Supporting Information

Chen et al. 10.1073/pnas.1015739108

SI Materials and Methods

Purification of P22 Procapsids and Virions. P22 procapsids were purified as previously described (1, 2). Briefly, *Salmonella typhimurium* (strain DB7136) was infected with P22 phage carrying two mutations: 2*am*H202, which blocks DNA packaging, and 13*am*H101, which delays lysis. After 3.5 h, the infected cells were pelleted by centrifugation, lysed by repeated freeze/thaw cycles and treated with DNase and RNase. The cell debris was removed by low-speed centrifugation ($5{,}000 \times g$) and the procapsids were collected by high-speed centrifugation ($155{,}000 \times g$) after 1 h. Finally, wild-type procapsids were purified by chromatography using an S-1000 column to remove membrane vesicles. Fractions containing procapsids were concentrated by centrifugation at $140{,}000 \times g$ for 1 h and then resuspended in buffer B (50 mM Tris pH 7.6, 25 mM NaCl, 2 mM EDTA). The resuspended procapsids were sedimented through a 5–45% sucrose gradient in a Beckman SW28 swinging bucket rotor at 28,000 rev/ min for 3 h. A blue opalescent band was removed and dialyzed against buffer B. SDS gel electrophoresis of the sample with visualization by Coomassie blue staining revealed major bands corresponding to the coat and scaffolding proteins, and minor bands corresponding to the gene 1 portal protein and gene 16 and gene 20 ejection proteins (3–5). The protein bands exhibited the mobilities expected of intact polypeptide chains.

The infectious P22 virion particles were prepared in a similar way to the procapsids. The phage strain infecting *Salmonella typhimurium* did not carry the 2*am*H202 mutation but carried the 13*am*H101 mutation to delay lysis to accumulate more virus particles. Samples were sedimented through 15–30% sucrose gradients to separate the DNA-containing infectious virion particles from both procapsids and virion particles that lacked DNA. The samples were dialyzed to remove sucrose and then concentrated.

Electron Cryomicroscopy (Cryo-EM) and Image Processing. R1.2/1.3 400-mesh copper Quantifoil grids (Quantifoil Micro Tools GmbH) were soaked in acetone overnight, then washed several times in both acetone and water before being dried overnight to minimize the charging and/or beam-induced movement at liquid helium specimen temperature. The grids were preirradiated for at least 4 h at 10,000× magnification in a JEM1200 electron microscope (JEOL Ltd.) with a tungsten filament operated at 100 kV (6, 7). The pretreated Quantifoil grids were glow discharged for 10 s just before the sample freezing. No continuous carbon film was applied to the holey grids. A droplet (~3 μL) of the sample was applied to the holey film side of the grids, blotted for 2 s at 95–100% humidity and room temperature, and immediately plunged into liquid ethane using a Vitrobot (FEI, Einhoven). The frozen grids were stored in liquid nitrogen. Data for the icosahedral reconstructions of the procapsid and the virion were collected on a JEM-3000SFF electron cryomicroscope (JEOL Ltd.) with a field emission gun and a top-entry specimen stage, operated at 300 kV at liquid helium specimen temperature (~4 K). The images were collected on Kodak SO-163 photographic film at 60,000× nominal magnification with ~36 e/Å^2 dose per image and with 2.0-s exposure time. Films were developed for 12 min in Kodak developer D-19 and then fixed for 10 min in Kodak fixer at 20 °C. The photographic films were digitized on a Nikon Super Coolscan 9000 scanner with a 6.35-μm step size yielding a final pixel size of 1.06 Å. Additional data were collected for the procapsid asymmetric reconstruction on a Gatan 10×10 k CCD (US10000XP, model 990) with 2× binning on a 300-kV JEM-3200FSC cryo-EM (JEOL Ltd.) with an energy slit of 20 eV in the in-column energy filter at ~20 e/Å^2 dose per frame with an effective magnification of 70,600× (yielding 2.55 Å/pixel at 2× binning). The specimen temperature was kept at ~101 K.

Particle images were automatically picked using *ethan* (8). The boxed-out particles were screened manually using *boxer* (9) to eliminate bad particles and manually add particles missed by *ethan*. A box size of 864×864 was chosen for both the procapsid and virion images. A box size of 432×432 was used for the icosahedral reconstruction of the empty procapsid (10) and the procapsid asymmetric reconstruction. Contrast transfer function (CTF) fitting was performed on micrographs or CCD frames without drift, charging or astigmatism. The initial CTF fitting was done automatically using *fitctf.py* (11) and fine-tuned manually using *ctfit* (9).

