdisciplinary Studies, Hunan Normal University, Changsha 410082, China.

²Key Laboratory for Matter Microstructure and Function of Hunan Province, Key Laboratory of Low-dimensional Quantum Structures and Quantum Control, School of Physics and Electronics, Hunan Normal University, Changsha 410082, China.

³Hunan Research Center of the Basic Discipline for Quantum Effects and Quantum Technologies, Hunan Normal University, Changsha 410081, China.

⁴School of Physics, Huazhong University of Science and Technology, Wuhan 430074, China.

*Correspondence author: hrliu@hunnu.edu.cn

#These authors contributed equally to this work.

This PDF file includes:

Supplementary Figures S1 to S19

Supplementary Table S1 and S2

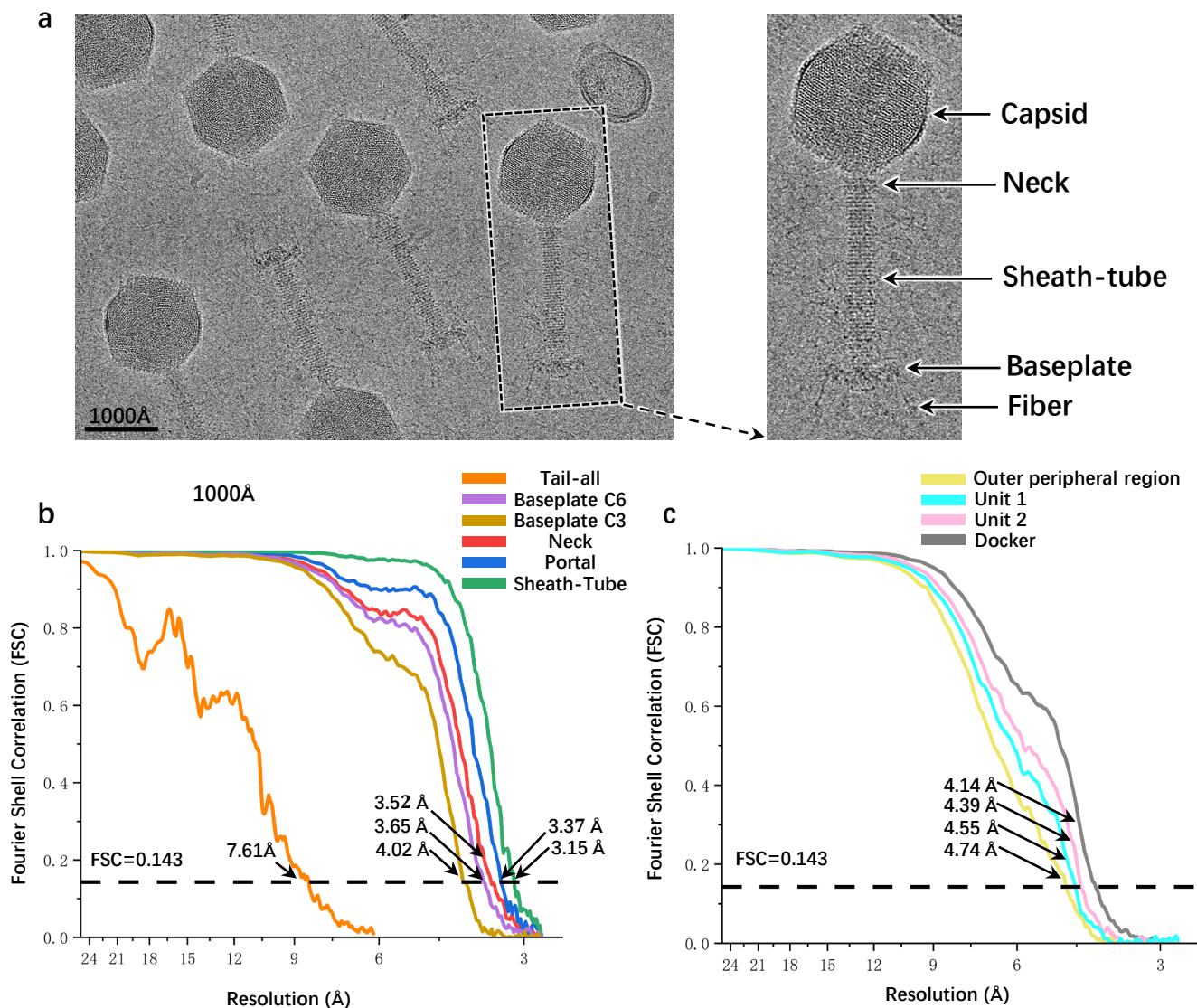
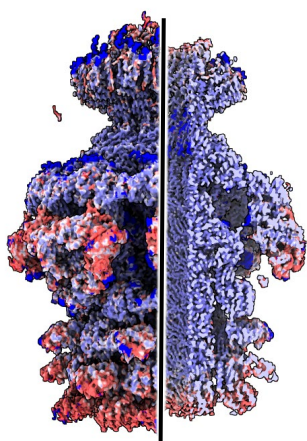
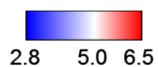
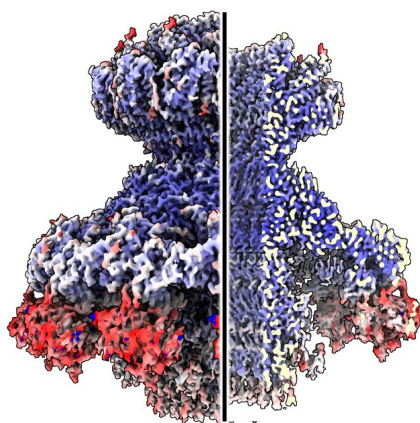
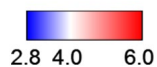
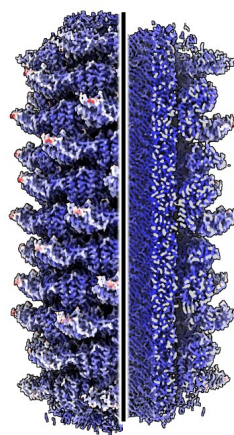
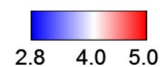
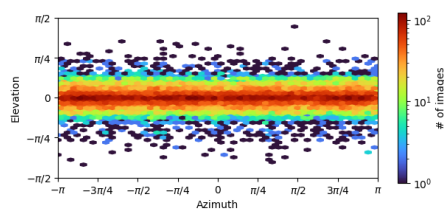
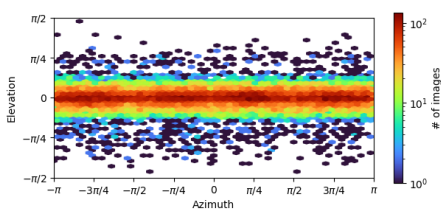


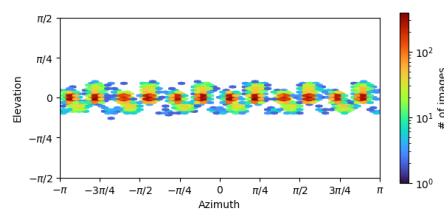
Figure S1. Cryo-EM images and Fourier shell correlation curves of the phiKZ virion. (a) Cryo-EM images of the phiKZ particles. **(b)** Structural resolutions of the overall tail, portal, neck, sheath-tube, and baseplate in the phiKZ virion. **(c)** Structural resolutions of the overall outer peripheral region, unit1, unit2, and docker in the phiKZ virion.



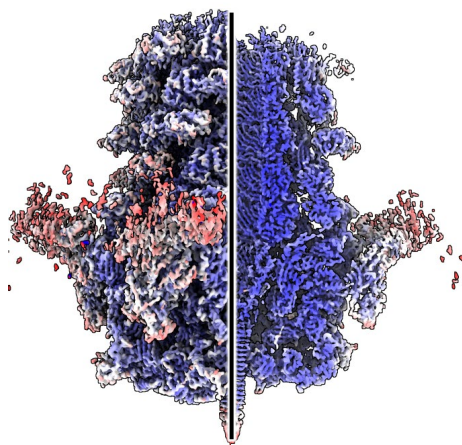
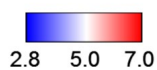
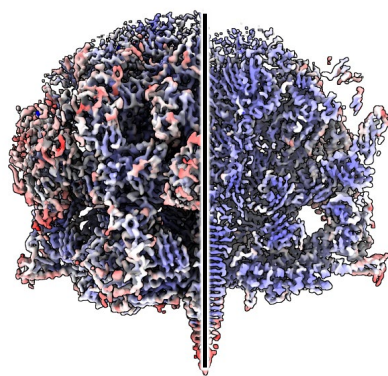
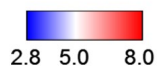
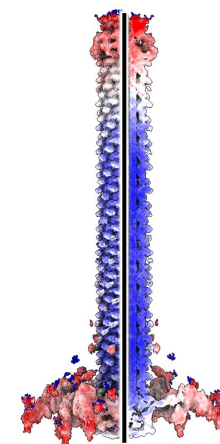
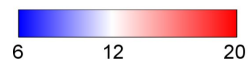
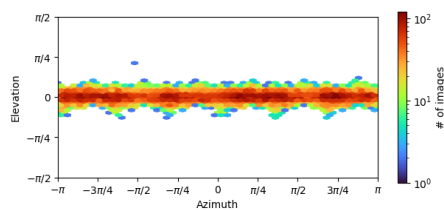
Neck 3.52Å



