

Supplementary Materials for  
**High-resolution cryo-EM structure of the *Shigella* virus Sf6 genome delivery  
tail machine**

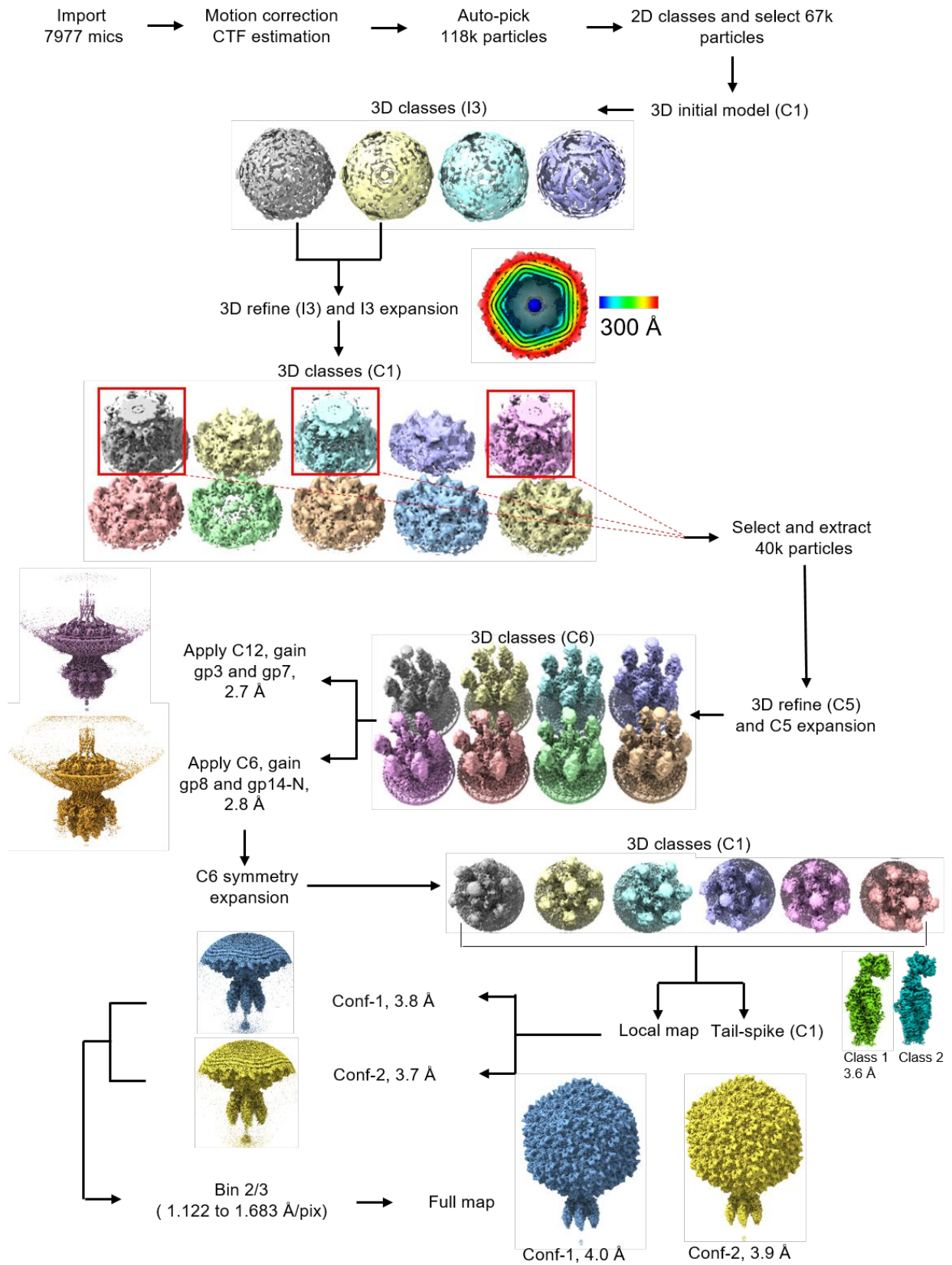
Fenglin Li *et al.*

Corresponding author: Gino Cingolani, gino.cingolani@jefferson.edu

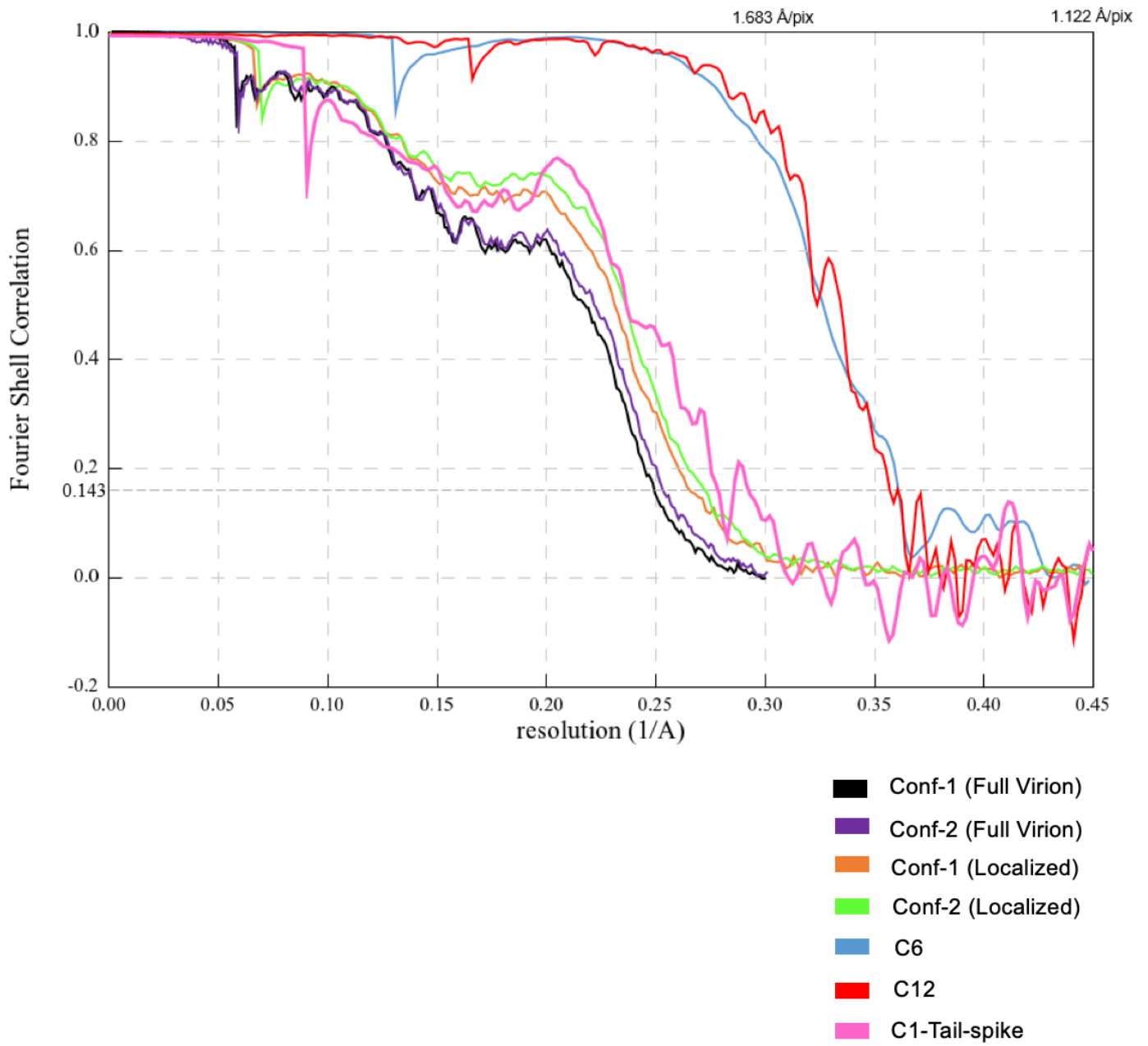
*Sci. Adv.* **8**, eadc9641 (2022)  
DOI: 10.1126/sciadv.adc9641