Structure refinement for both the procapsid and virion was done using real-space projection matching in EMAN with newly implemented iterative programs for global optimization and defocus compensation. An initial 3D reference map was reconstructed from the particles without binning. Using newly developed local refinement methods, the density map was refined using the previous iteration's alignment parameters as a starting point. After several iterations of local orientation refinement, the reconstruction was improved to ~4.5-Å resolution based on Fourier shell correlation (FSC) = 0.5 criterion for two split datasets. Defocus refinement was applied to determine a more accurate defocus for individual particles. Then new defocus parameters were used for a local orientation refinement; each iteration included a step of defocus refinement followed by a step of local orientation refinement. Several such iterations were performed. Resolutions for the final maps of both the procapsid and virion were determined to be 3.8 and 4.0 Å based on the 0.5 criterion of FSC (12) from two independent half datasets.

The asymmetric reconstruction for the procapsid was accomplished using the method described previously (13). Briefly, initial particle orientations were determined using icosahedral symmetry. The faint density for the portal was segmented out and 12-fold averaged to aid in the subsequent steps for determining the location and orientation of the portal in each particle image. The subsequent reconstruction was done with C1 symmetry, producing a new density map of the particle including the portal. The choice of particles for asymmetric reconstructions was based on the statistical significance (2σ) of the lowest phase residual in the orientation search of each particle. This process was iterated until the 12 portal subunits and the long helix of the coat protein were readily identifiable. The resolution for the procapsid asymmetric reconstruction was assessed to be 8.7 Å based on the comparison with the 3.8-Å icosahedral reconstruction of the procapsid. The ~7.0-Å icosahedral reconstruction for the empty procapsid, in which the scaffolding proteins were removed by GuHCl, was performed from the published data (10) using the MultiPath Simulated Annealing optimization algorithm described previously (14).

Segmentation of Subunits. The original density maps of both the procapsid and virion were cropped to include one asymmetric unit, containing a complete hexamer and its neighboring pentamer subunit. Initially, the HK97 capsid protein structure [Protein Data Bank (PDB) ID: 1OHG] (15) was fit to each of the seven P22 monomers in both the procapsid and virion using Chimera (16). Preliminary segmentation of the P22 monomers in both structures were then done using Chimera's "Zone" utility with

a radius of approximately 20 pixels. The segmented subunits were then normalized and cropped to $128 \times 128 \times 128$ using *proc3d* in EMAN (9). Segmented subunits were aligned using *Foldhunter* (17) and averaged using *proc3d* for both the procapsid and virion subunits. In each of the averaged maps, a single subunit was manually segmented using Chimera based on density connectivity. A second-round of alignment with *Foldhunter* (17) was performed in which the initial segmented subunits were fit to their respective averaged subunit and averaged a second time. A final manual segmentation of the "second-round" averaged subunits was done in Chimera.

Cα Backbone Trace and Model Optimization. Initial model building for individual subunits was done as described recently (18). Briefly, *SSEHunter* was used to identify the secondary structure elements within the procapsid and virion coat protein density maps. Secondary structure programs, *JPred*, *SSPro*, and *PsiPred*, were used to generate a consensus secondary structure prediction for gp5, the P22 coat protein (sequence ID: NP_059630.1). Using Gorgon (18), a correspondence between the sequence-predicted helices and the *SSEHunter*-identified helices was calculated. A Cα trace connecting the helices was then constructed in Gorgon and optimized in Coot (19) for both the procapsid and virion subunits. Both Cα models were then placed back into their corresponding positions in an asymmetric unit in the original, nonaveraged, density maps. Manual refinement of the Cα positions was done in Coot to eliminate visible clashes among neighboring subunits and fit local regions of conformational variance in the individual subunits.