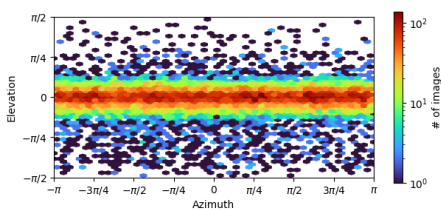
Portal 3.37Å



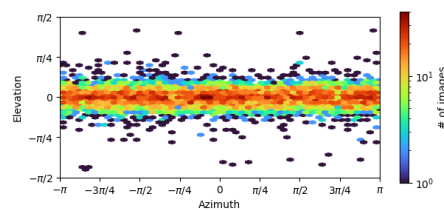
Tail sheath and tube 3.17Å



Baseplate-C6
3.65Å



Baseplate-C3
4.02Å



Tail-overall 7.61Å

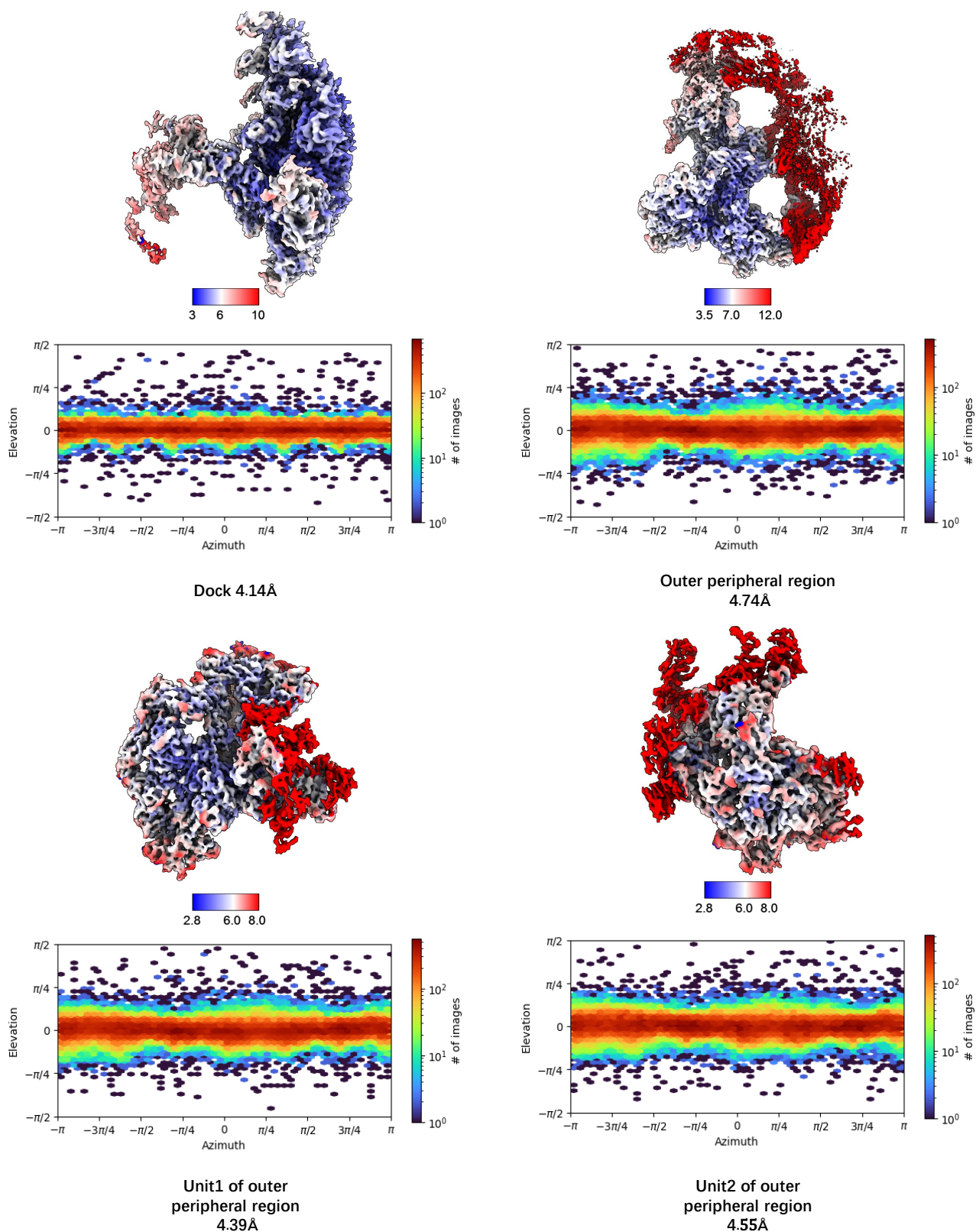


Figure S2. Local-resolution maps and angular distributions of all phiKZ reconstructions used for model building. The global resolution of each structure is indicated.

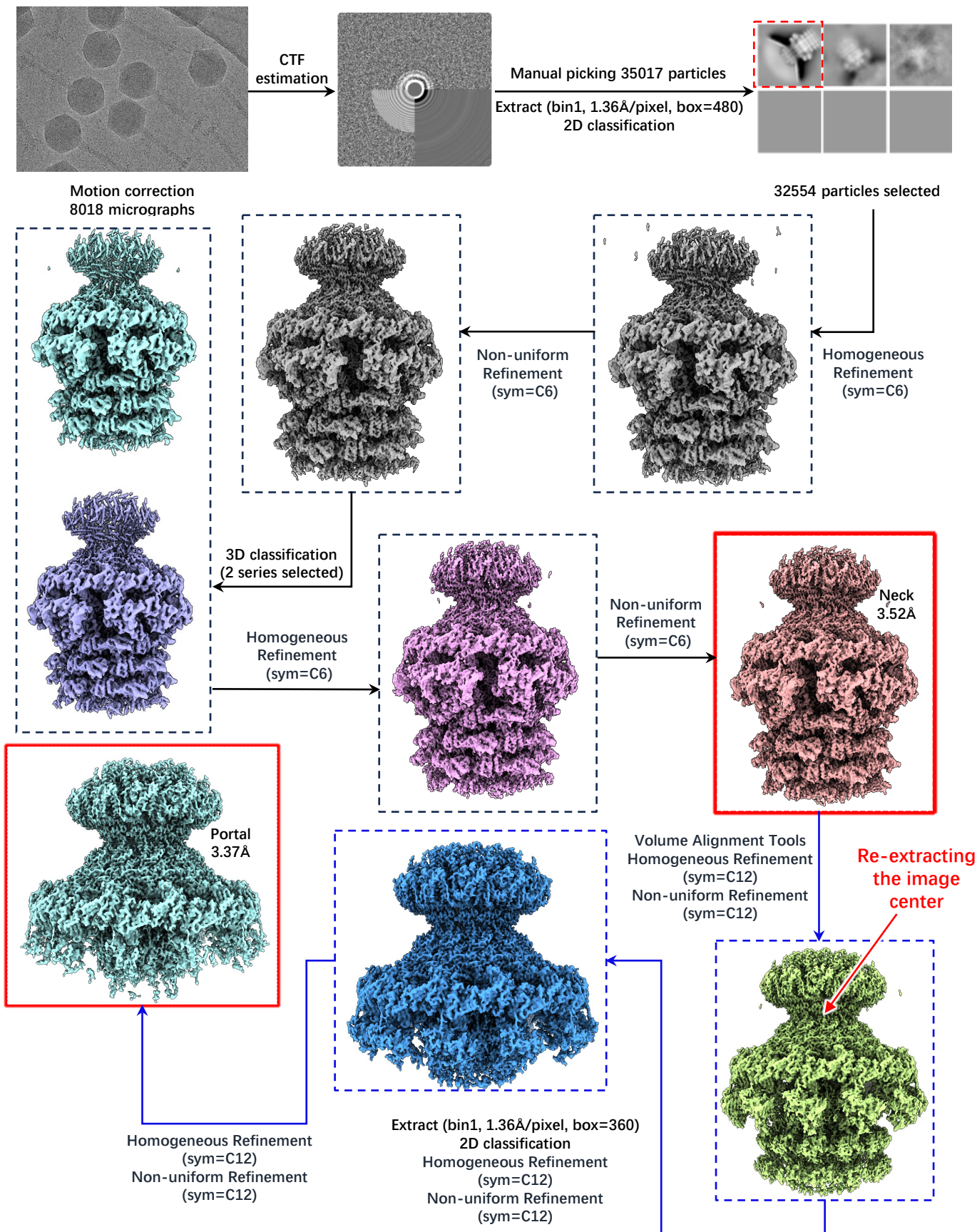


Figure S3. Flowchart of Cryo-EM data processing of the phiKZ portal and neck. Workflow illustrating the step-by-step process for phiKZ neck reconstruction.

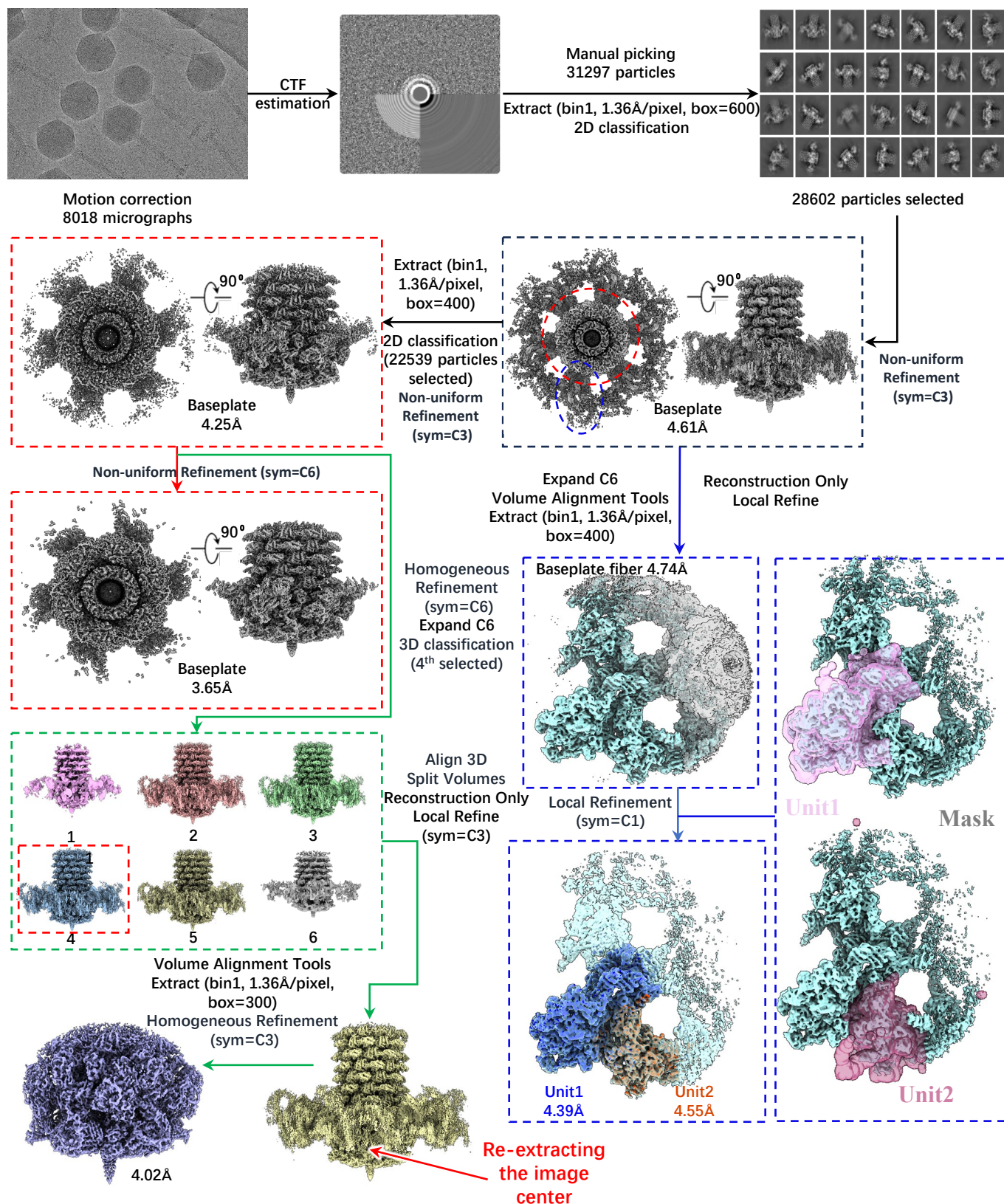
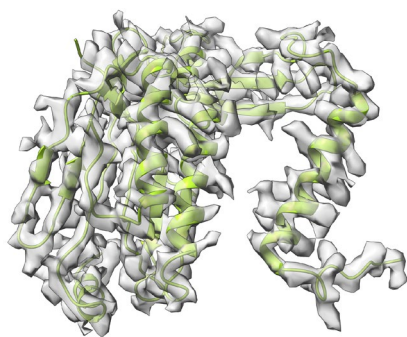
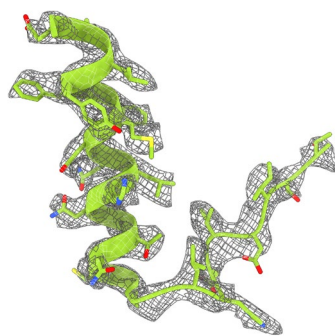


Figure S4. Flowchart of Cryo-EM data processing of the phiKZ baseplate-fiber complex.

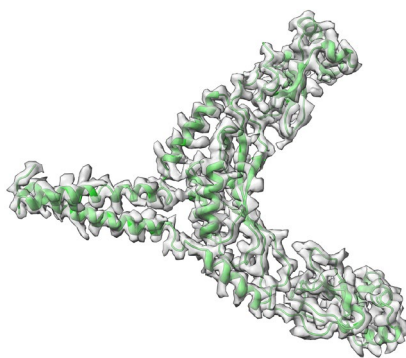
Workflow illustration the step-by-step process for reconstruction of phiKZ baseplate and its associated baseplate-fiber complex.