**This PDF file includes:**

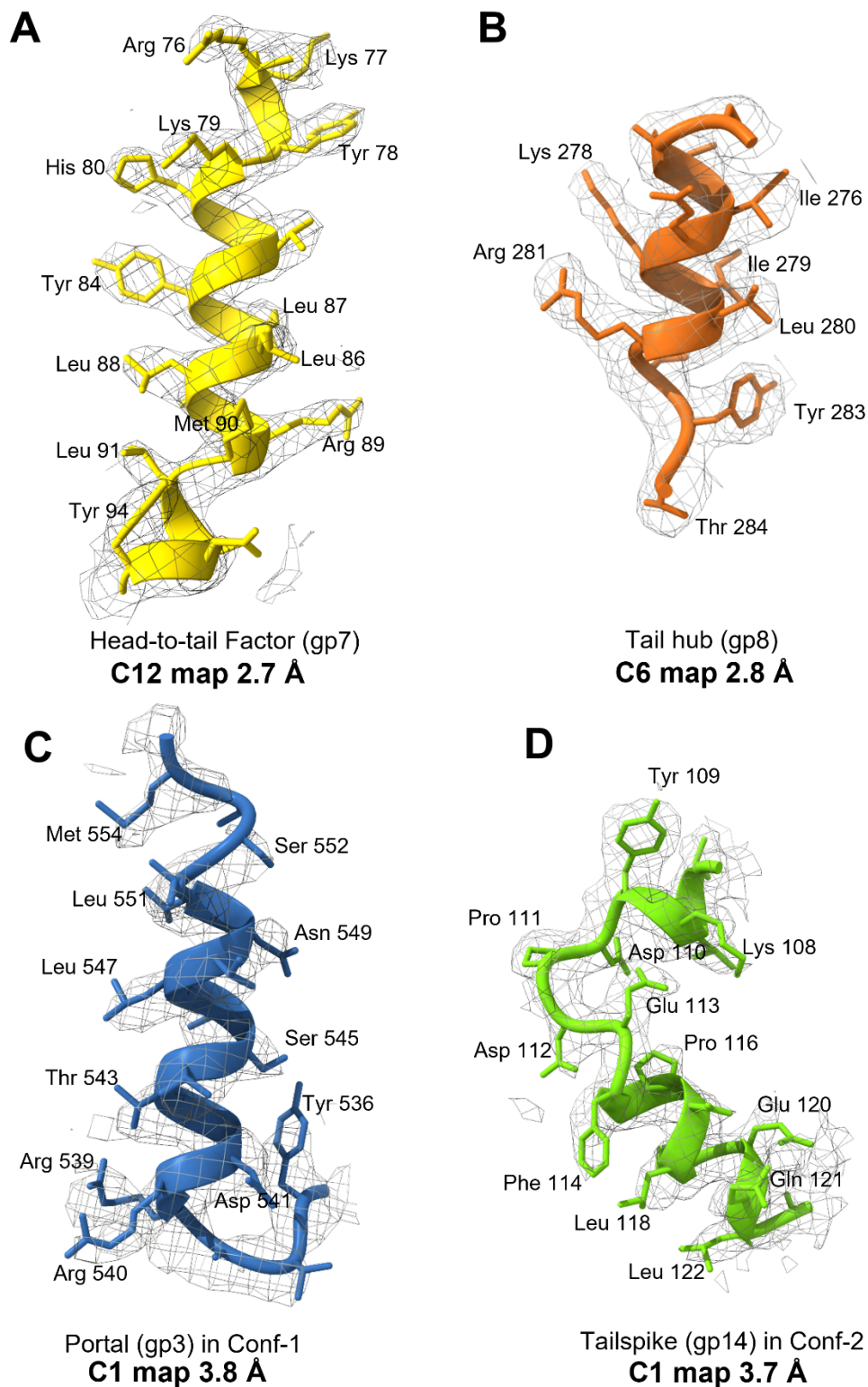
Figs. S1 to S13  
Tables S1 and S2



**Figure S1. Flowchart for cryo-EM data processing and structure determination.** The resolution of all reconstructions in this paper was estimated at a Fourier Shell Correlation FSC=0.143.