To improve our de novo models, we used Rosetta with cryo-EM density map constraints to generate full atomic models for one coat protein subunit from the procapsid and one coat protein subunit from the virion. In this procedure, a sequence alignment between gp5 and possible structural templates is required for model generation. As mentioned, no sequence homology was detectable over the entire coat protein (430 aa residues), though HHsenser (20) was used to generate an initial sequence alignment to the HK97 coat protein (PDB ID: 1OHG) over the P22 coat protein sequence, excluding residues in the extra density (ED) domain (residues 217–370) and the N arm (residues 1–29). An additional sequence alignment was also generated from the structural superposition of 1OHG and 2E0Z (a 38.8 kDa protein from *Pyrococcus furiosus* virus-like particles), a structurally homologous protein. In addition, two more sequence alignments were generated based on a structural superposition of our de novo Cα model to the aforementioned 1OHG and 2E0Z structures. These four alignments were used as starting points for the rigid-body perturbation protocol described previously for optimizing initial cryo-EM based model (21). Following three rounds of iterative model rebuilding and optimization, the final procapsid and virion models for gp5 (excluding the N arm and ED domain) were chosen based on a combination of Rosetta score and fit-to-density score.

The full atom model for a single subunit (excluding the N arm and ED domain) was fit to the coat protein positions for all subunits in a pentamer and hexamer of the procapsid and the virion separately. In each case, iterative rebuilding and optimization was again used to eliminate clashes in the assembly and improve regions with poor fit-to-density scores in the capsomeres. Refinement of the pentamers and close-to-symmetric virion hexamers used Rosetta's symmetric modeling protocol (22); the skewed hexamer assembly in the procapsid was refined asymmetrically. A final round of side-chain rotamer optimization and backbone and side-chain torsion minimization was performed with an increased fit-to-density weight. The ED domain and N arm, for which no structural templates were obtained, were added back to the models based on our Cα de novo model. Through this process, complete models for the pentamers and hexamers in both particles were derived. The Cα positions in these models were nearly identical to those from our original de novo models though intracapsomere clashes were removed. In order to obtain the final model, another round of refinement with Coot and Gorgon was used to eliminate clashes within and between asymmetric units.

1. Fuller MT, King J (1981) Purification of the coat and scaffolding proteins from procapsids of bacteriophage P22. *Virology* 112:529–547.
2. Prevelige PE, Jr, Thomas D, King J (1988) Scaffolding protein regulates the polymerization of P22 coat subunits into icosahedral shells in vitro. *J Mol Biol* 202:743–757.
3. King J, Lenk EV, Botstein D (1973) Mechanism of head assembly and DNA encapsulation in Salmonella phage P22. II. Morphogenetic pathway. *J Mol Biol* 80:697–731.
4. Casjens S, King J (1974) P22 morphogenesis. I: Catalytic scaffolding protein in capsid assembly. *J Supramol Struct* 2:202–224.
5. Greene B, King J (1996) Scaffolding mutants identifying domains required for P22 procapsid assembly and maturation. *Virology* 225:82–96.
6. Chen DH, Jakana J, Chiu W (2007) Single-particle cryo-EM data collected on a 300-kV liquid helium-cooled electron cryomicroscope. *J Chinese Elec Microsc Soc* 26:473–479.
7. Miyazawa A, Fujiyoshi Y, Stowell M, Unwin N (1999) Nicotinic acetylcholine receptor at 4.6 A resolution: Transverse tunnels in the channel wall. *J Mol Biol* 288:765–786.
8. Kivioja T, Ravantti J, Verkhovsky A, Ukkonen E, Bamford D (2000) Local average intensity-based method for identifying spherical particles in electron micrographs. *J Struct Biol* 131:126–134.
9. Ludtke SJ, Baldwin PR, Chiu W (1999) EMAN: Semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* 128:82–97.
10. Jiang W, et al. (2003) Coat protein fold and maturation transition of bacteriophage P22 seen at subnanometer resolutions. *Nat Struct Biol* 10:131–135.
11. Yang C, et al. (2009) Estimating contrast transfer function and associated parameters by constrained non-linear optimization. *J Microsc* 233:391–403.
12. Harauz G, van Heel M (1986) Exact filters for general geometry three dimensional reconstruction. *Optik* 73:146–156.
13. Liu X, Rochat RH, Chiu W (2010) Reconstructing Cyano-bacteriophage P-SSP7 structure without imposing symmetry. *Protoc Exch*, 10.1038/nprot.2010.96.
14. Liu X, Jiang W, Jakana J, Chiu W (2007) Averaging tens to hundreds of icosahedral particle images to resolve protein secondary structure elements using a Multi-Path Simulated Annealing optimization algorithm. *J Struct Biol* 160:11–27.
15. Wikoff WR, et al. (2000) Topologically linked protein rings in the bacteriophage HK97 capsid. *Science* 289:2129–2133.
16. Pettersen EF, et al. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612.
17. Jiang W, Baker ML, Ludtke SJ, Chiu W (2001) Bridging the information gap: Computational tools for intermediate resolution structure interpretation. *J Mol Biol* 308:1033–1044.
18. Baker ML, Zhang J, Ludtke SJ, Chiu W (2010) Cryo-EM of macromolecular assemblies at near-atomic resolution. *Nat Protoc* 5:1697–1708.
19. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr, Sect D: Biol Crystallogr* 60:2126–2132.
20. Soding J, Remmert M, Biegert A, Lupas AN (2006) HHsenser: Exhaustive transitive profile search using HMM-HMM comparison. *Nucleic Acids Res* 34:W374–378.
21. DiMaio F, Tyka MD, Baker ML, Chiu W, Baker D (2009) Refinement of protein structures into low-resolution density maps using rosetta. *J Mol Biol* 392:181–190.
22. Andre I, Bradley P, Wang C, Baker D (2007) Prediction of the structure of symmetrical protein assemblies. *Proc Natl Acad Sci USA* 104:17656–17661.