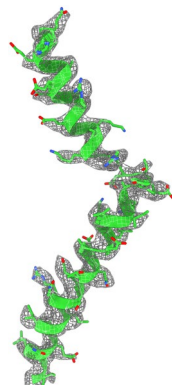
Gp34



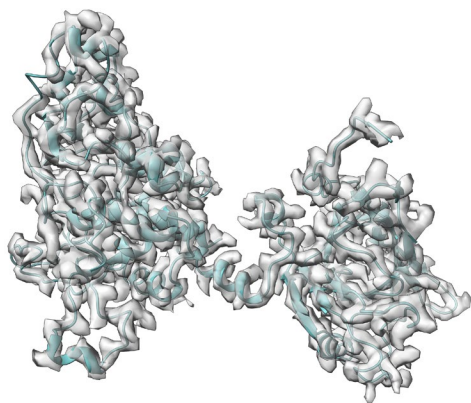
T12-E38



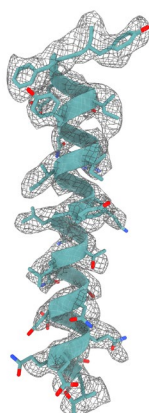
Gp161



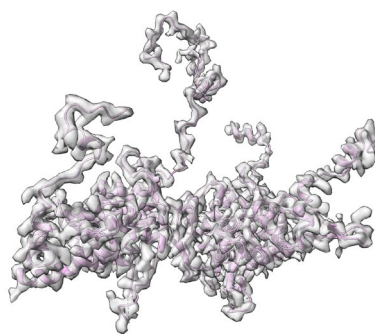
Q153-I197



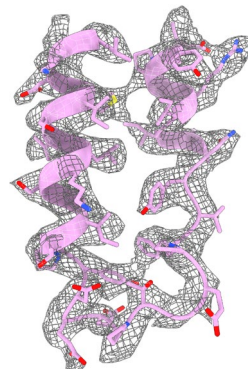
Gp99



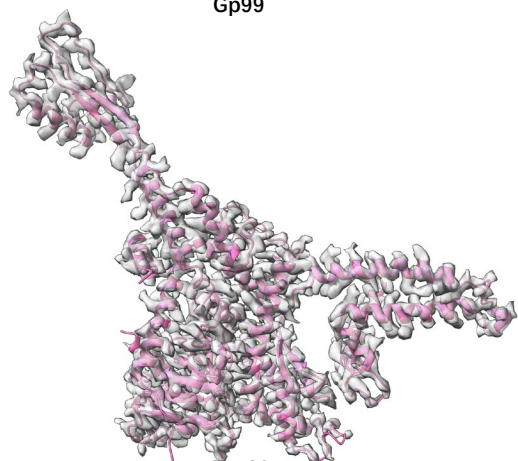
T12-Y223



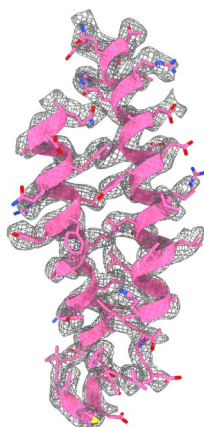
Gp174



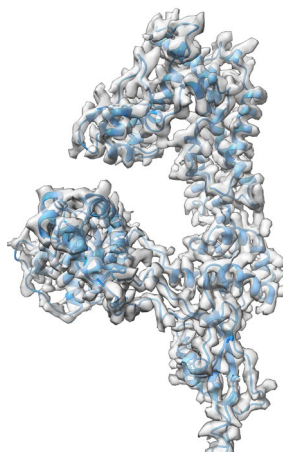
D253-L287



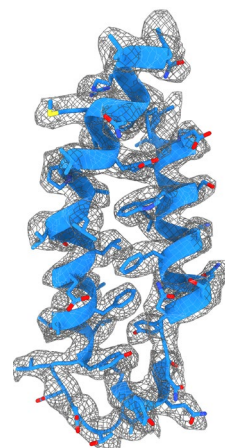
Gp129



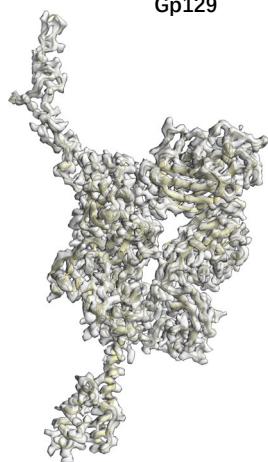
D687-E748



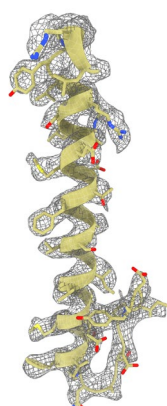
Gp128



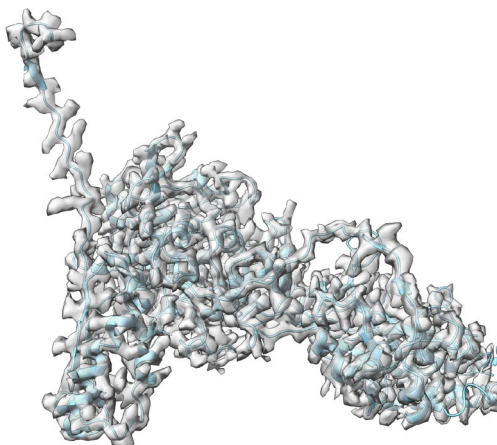
F155-N205



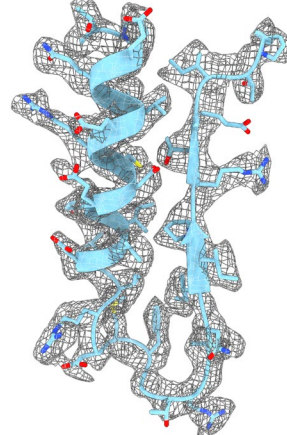
Gp87



Y39-L77



Gp29



T621-P655

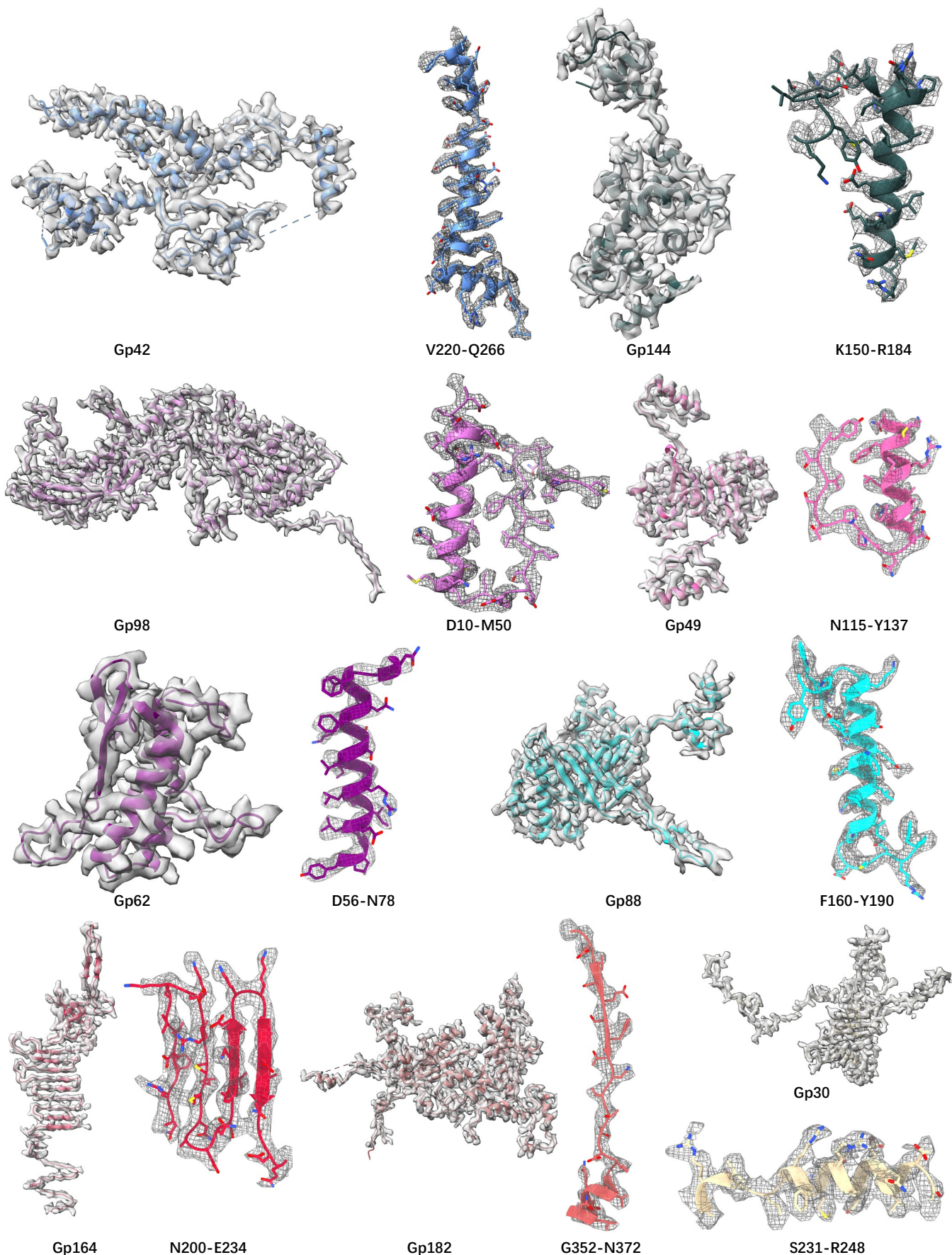


Figure S5. Density maps (transparent or mesh) superimposed on atomic models for proteins gp29, gp30, gp34, gp42, gp49, gp62, gp87, gp88, gp161, gp98, gp99, gp128, gp129, gp144, gp164, gp174, and gp182.

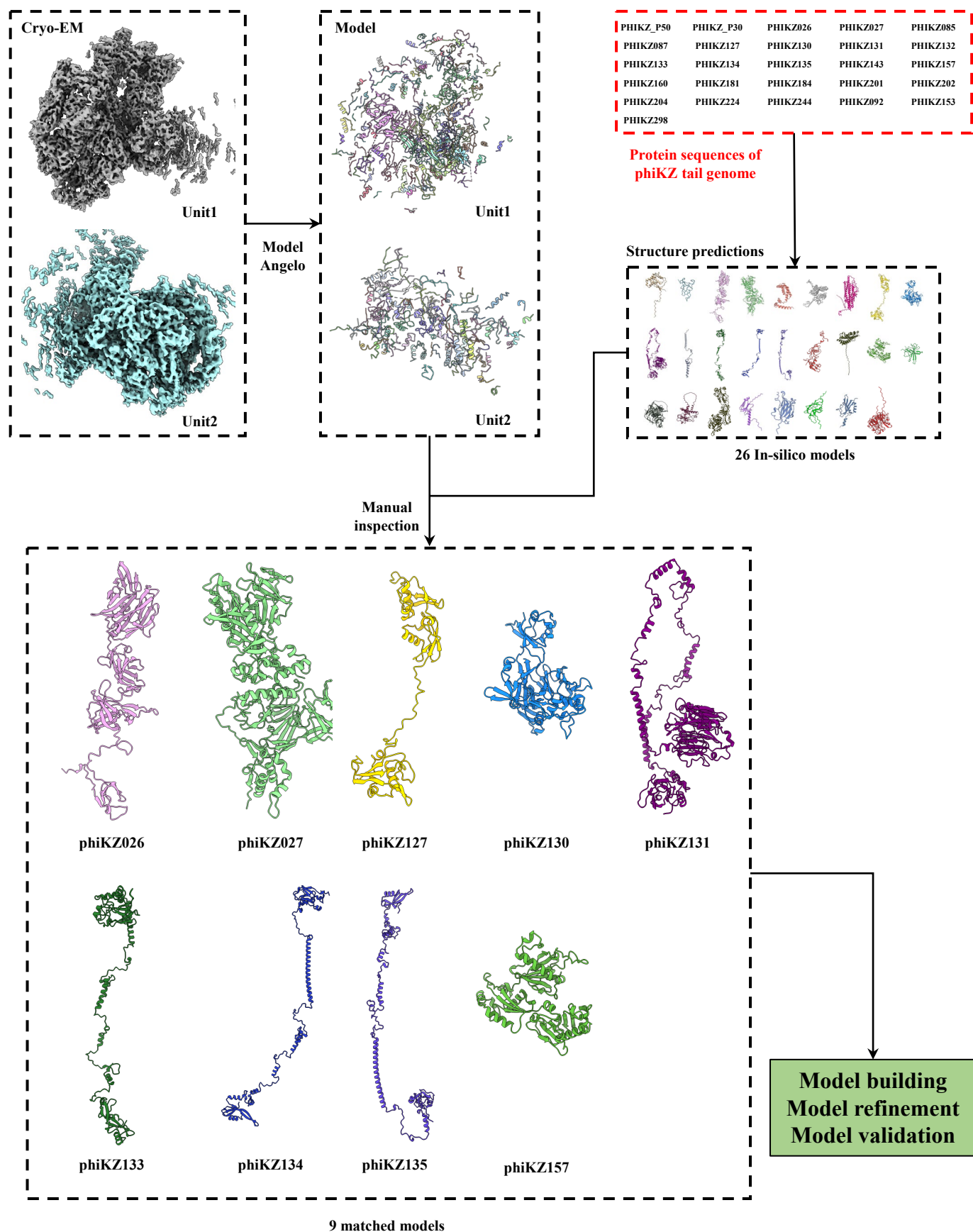


Figure S6. Flowchart for the identification and characterization of fiber proteins. The structural models presented were generated by AlphaFold3.