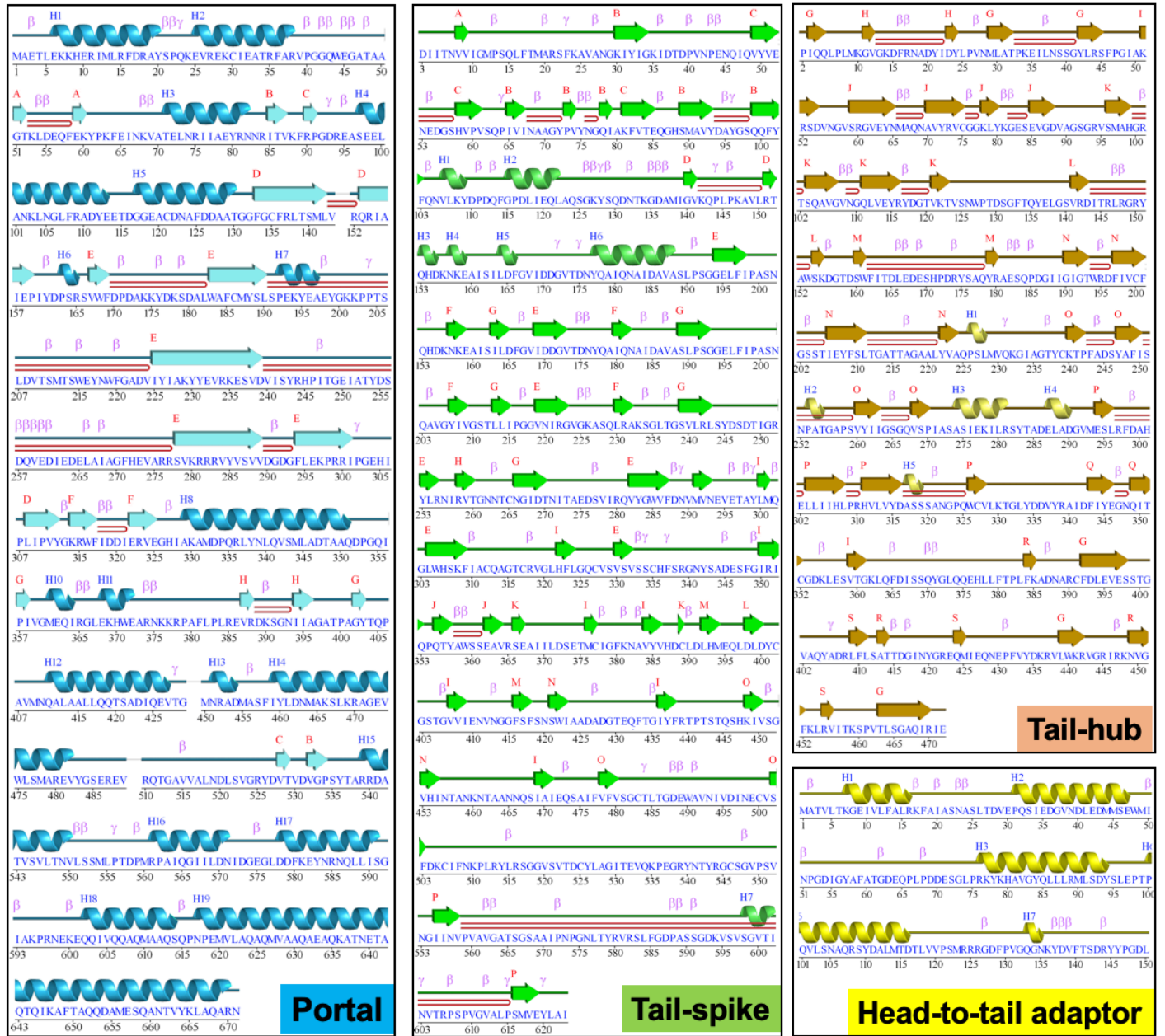


**Figure S2. Fourier Shell Correlation (FSC) curves for all cryo-EM reconstructions presented in this paper.**

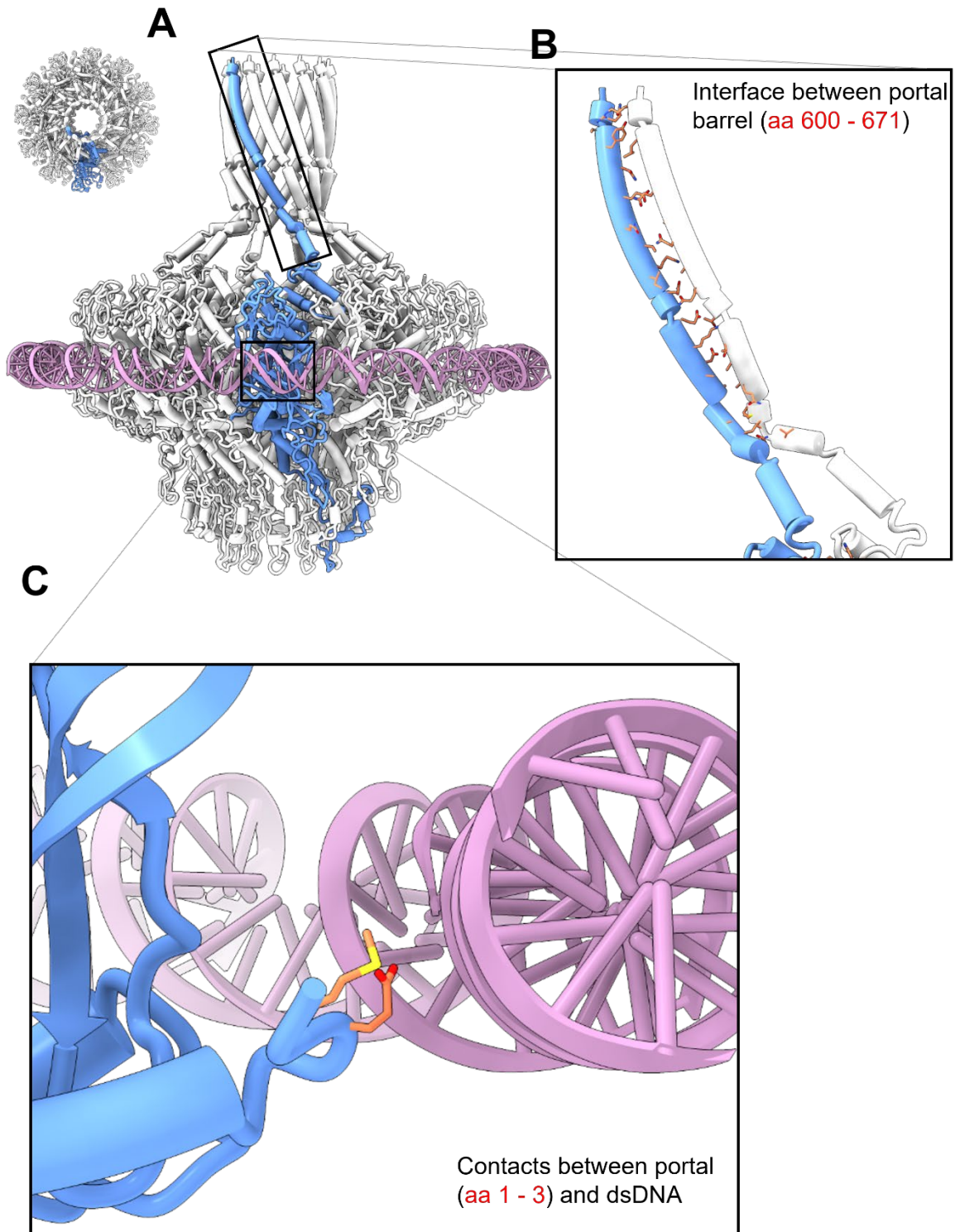


**Figure S3. Cryo-EM map quality.** Representative regions of the C12 (**A**) and C6 (**B**) symmetrized localized reconstructions contoured at  $4.5 \sigma$  and overlaid to the refined models. (**C**) Representative cryo-EM maps of the asymmetric C1 reconstructions overlaid to portal protein in Conf-1 (*left*) and tailspike in Conf-2 (*right*).