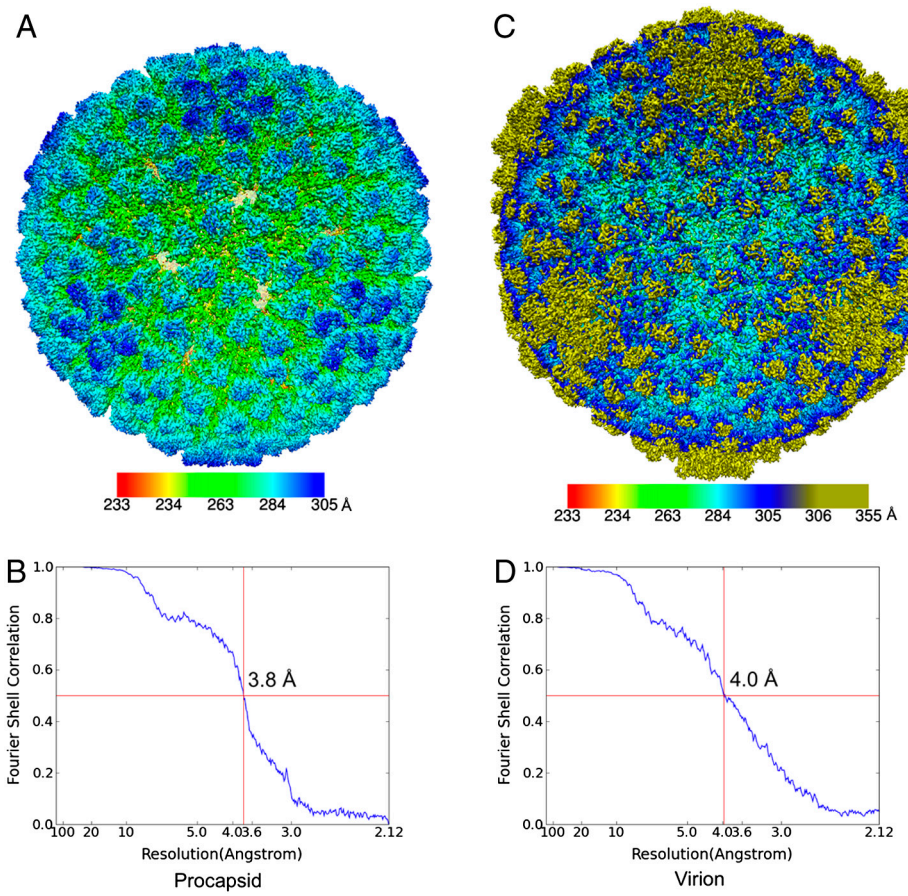


Fig. S1. Cryo-EM density maps of the P22 procapsid and virion. Radially colored cryo-EM density maps of procapsid (*A*) and virion (*C*), viewed along an icosahedral threefold axis with their corresponding color schemes shown below each map. Resolution measurements with the FSC = 0.5 criterion show 3.8- and 4.0-Å resolution for the procapsid (*B*) and virion (*D*) reconstructions, respectively.