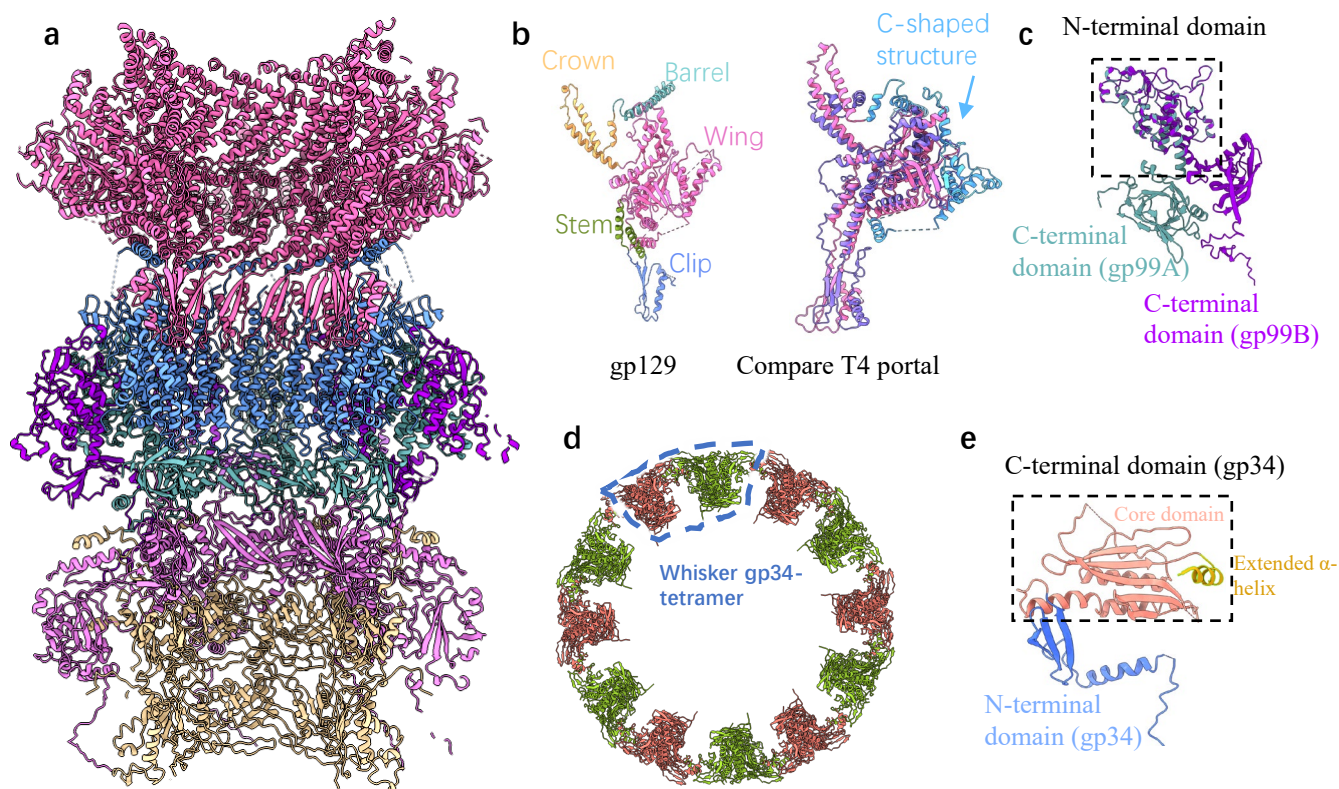


Figure S7. Structures of the phiKZ portal-neck complex. **(a)** The neck channel assembled from the portal, adaptor, stopper, and terminator. **(b)** Ribbon models of the portal gp129, with distinct domains highlighted in different colors. Structural comparison of the portal proteins from phage phiKZ (pink) and phage T4 (purple), with the additional α -helical region in phiKZ highlighted in blue. **(c)** Two conformations of gp99. **(d)** The hexameric whisker ring comprises six gp34 tetramers. Adjacent tetramers are alternately colored green and salmon to illustrate their arrangement. **(e)** Ribbon models of the whisker gp34, with distinct domains highlighted in different colors. The N-terminal domain is shown in blue, the C-terminal core domain in salmon, and the extended α -helix in yellow.

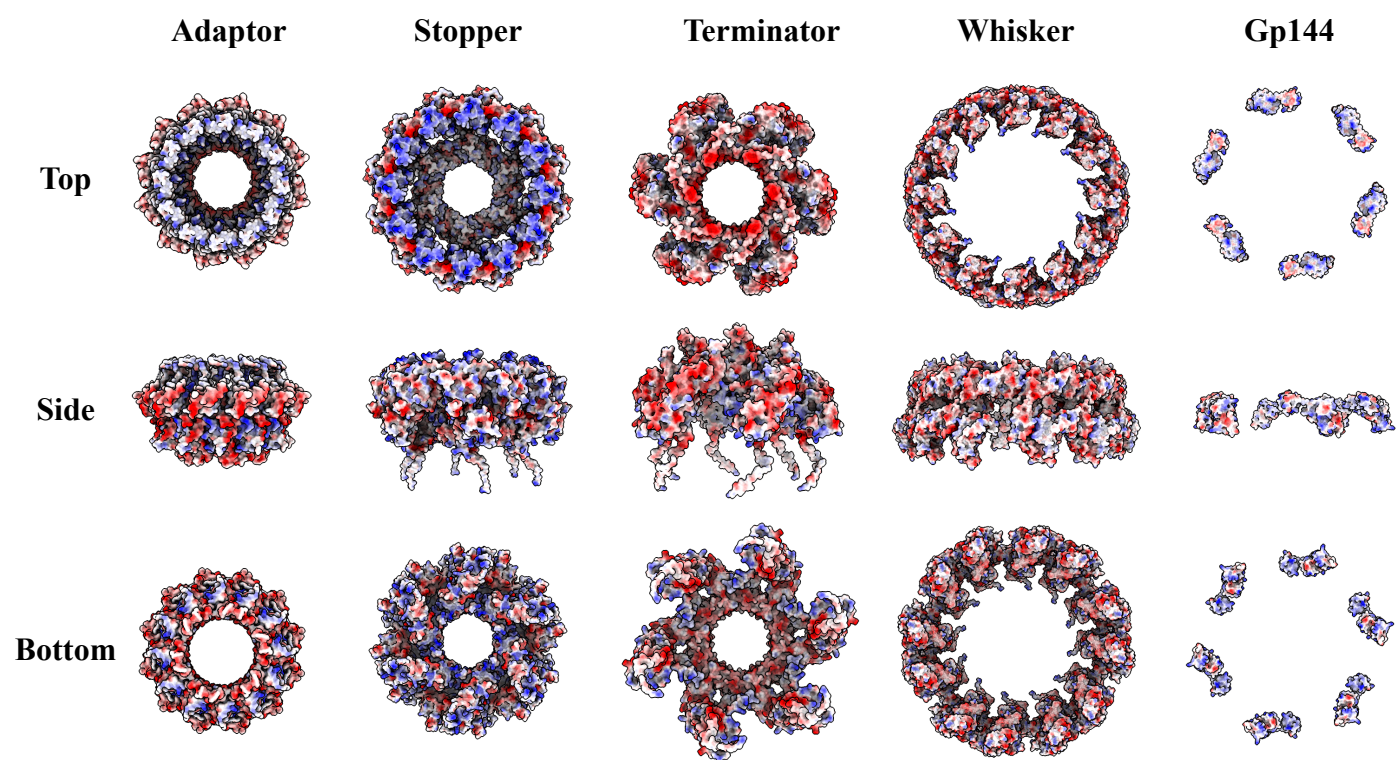


Figure S8. Electrostatic potential surface of the neck protein. Electrostatic potential surfaces of the neck protein are displayed from three top, lateral, and bottom views. The dimeric interfaces between neck proteins exhibit both electrostatic and structural complementarity.