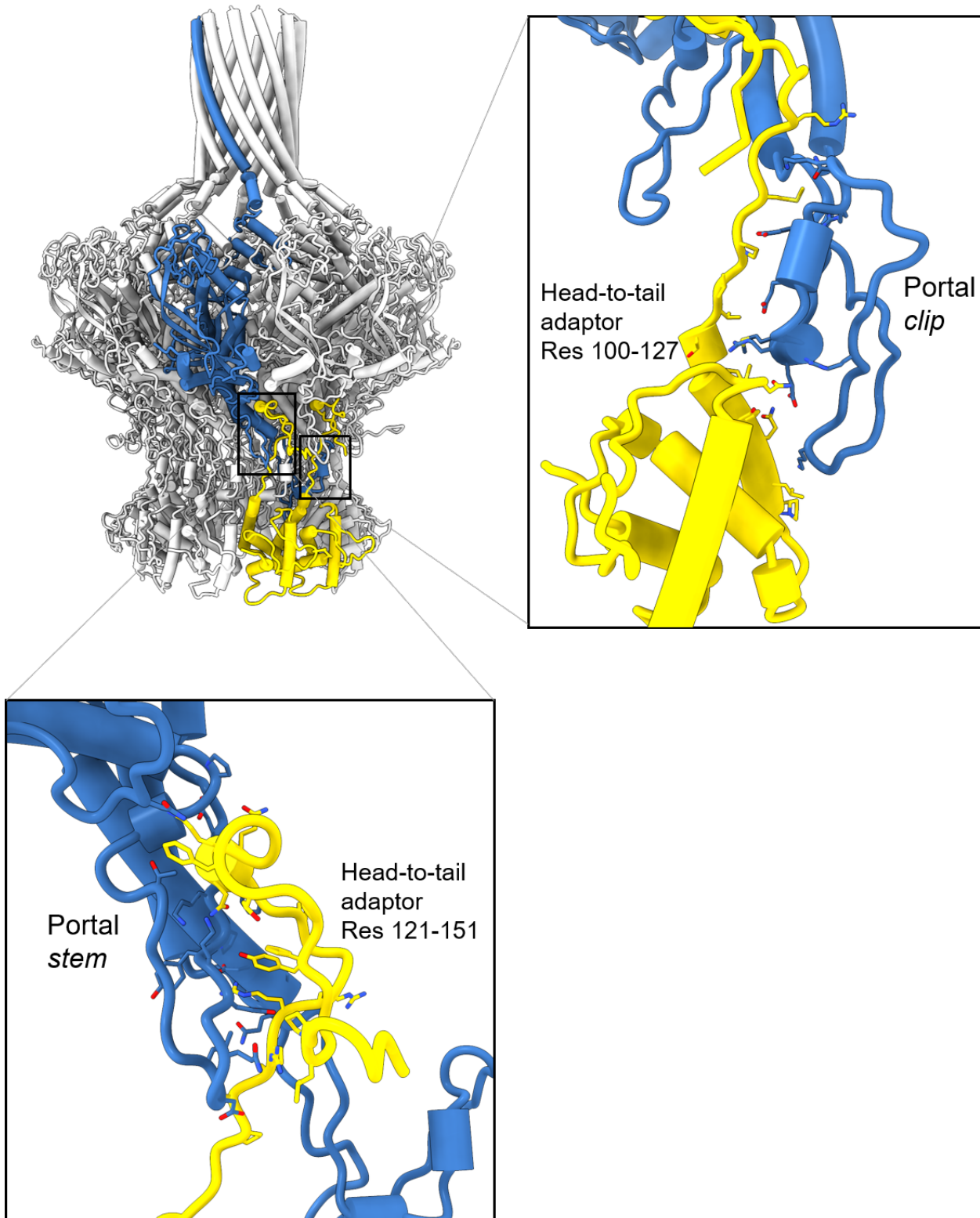




**Figure S4. Secondary structure elements and amino acid sequences of all Sf6 tail factors built *de novo* in this study. The diagrams were generated using PDBsum.**

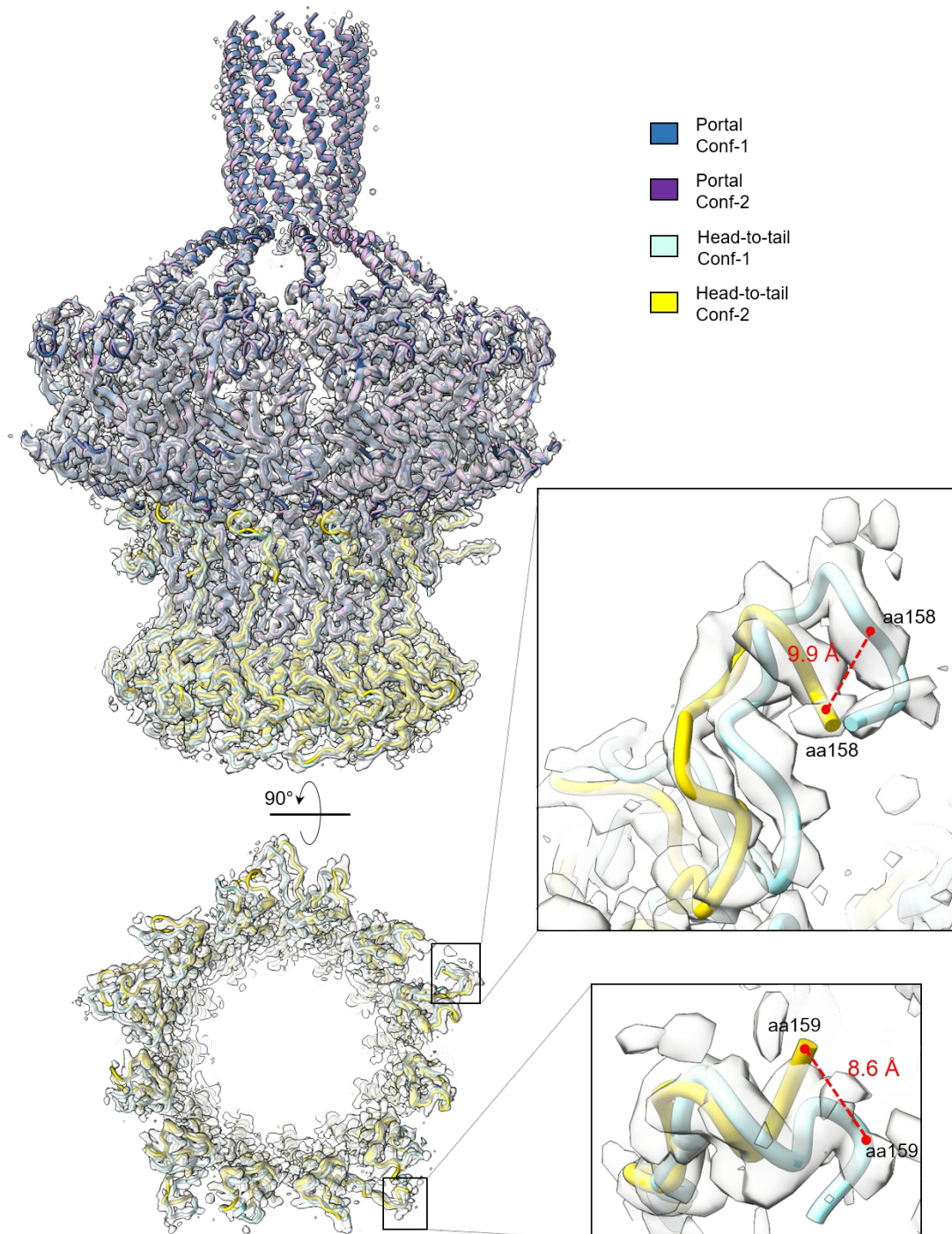


**Figure S5. Sf6 portal protein.** (A) Ribbon diagram of the Sf6 portal protein showing the ring of DNA seen in the reconstruction (colored in purple) around the portal perimeter. Only one portal protomer is colored in blue, while the other 11 protomers are in light gray. (B) Magnified view of the interactions between two barrel helices. Sidechains in 3.5 Å bonding distance are colored in orange. (C) Magnified view of the putative interaction between DNA and the portal wing with residues Met1 and Glu3 colored in orange.

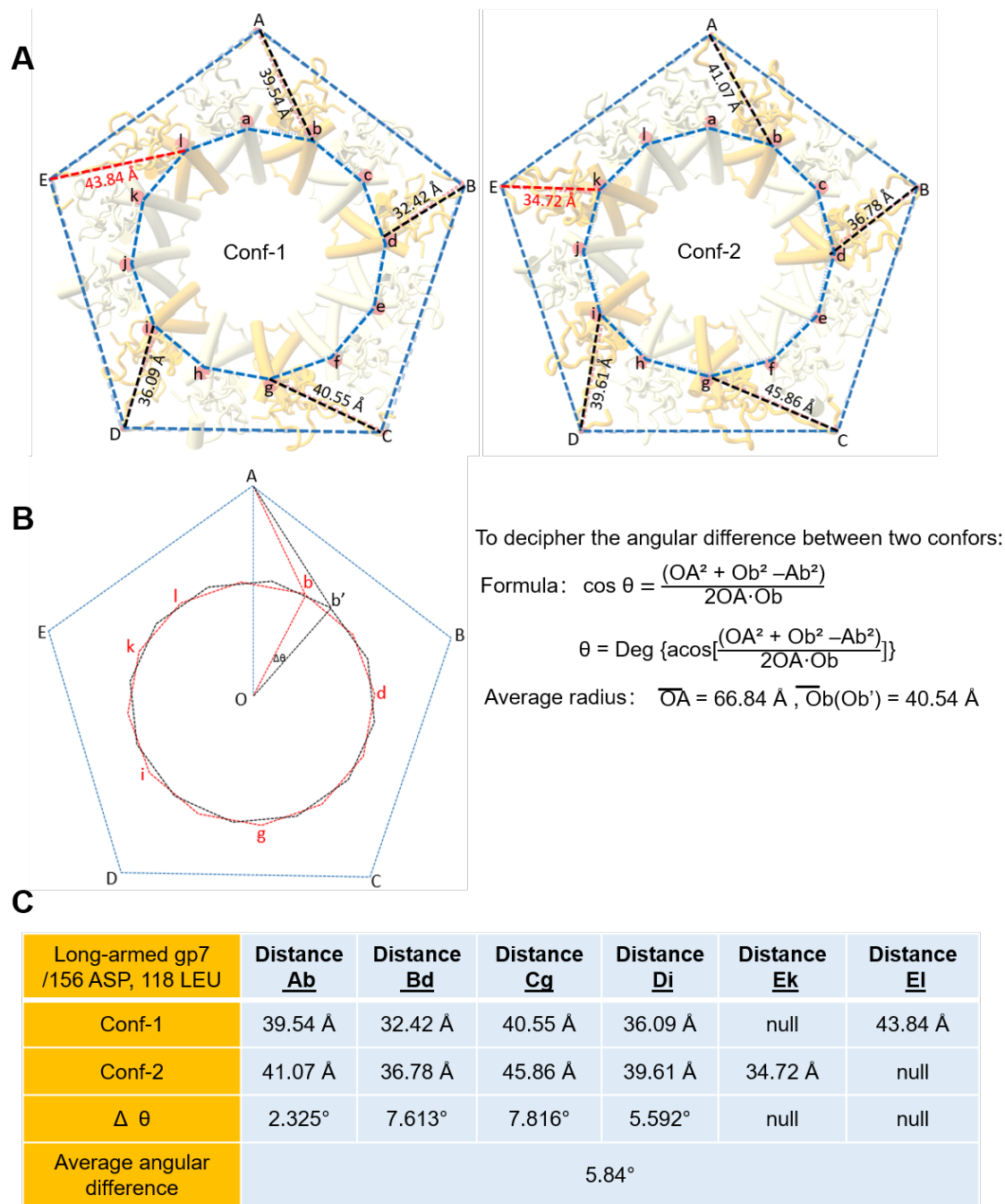


**Figure S6. Sf6 portal protein:head-to-tail complex.** Contact points between the Sf6 portal protein (blue) and the head-to-tail factor gp7 (yellow) within a 3.5 Å bonding distance.

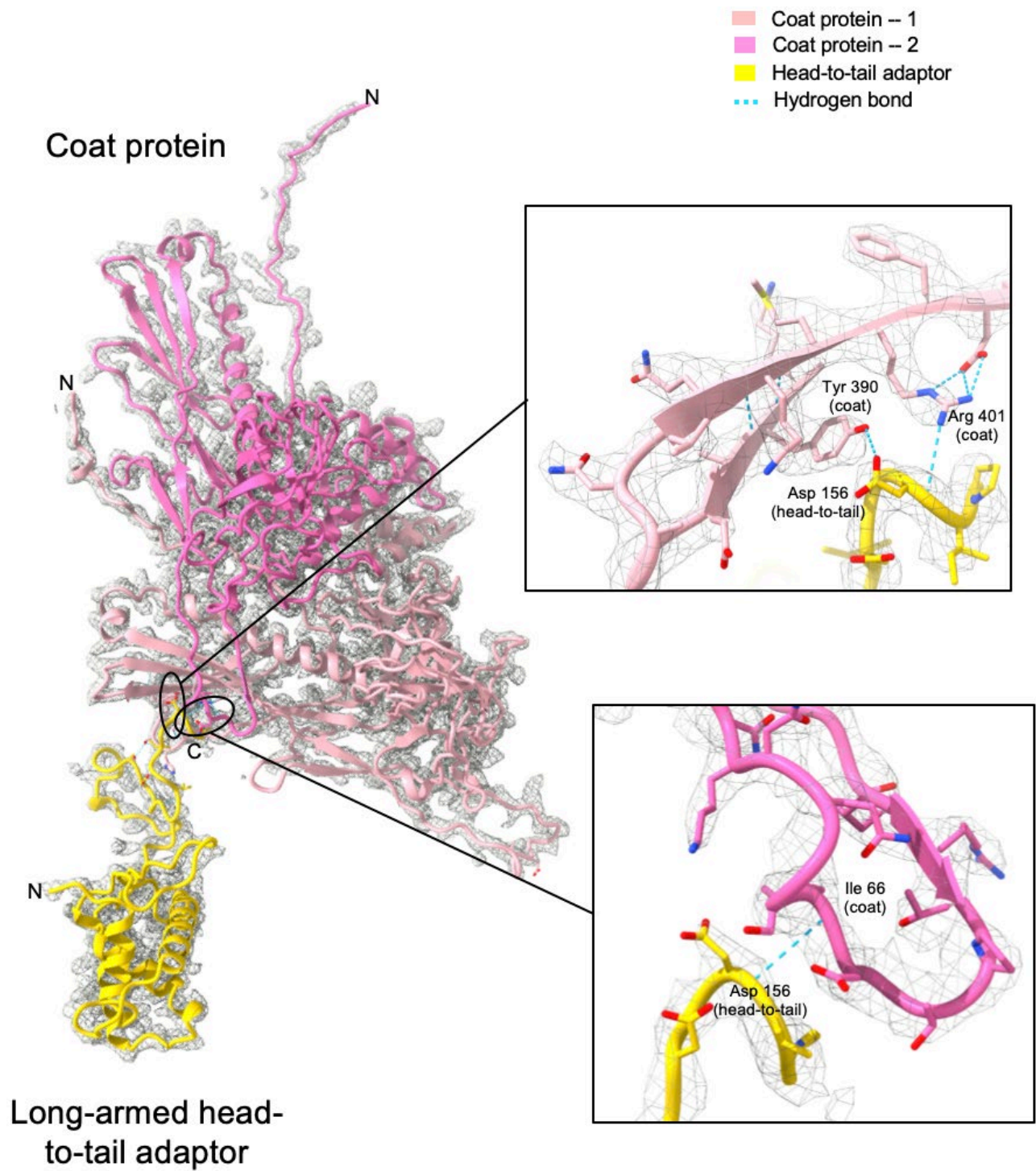




**Figure S7. Overlay of the cryo-EM maps of Conf-1 (yellow) and Conf-2 (cyan) rotated by 144°.** The maps are contoured at  $3.4\sigma$ . If the two maps were identical (but had different origins), rotating Conf-2 by 144° should result in a map identical to Conf-1 (at least in the C12-symmetric portion of this map that includes portal protein and gp7). Instead, the overlay shows that two maps display significant structural differences in the head-to-tail factors, with a maximum displacement of up to ~10 Å (see zoom-in panels), while the portal proteins are nearly identical. The overall cross-correlation coefficient between the two maps is 0.87. The RMSD between head-to-tail atomic models refined against Conf-1 and Conf-2 is RMSD = 2.3 Å.