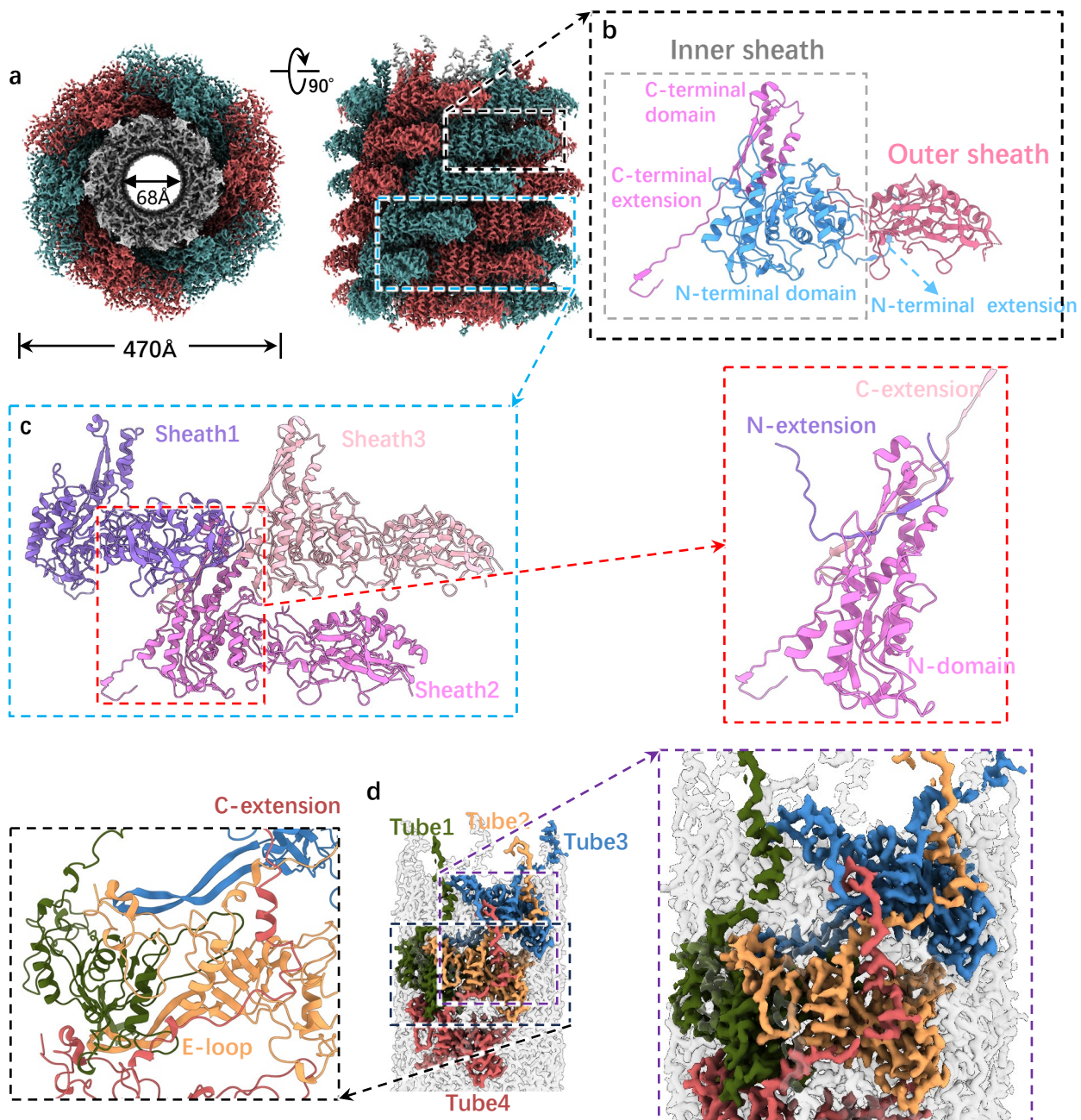


Figure S9. Organization of the sheath and tube. **(a)** Helical tail structure shown in top (left panel) and side (right panel) views. Adjacent strands of the six-start helix are colored distinctly. **(b)** Tail sheath composition: the inner sheath (dashed outline, pink/light blue) and extended outer sheath components (rose red). **(c)** Lower-layer sheath subunits (pink) are highlighted to elucidate inter-ring interactions. The interface between three sheath monomers spanning two consecutive rings is explicitly resolved in the zoomed-in view panel (red dashed box). **(d)** Cryo-EM density of three hexameric rings in the phiKZ tail tube, with subunits in the three discs highlighted in four distinct colors (central panel). The C-terminal extension of a lower-disc subunit (red) penetrates into the core component of an upper-disc subunit (blue) in the right panel. The N-terminal domain of this lower-disc subunit (red) engages with both: (i) the N-terminal domain of a middle-left-disc subunit (green), and (ii) the E-loop of a middle-right-disc subunit (yellow) in the left panel. The E-loop and N-terminal domain of the middle-right subunit (yellow) collaboratively form a clamp-like structure that encircles the core domain of the middle-left subunit (green) in the left panel.

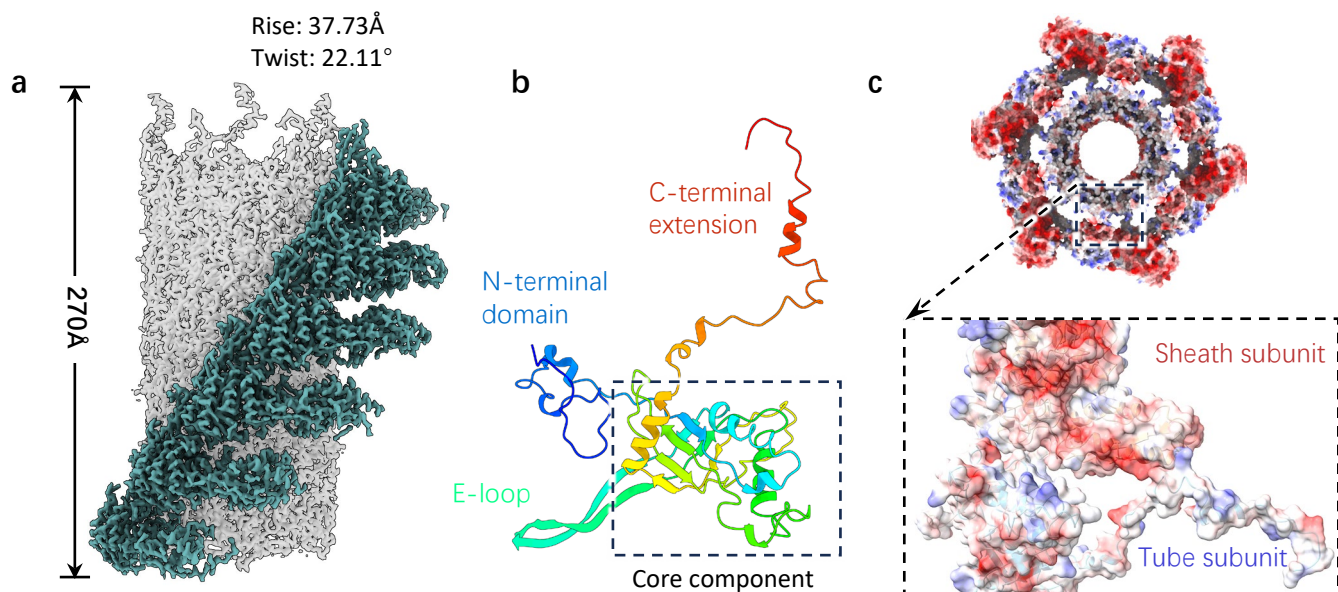


Figure S10. Structure of the tail sheath and tube. (a) Cryo-EM density map of all sheath subunits (dark cyan) within a single helical pitch, arranged around the tube ring (gray). **(b)** Ribbon models of the tail tube gp30. **(c)** The complementary negative-charged interface between the tail tube and sheath.

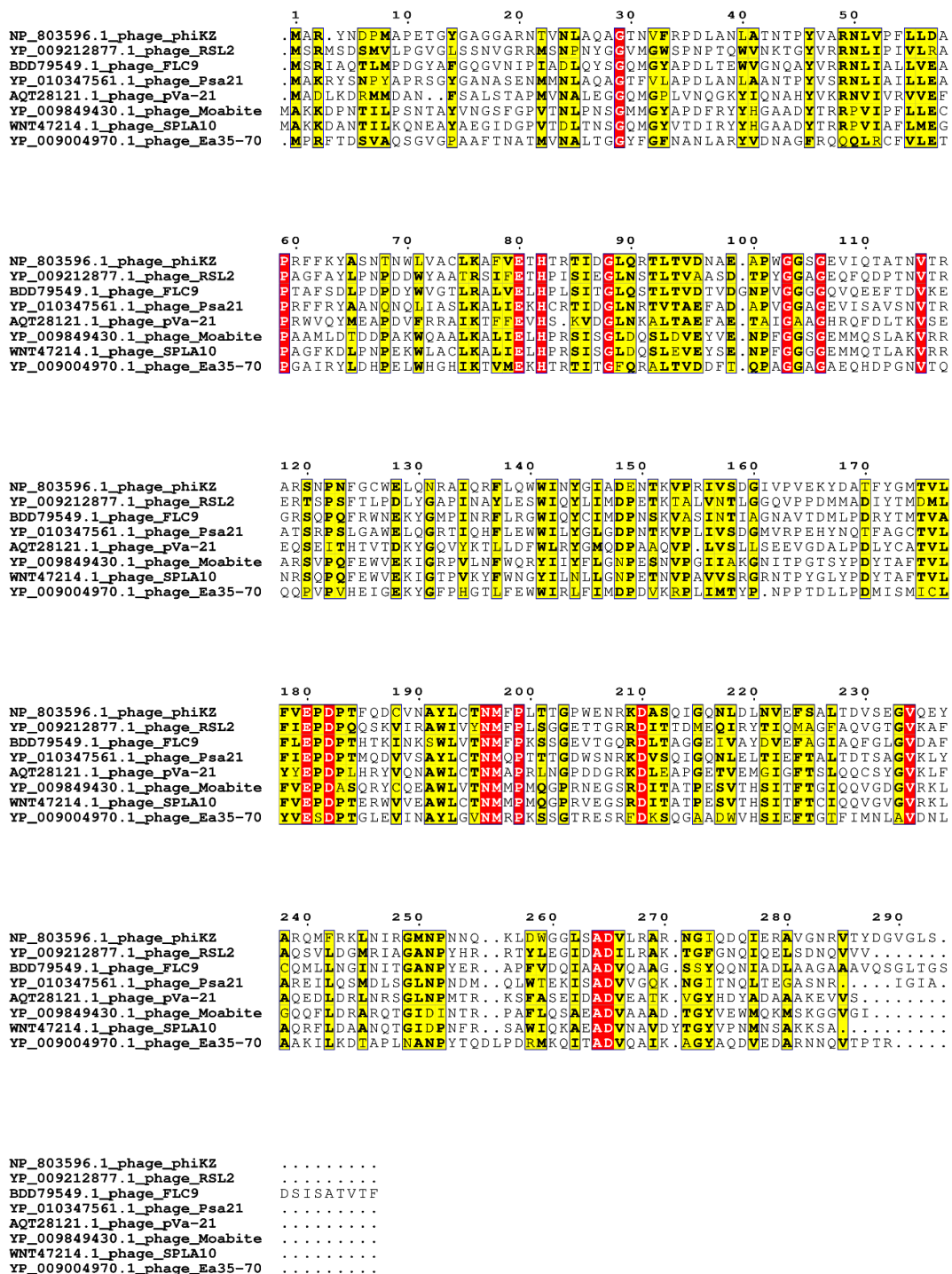


Figure S11. Multiple sequence alignment of the tail tube in jumbo phage phiKZ, RSL2, FLC9, Psa21, pVa-21, Moabite, SPLA10 and Ea35-70. Sequence numbering is based on the phiKZ genome.

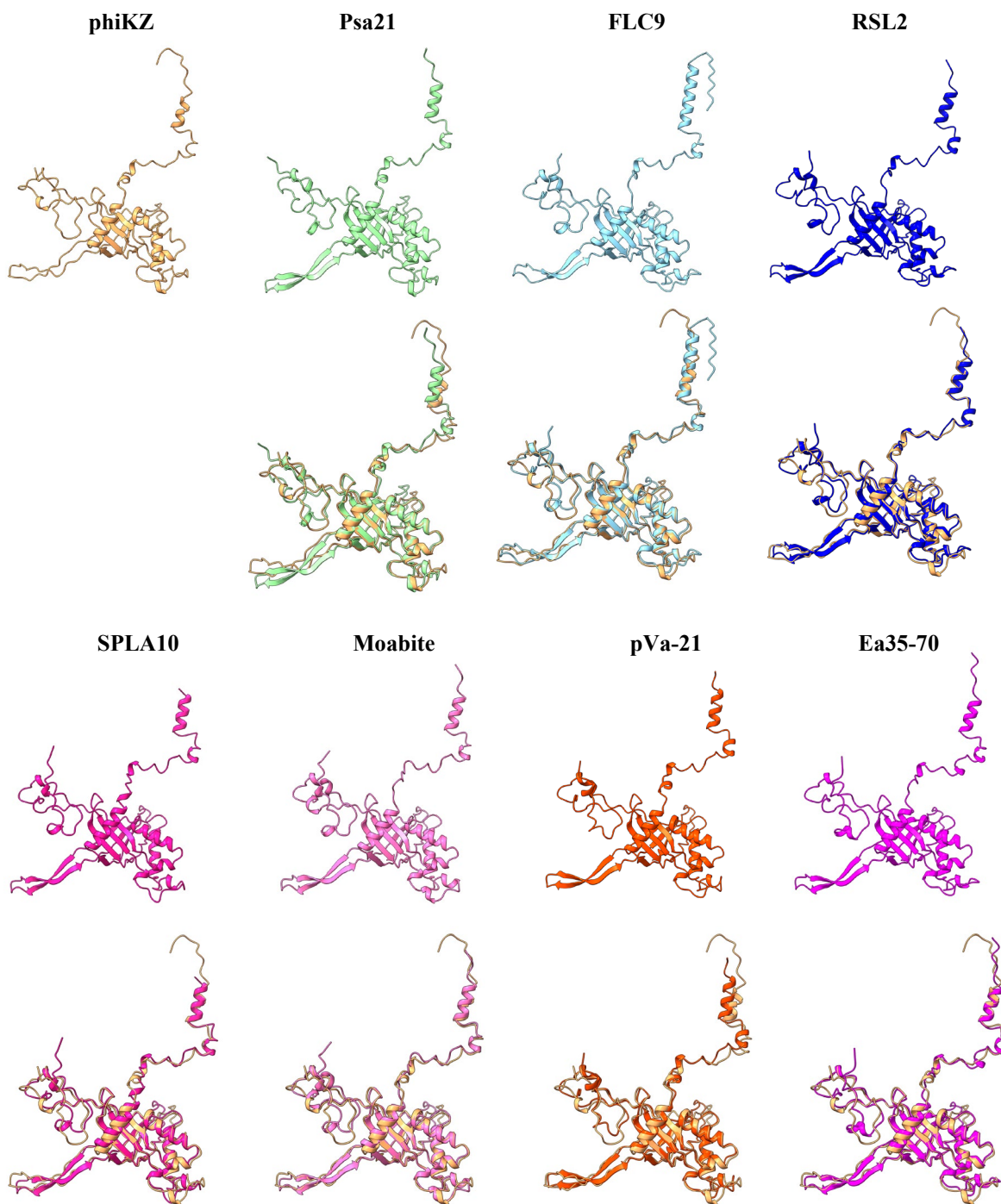


Figure S12. Structural comparison of the phiKZ tail tube protein with AlphaFold 3-predicted homologs from other jumbo phages. The experimental structure of the phiKZ tail tube subunit is compared with predicted models from phages RSL2 (34.15% sequence identity), FLC9 (35.00%), Psa21 (60.42%), pVa-21 (30.49%), Moabite (30.85%), SPLA10 (32.37%), and Ea35-70 (30.34%).

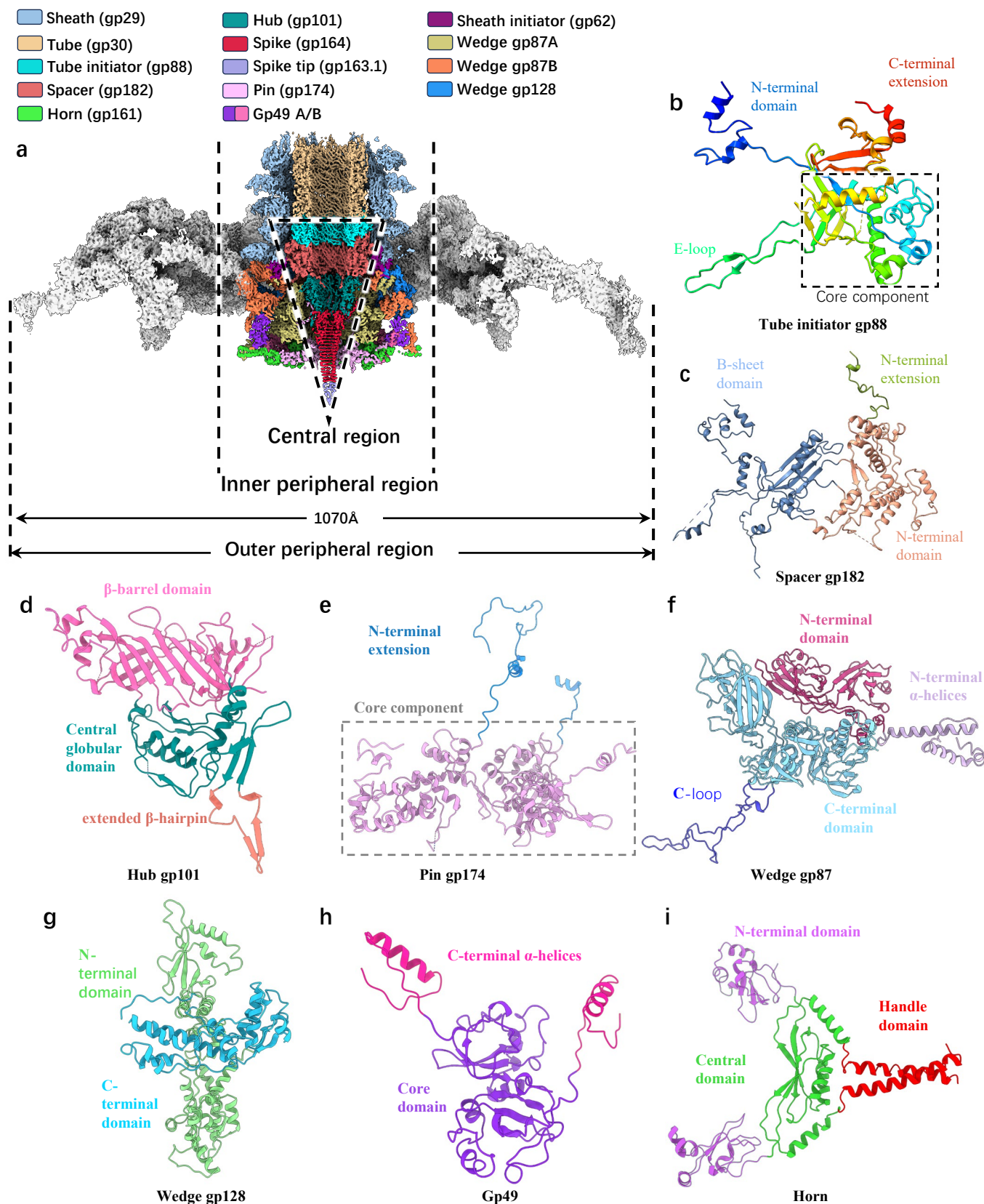


Figure S13. Ribbon models of the baseplate proteins. (a) The phiKZ baseplate complex is organized into three distinct regions. (b-i) Ribbon representations are shown for tube initiator gp88 (b), spacer gp182 (c), hub gp101 (d), pin gp174 (e), wedge gp87 (f), wedge gp128 (g), gp49 (h), horn gp161 (i). Distinct domains within each protein are labeled and individually colored.

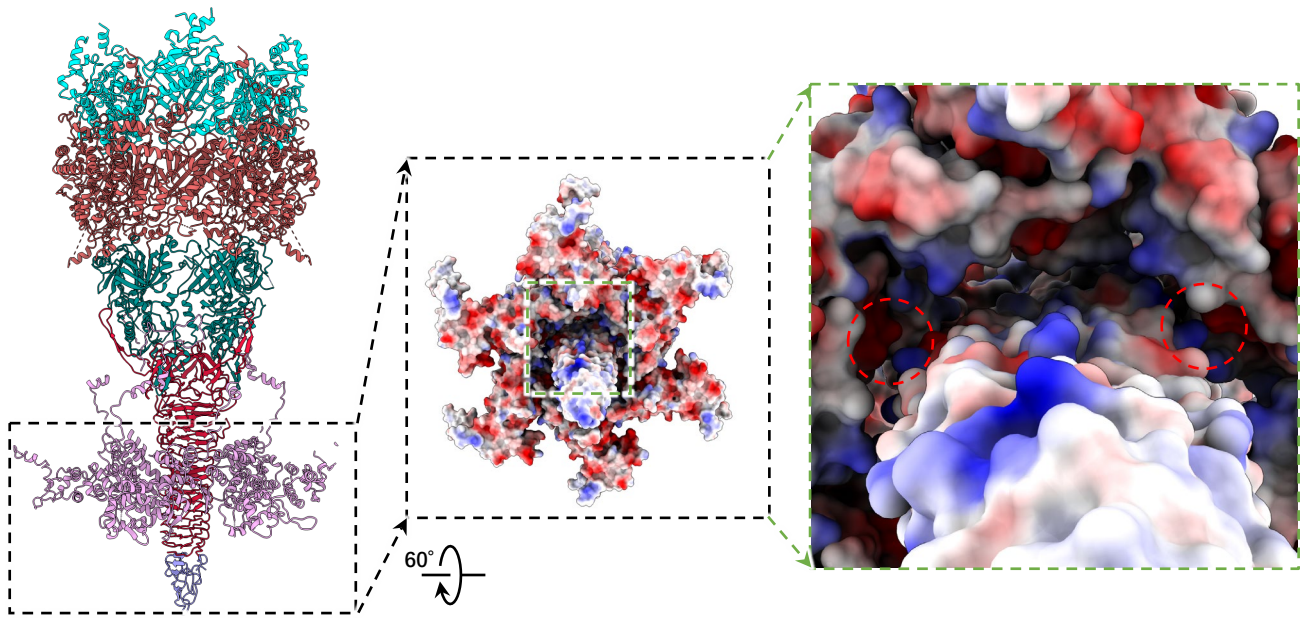


Figure S14. Electrostatic potential surfaces of pin protein gp174 and spike protein gp164. The enlarged view on the right illustrates their interaction sites.

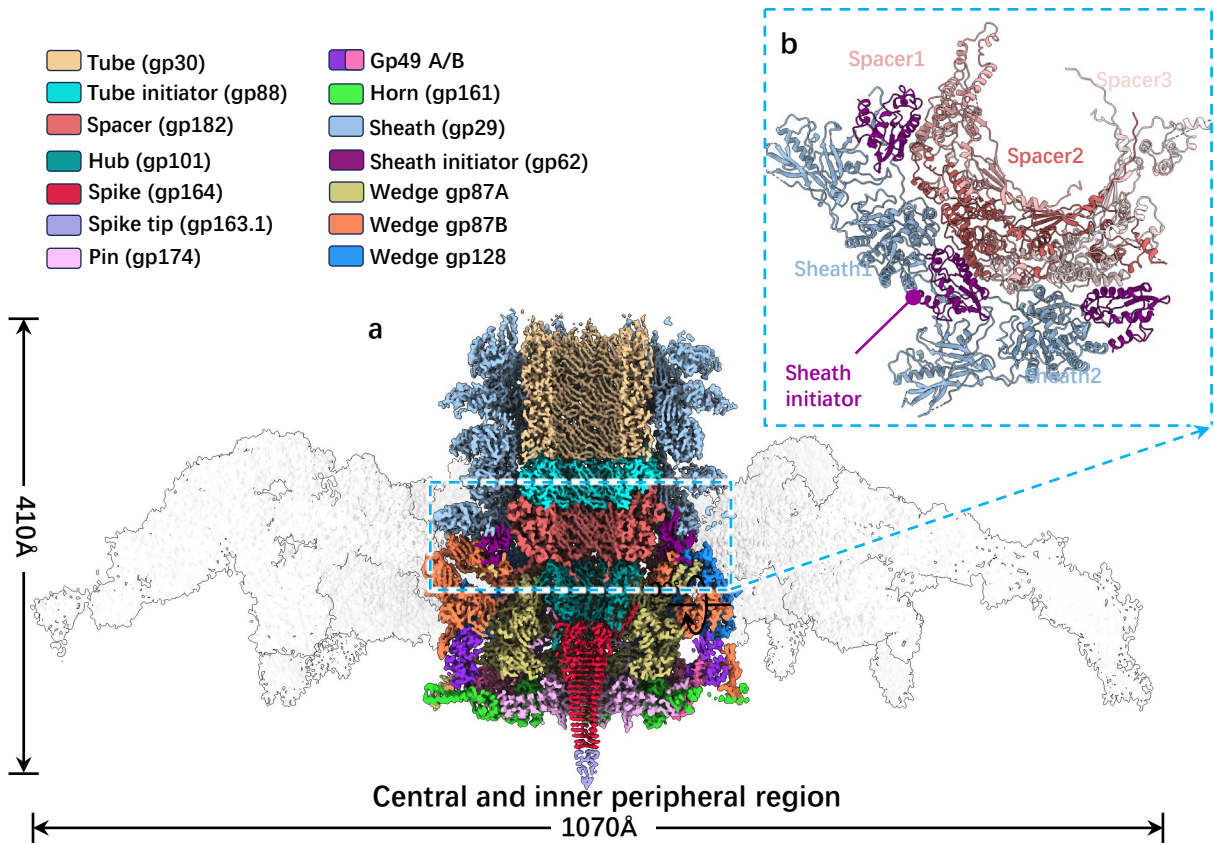


Figure S15. Structure of the central region and inner peripheral region in the baseplate complex.