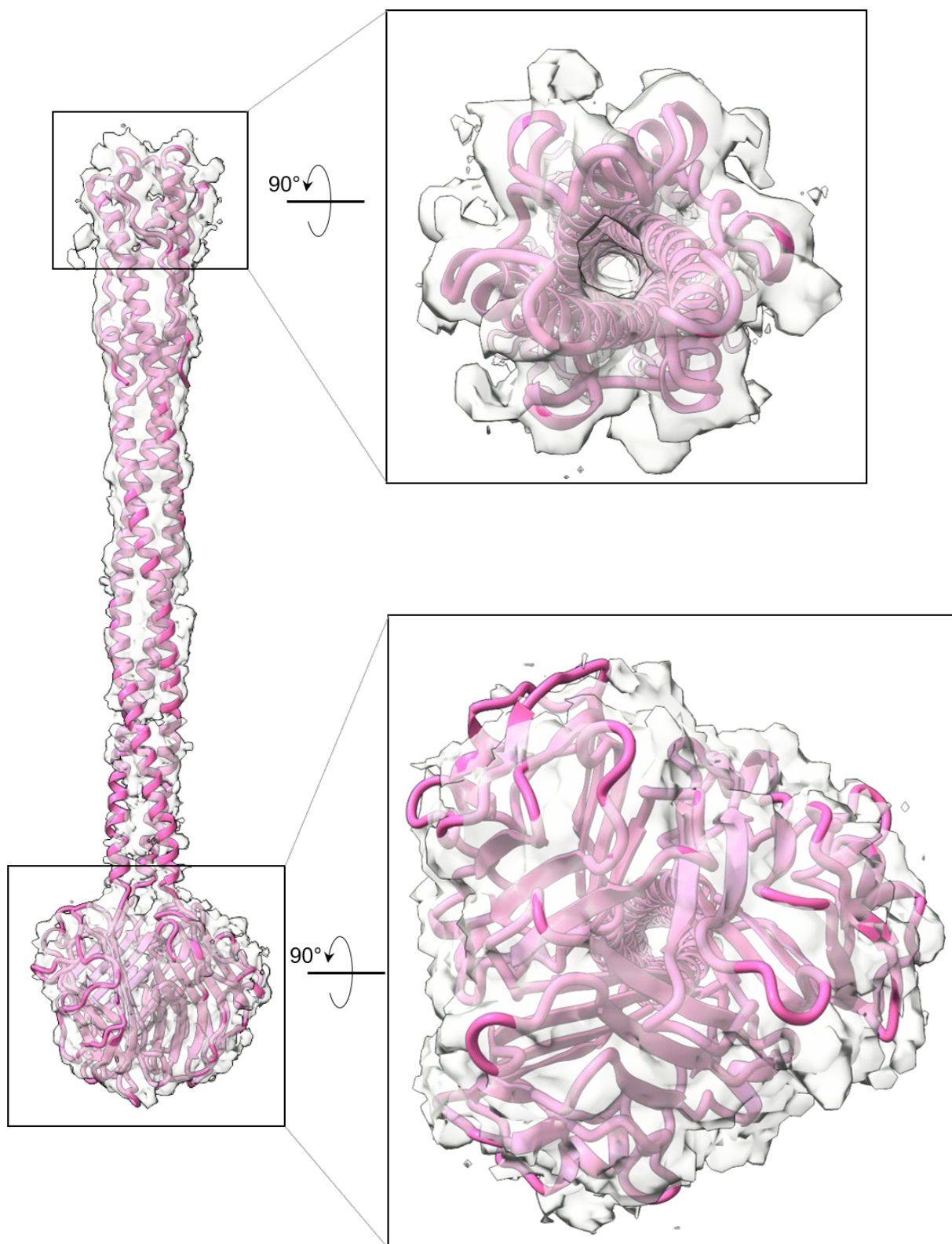


**Figure S8. Estimation of the angular differences between Conf-1 and Conf-2.** The angular differences between two head-to-tail conformations can be estimated trigonometrically by calculating the cosine of the  $\theta$  angle (see panel B). **(A)** The capsid five-fold vertexes are labeled **A** to **E** (see pentagon 'ABCDE'), and gp7 protomers are **a** to **l** (see dodecagon 'a-l'). The atomic distances between the pentagon vertexes and five fixed points on the long-armed head-to-tail adaptors (measured using ChimeraX) are shown. **(B)** Overlay of two gp7 dodecagons (conf-1 = **a** - **l**; conf-2 = **a'** - **l'**) in the pentagon ABCDE with the same origin **O**. The angle between the two octagons ( $\theta$ ) can be calculated trigonometrically:  $\theta = \text{Deg} \left\{ \arccos \left[ \frac{(OA^2 + Ob^2 - Ab^2)}{2OA \cdot Ob} \right] \right\}$ . The angular difference between two dodecagons is  $\Delta\theta$ . Assuming the average radius of the pentagon is  $OA$  and for two dodecagons is  $Ob$  (or  $Ob'$ ), then the angular differences ( $\Delta\theta$ ) can be calculated, as shown in panel **(C)**. The average angular difference between two conformations is about  $6^\circ$ .