(a) Cut-open view of the central and inner peripheral region in the phiKZ baseplate complex.

Individual protein names are annotated at the upper left. The outer peripheral region and fibers of the phiKZ are rendered in white. **(b)** Zoomed-in view of the interface connecting the spacer (red), the sheath initiator (purple), and the proximal tail sheath ring (blue). The sheath initiator is strategically positioned to bridge the spacer and the tail sheath assembly.

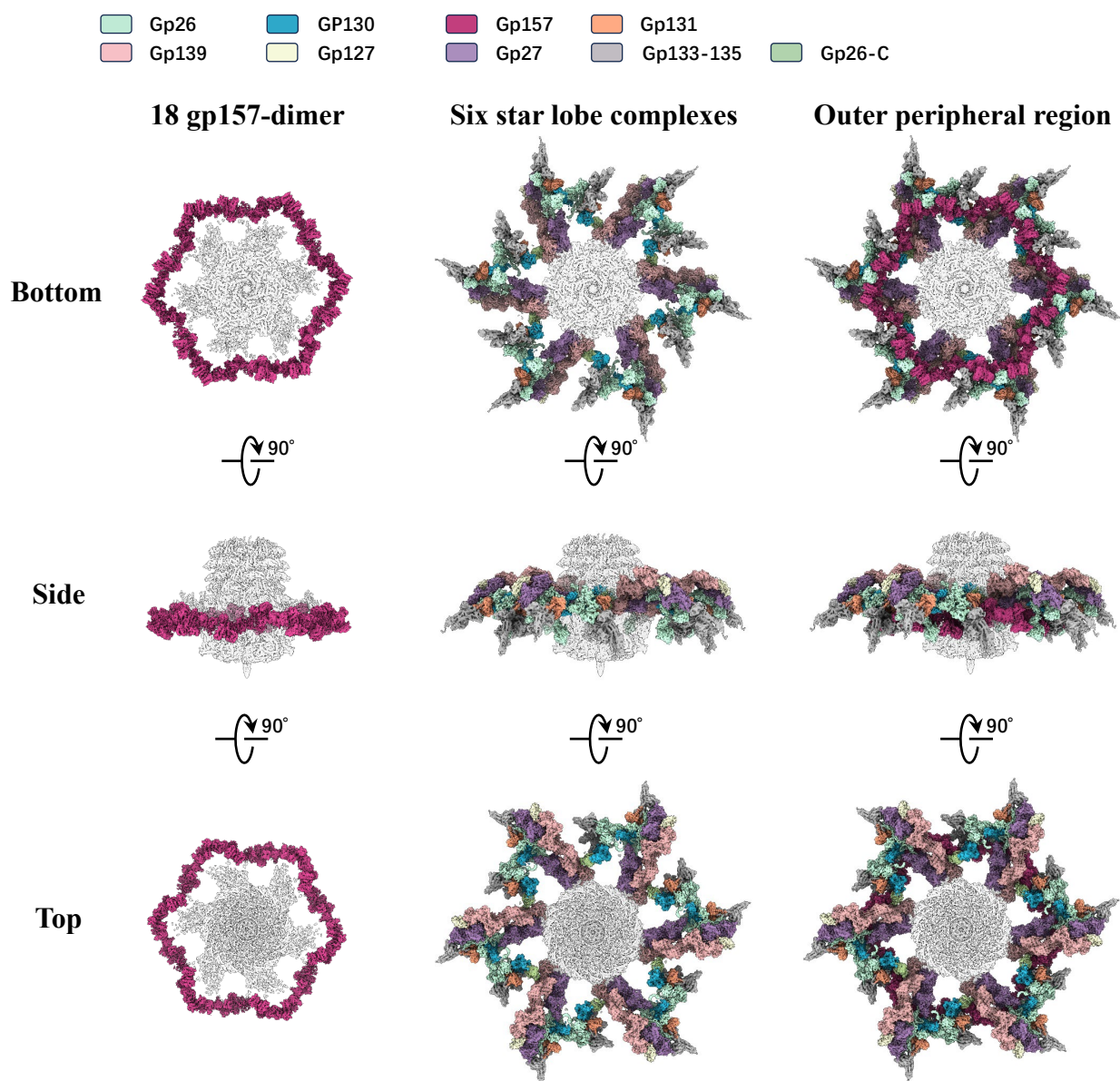


Figure S16. Spatial organization of the outer peripheral region in phiKZ baseplate complex. The outer peripheral region of the phiKZ baseplate comprises a hexagram-shaped ring of 18 gp157 dimers and six star-lobe complexes. The top, side, and bottom views of the hexagram-shaped ring, and six star-lobe complexes, and their combination are shown, respectively. The key protein gp157, which forms a continuous ring around the baseplate, is highlighted.

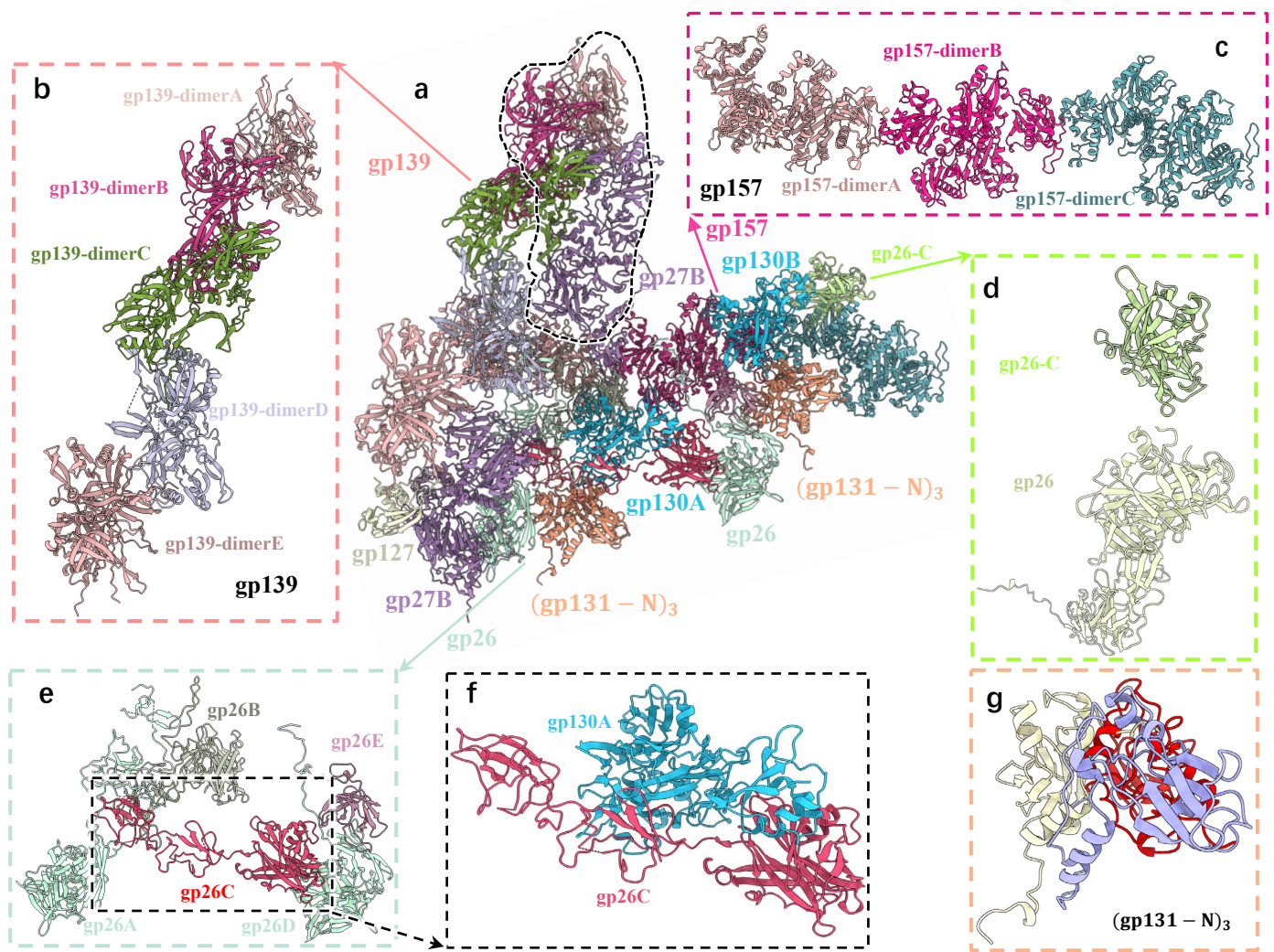


Figure S17. Structure of a single star lobe complex in outer peripheral region. Key proteins gp139, gp26, and gp157 are highlighted in different colors to illustrate their distribution. **(a)** Atomic model of a single phiKZ star lobe complex, with each protein colored distinctly. Within the dashed box are two gp139 dimers and one gp27 monomer connecting the inner peripheral region. **(b)** The assembly of five gp139 dimers. Gp139 dimers A and B connect the inner peripheral region, as shown in panel a. **(c)** The arrangement of three gp157 dimers. **(d)** The C-terminal domain of gp26 fits into the connecting region between two adjacent units within star lobes. **(e)** The assembly of five gp26 molecules. **(f)** Gp26C and gp130 form an interlocking interface that bridges the unit1 and unit2 of star lobe. **(g)** The region formed by the association of three gp131 N-termini.