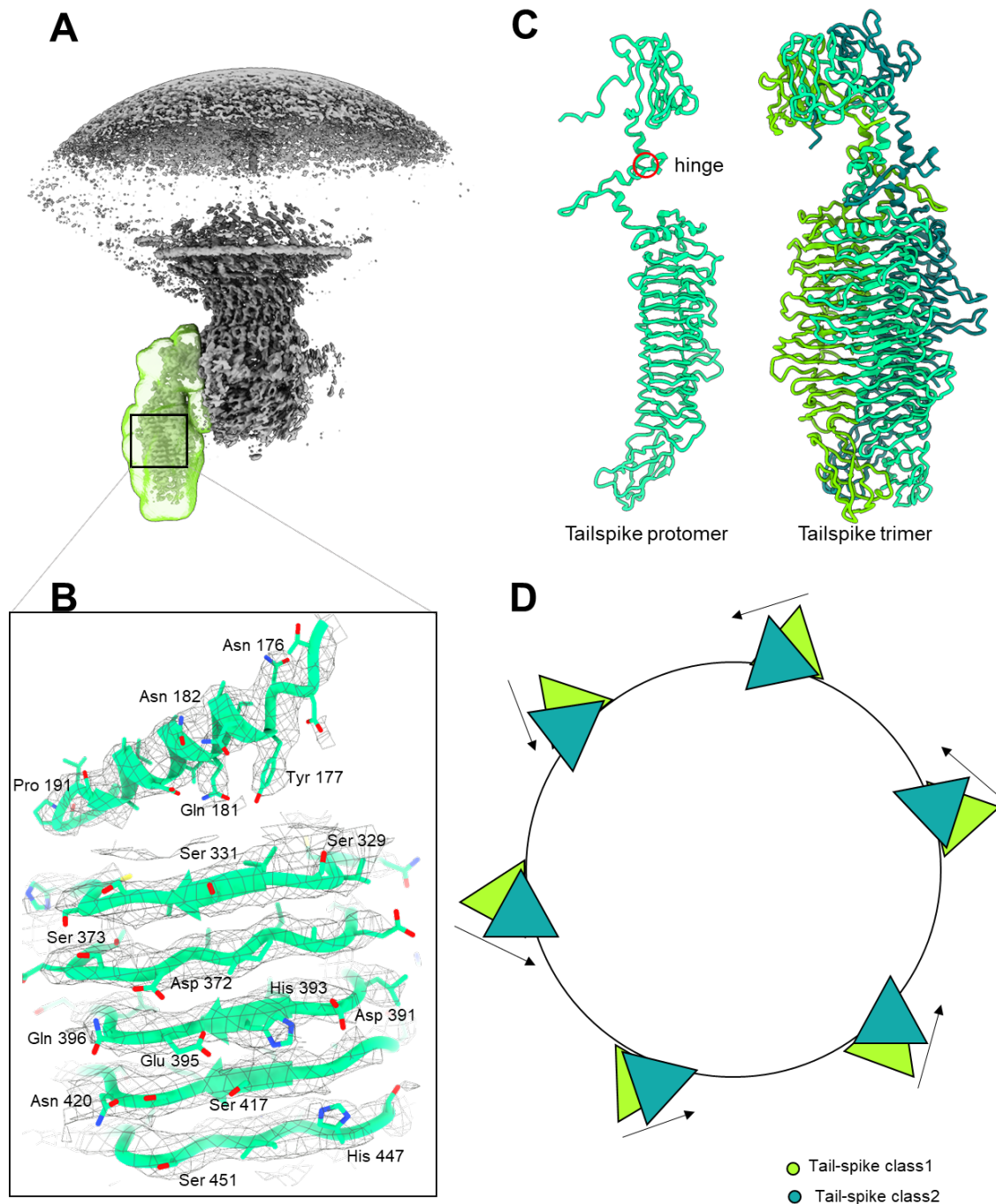


**Figure S9. Interaction between a representative long-armed tail-to-head adaptor (yellow) and two coat protein subunits (pink and magenta).** The refined model is overlaid with the C1 density (contoured at  $4.3\sigma$ ). The magnified panels on the right-hand side reveal atomic details of the bonds formed by two coat subunits and the head-to-tail adaptor prong.



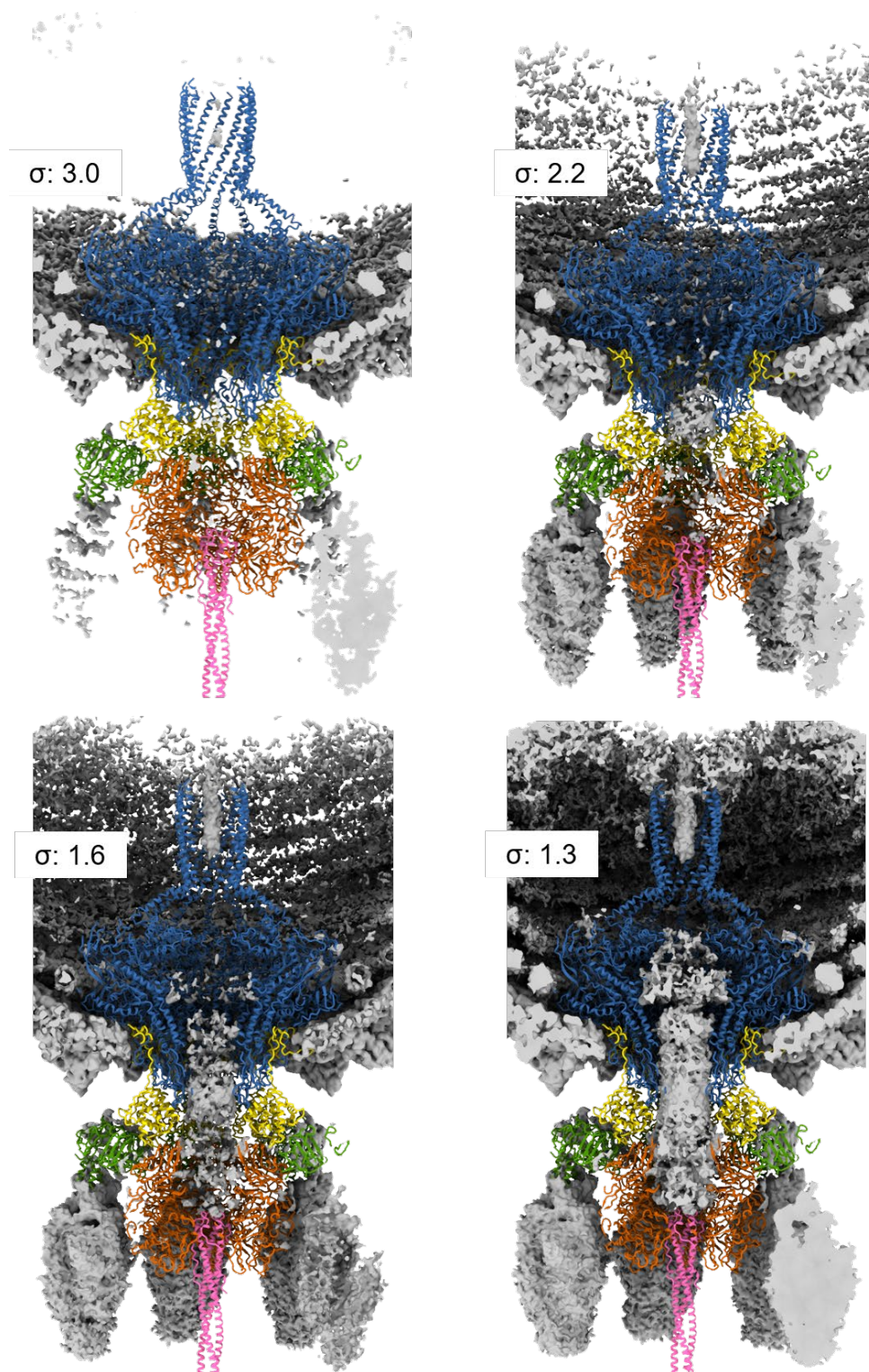


**Figure S10. Sf6 tail needle.** The cryo-EM map of the tail needle was extracted from the C1 reconstruction and overlaid to the refined atomic model. The cryo-EM map (light gray) is contoured at  $1.5\sigma$  at 7 Å  $FSC_{0.143}$ . The model (magenta) was real-space refined against the cryo-EM map yielding a final correlation coefficient  $CC = 0.61$ .

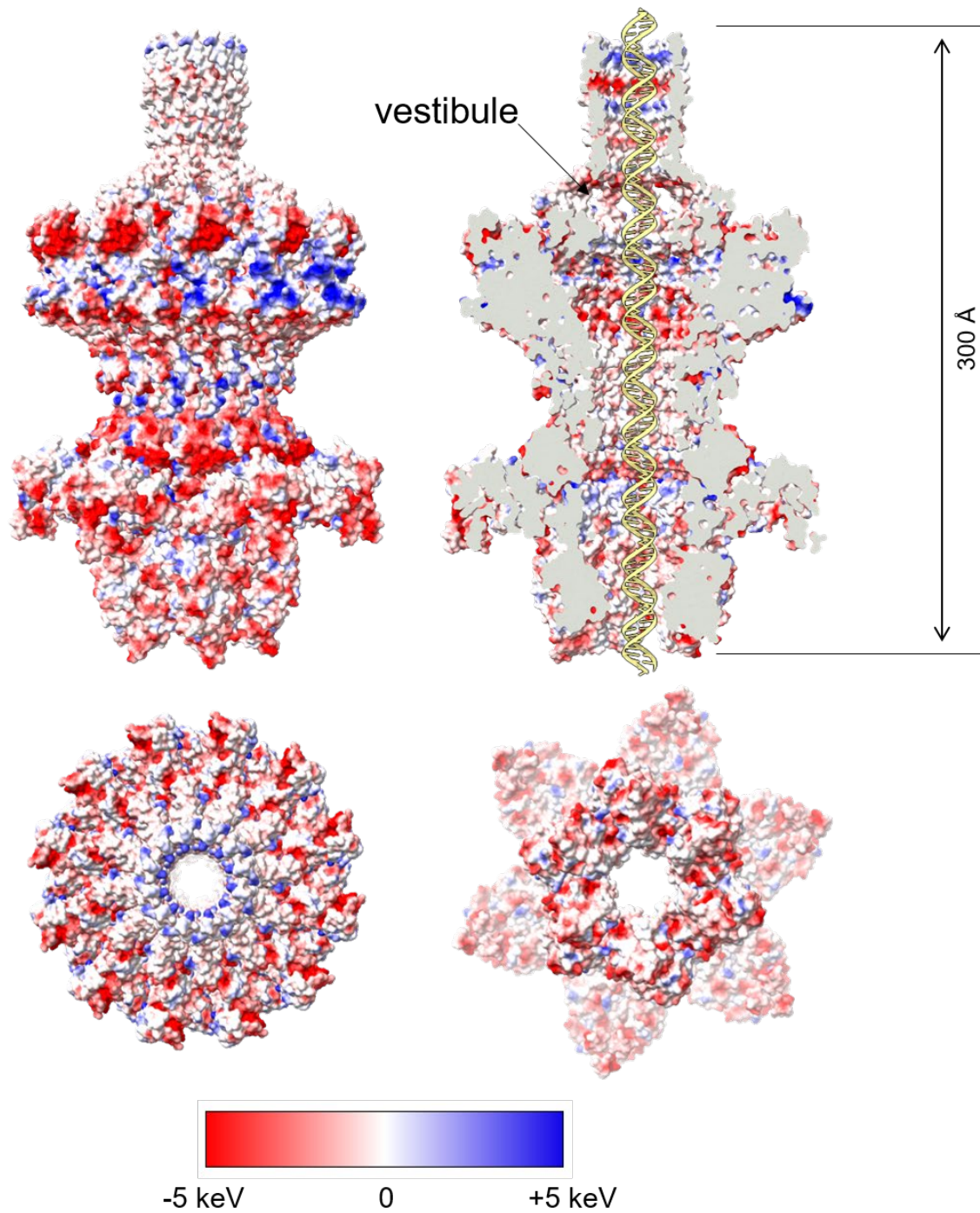


**Figure S11. Reconstruction of the full-length Sf6 tailspike.** (A) C6-symmetrized localized reconstruction of the Sf6 tailspike protomer showing the mask applied during classification and refinement. The mask was generated using SCIPION 3.0. (B) Final 3.1 Å map of the tailspike core (contoured at  $3.6\sigma$ ) overlaid to the refined atomic model. (C) A composite structure of the full-length tailspike protomer (left) and trimer (right). The hinge between the tailspike N-termini and core is shown by a red circle. (D) Schematic diagram illustrating the degree of motion of each tailspike timer (shown as two triangles).





**Figure S12. Density inside the Sf6 tail channel.** The C1 map was displayed in ChimeraX at different contour levels between 1.3 - 3.0 $\sigma$ . The density in the portal barrel did not vary significantly and likely corresponds to dsDNA. Instead, the sparse density at the bottom of the tail channel, visible only at the lower contour and absent in the portal vestibule, may represent a proteinaceous component, possibly one or more of Sf6 ejection proteins.



**Figure S13. Sf6 genome-delivery tail channel.** Side (left) and section (right) views of the Sf6 genome-delivery tail channel showing the Coulombic electrostatic potential surface calculated using ChimeraX. A 100-mer dsDNA was docked inside the channel.

**Table S1. Conservation between Sf6 and P22 capsid and tail factors.**

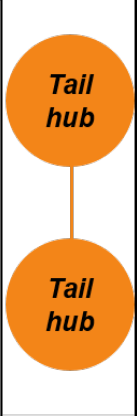
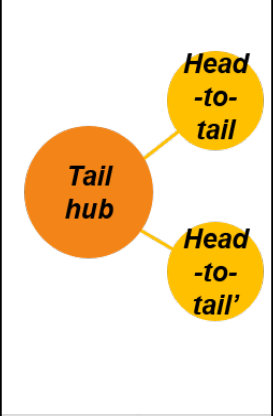
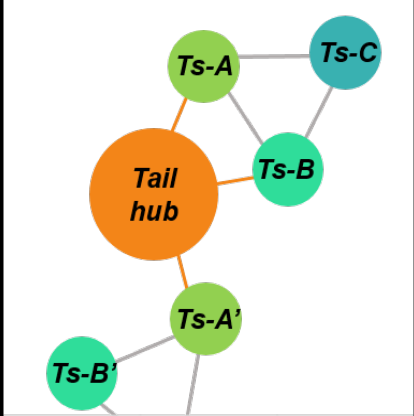
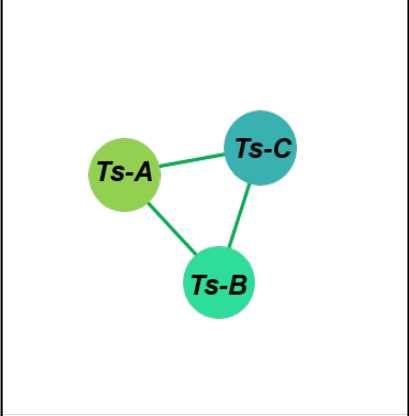
Protein name	Gene product		Amino acids		Amino acid sequence Similarity/Identity (%)	Copies per Virion	PDB IDs	
	<i>Sf6</i>	<i>P22</i>	<i>Sf6</i>	<i>P22</i>			<i>Sf6</i>	<i>P22</i>
Coat protein	Gp5	Gp5	423	430	33.5 / 17.7	415	5UU5 (cryo-EM)	2XYZ (cryo-EM)
Portal protein	Gp3	Gp1	708	725	46.0 / 29.2	12	n/a	3LJ5 (X-ray)
Head-to-tail adaptor	Gp7	Gp4	160	166	51.8 / 33.3	12	5VGT (X-ray)	4V4K (X-ray)
Tail hub	Gp8	Gp10	472	472	95.6 / 93.2	6	n/a	n/a
Tailspike	Gp14	Gp9	623	667	33.3 / 21.1	18	2VBE * (X-ray)	1TYV ** (X-ray)
Tail Needle	Gp9	Gp26	282	233	44.1 / 33.3	3	3RWN *** (X-ray)	2POH (X-ray)

\* Partial structure comprising only the tailspike core (residues 113-622)

\*\* Partial structure comprising only the tailspike core (residues 113-666)

\*\*\* Partial structure comprising only the tail needle knob (residues 128-282)

**Table S2. Summary of all binding interfaces, bonding, and energetics in Sf6 tail determined using PISA**

												
<b>Interface</b>	<b>Tail hub / Tail hub</b>	<b>Tail hub / Head-to-tail</b>	<b>Tail hub / Head-to-tail'</b>	<b>Tail hub / Ts-A</b>	<b>Tail hub / Ts-B</b>	<b>Tail hub / Ts-A'</b>	<b>Ts-A / Ts-B</b>	<b>Ts-A / Ts-C</b>	<b>Ts-B / Ts-C</b>			
No. of interface residues	37/23	8/8	6/4	11/10	7/7	5/7	28/26	20/24	23/20			
Interface areas (Å <sup>2</sup> )	1868/1995	563/634	414/439	558/549	414/430	199/201	1619/1784	1440/1385	1277/1359			
No. of salt bridges	6	-	-	2	2	-	-	1	-			
No. of hydrogen bonds	9	2	3	6	3	2	8	7	6			
No. of non-bonded contacts	119	140	25	47	25	16	91	68	69			