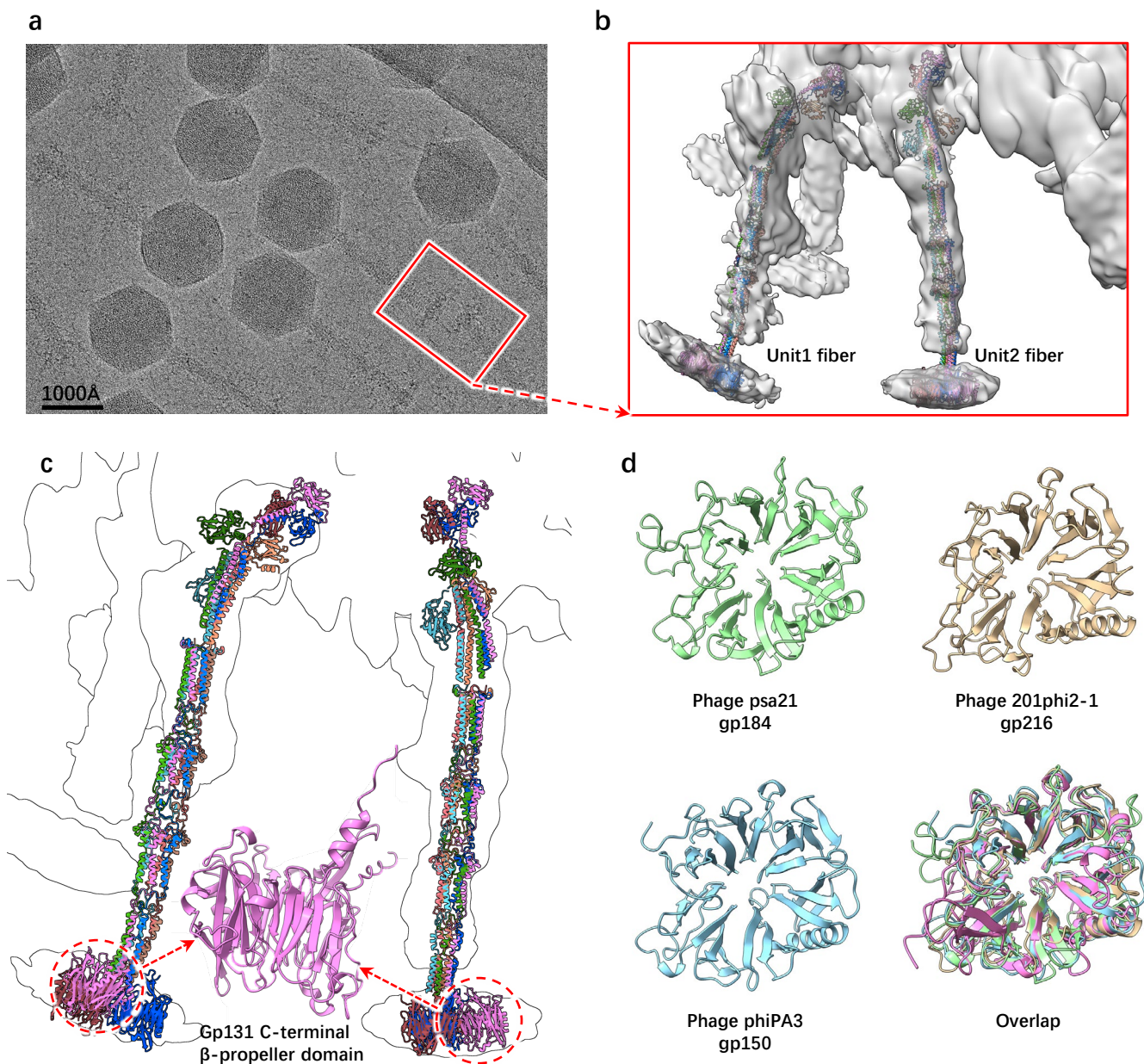


Figure S18. Architectural features of the phiKZ tail fiber. (a) Representative cryo-EM micrographs of phiKZ show the baseplate; both tail fibers protrude prominently from it. (b) AlphaFold3-predicted models of the unit1 and unit2 fibers were fitted into the C6-symmetrized, low-pass filtered cryo-EM density of the fiber complex. (c) Density map blurred to show fiber hexamer binding mode; low resolution at the fiber distal end precludes definitive binding mode assignment, but the seven-bladed β -propeller C-terminus of the gp131 trimer is fitable to this region. Center (pink): AlphaFold3-predicted seven-bladed β -propeller C-terminus of gp131. (d) Structural comparison of the C-terminal domain of phiKZ gp131 with AlphaFold3-predicted homologs from other jumbo phages, including putative tail fiber or tailspike proteins: gp184 from phage psa21, gp216 from phage 201phi2-1, and gp150 from phage phiPA3. The bottom right panel shows structural overlap of all four models.

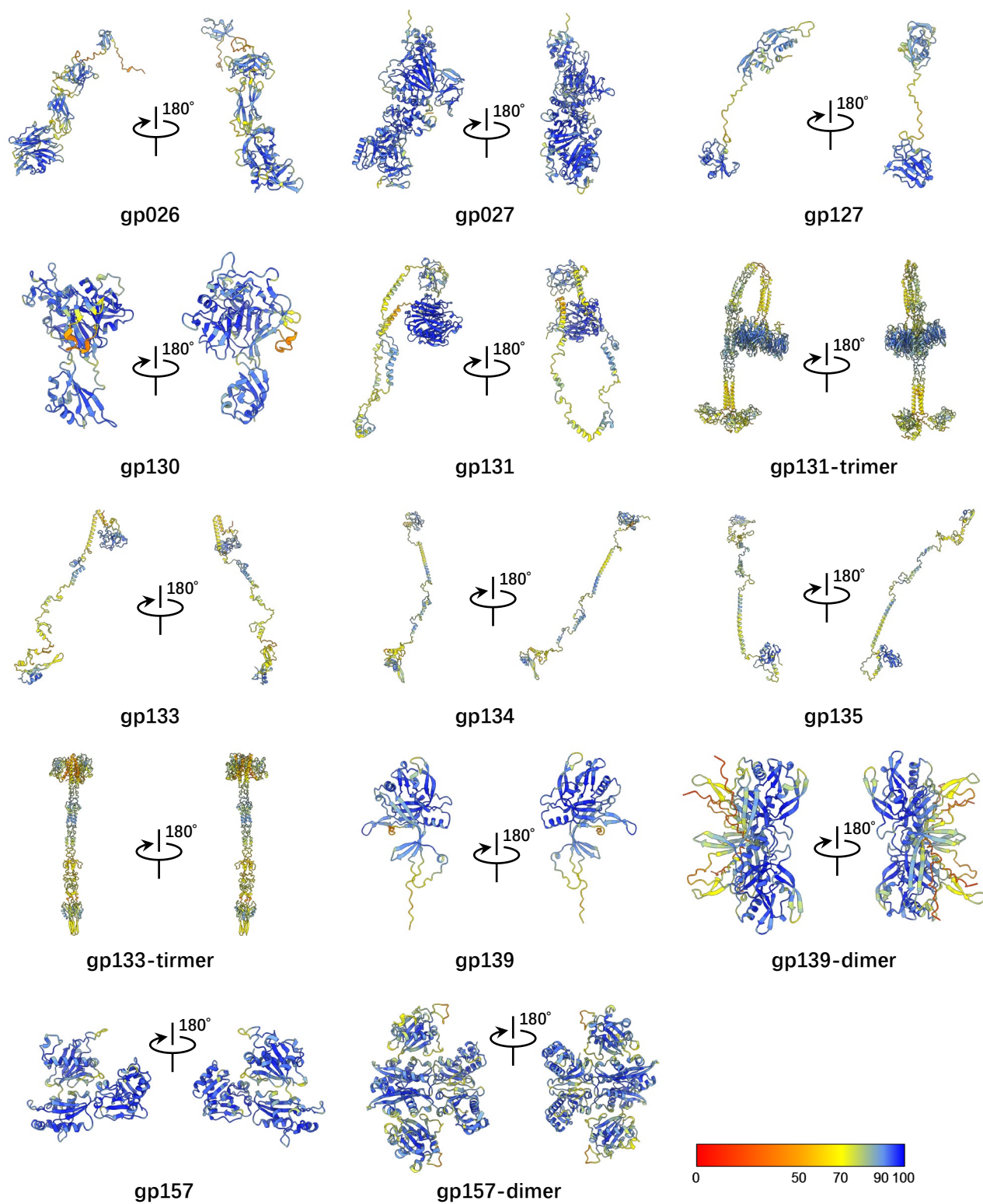


Figure S19. Predicted model confidence (pLDDT/pTM). All AlphaFold3-predicted atomic models were colored by per-residue pLDDT scores in UCSF ChimeraX.

Table S1. Virion proteins of phiKZ

Gene	Protein name	aa	Modeled residues	Interacts with	Copies per phage	symmetry
26	unknownA	552			6	C6
	unknownB				6	
	unknownC				6	
	unknownD				6	
	unknownE				6	
27	unknownA	898			6	C6
	unknownB				6	
28	inner capsid complex	320				I
29	sheath	695	3-694		264	C6
30	tube	293	3-292		258	C6
34	whisker	256	4-162,176-255	99,144	48	C6
35	minor capsid protein	271				I
42	adaptor	266	1-103,109-155,164-266	34,99,129	12	C12
49	otherA	138	3-137	49B,87,128,161	12	C6
	otherB		3-137	49A,87A,161	12	
62	sheath initiator	136	1-135	29,87,128,182	6	C6
85	inner capsid complex	149				I
86	inner capsid complex	422				I
87	wedge2A	971	3-967	49,62,87B,128,161,174	6	C6
	wedge2B		3-968	29,49A,62,87A,128,161	6	
88	tube initiator	349	6-340	29,30,182	6	C6
91	inner capsid complex	176				I
93	fiber-like protein	435				I
98	terminator	542	2-540	30,99,144	6	C6
99	stopperA	469	32-468	34,42,98,99B,144	6	C6
	stopperB		19-468	34,42,98,99A,144	6	
101	hub	460	5-452		3	C3
119	inner capsid complex	177				I
120	major capsid protein	747				I
127	unknown	290			6	C6
128	wedge1	724	4-615	62,87,139	6	C6
129	portal	896	50-169,265-339,360-813	42	12	C12
130	unknownA	427			6	C6
	unknownB				6	
131	unknown	771			36	C6
133	unknown	462			12	C6
134	unknown	457			12	C6
135	unknown	463			12	C6
139	dockerA	298			12	C6
	dockerB				12	
	dockerC				12	
	dockerD				12	
	dockerE				12	
144	mat	260	14-116,132-211,243-257	29,34,98	6	C6
157	unknownA	445			12	C6
	unknownB				12	
	unknownC				12	
161	horn	203	2-202	49,87,174	12	C6
162I	inner capsid complex	522				I
162II	fiber-like protein	522				I
163.1	spike tip	88	2-88	164	1	C1
164	spike	293	18-293	163.1,174	3	C3
174	pin	354	4-353	87A,161,164	6	C3
182	spacer	664	6-661	29,62,87A,88,101	6	C6
184	inner capsid complex	230				I
244	minor capsid protein	118				I

Note: Green indicates proteins in the Outer Peripheral Region predicted by AlphaFold3; orange indicates capsid protein; black indicates proteins de novo modeled based on our high-resolution map.

Table S2. Refinement and model statistics.

	tail-all	unit1	unit2	portal- adaptor	neck	sheath- tube	baseplate C6	baseplate C3	docker	outer peripheral
	(EMDB- 67125)	(EMDB- 67123)	(EMDB- 67124)	(EMDB- 67095) (PDB 9XPE)	(EMDB- 67102) (PDB 9XPS)	(EMDB- 67096) (PDB 9XPF)	(EMDB- 67121) (PDB 9XQS)	(EMDB- 67098) (PDB 9XPH)	(EMDB- 67097) (PDB 9XPG)	(EMDB- 67111) (PDB 9XQD)
Data collection and processing										
Magnification						53,000				
Voltage (kV)						300				
Electron exposure (e- /Å²)						30				
Defocus range (µm)						-1.2 ~ -2.2				
Pixel size (Å)						1.36				
Symmetry imposed	C6	C1	C1	C12	C6	C6	C6	C3	C1	C1
Final particle images (no.)	8,843	134,673	134,673	31,906	32,544	26,028	22,539	27,796	135,070	131,125
Map resolution (Å)	7.41	4.39	4.55	3.37	3.52	3.15	3.65	4.02	4.14	4.53
FSC threshold	0.143	0.143	0.143	0.143	0.143	0.143	0.143	0.143	0.143	0.143
Refinement										
Map sharpening <i>B</i> factor (Å²)	\	\	\	90	73	111	77	79	101	118
Model composition										
Non-hydrogen atoms	\	\	\	82,512	37,859	89,568	91,840	32,802	8,267	71,613
Protein residues	\	\	\	10,248	4,648	11,376	11,520	4,179	1,068	8,914
Ligands	\	\	\	\	\	\	\	\	\	\
<i>B</i> factors (Å²)										
Protein	\	\	\	24	58	31	53	74	77	91
Ligand	\	\	\	\	\	\	\	\	\	\
R.m.s. deviations										
Bond lengths (Å)	\	\	\	0.003	0.004	0.004	0.004	0.002	0.004	0.003
Bond angles (°)	\	\	\	0.509	0.563	0.497	0.587	0.579	0.683	0.631
Validation										
MolProbity score	\	\	\	1.51	2.37	1.66	2.12	1.98	2.19	2.08
Clashscore	\	\	\	2.55	5.15	2.11	5.91	6.68	12.39	12.14
Poor rotamers (%)	\	\	\	1.43	4.8	2..1	2.66	0	0	0.03
Ramachandran plot										
Favored (%)	\	\	\	97.13	93.06	94.79	92.34	91.81	88.48	92.23
Allowed (%)	\	\	\	2.87	6.67	5.08	7.45	8.09	11.05	7.36
Disallowed (%)	\	\	\	0	0.27	0.12	0.22	0.1	0.48	0.42