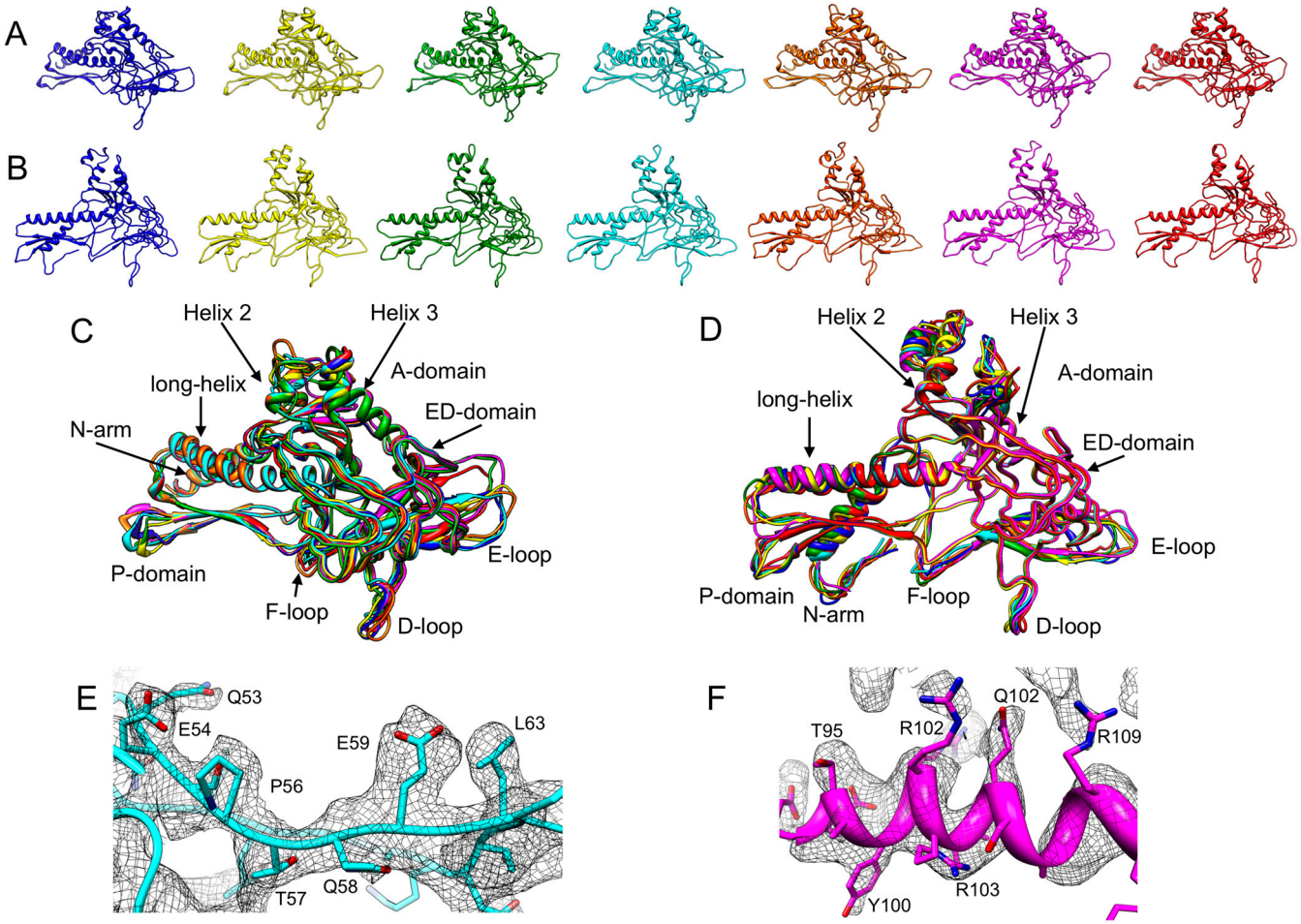


Fig. S2. Coat protein models for one asymmetric unit of the P22 procapsid and virion. Seven individually colored subunits from one asymmetric unit of the procapsid (*A*) and virion (*B*) models. Superimposition of those seven subunits from the procapsid (*C*) and virion (*D*) models. The eight domains and the helix 2 and 3 in the A domain of the coat protein are labeled. (*E* and *F*) Part of models for the procapsid and virion, respectively, superimposed with side-chain densities.

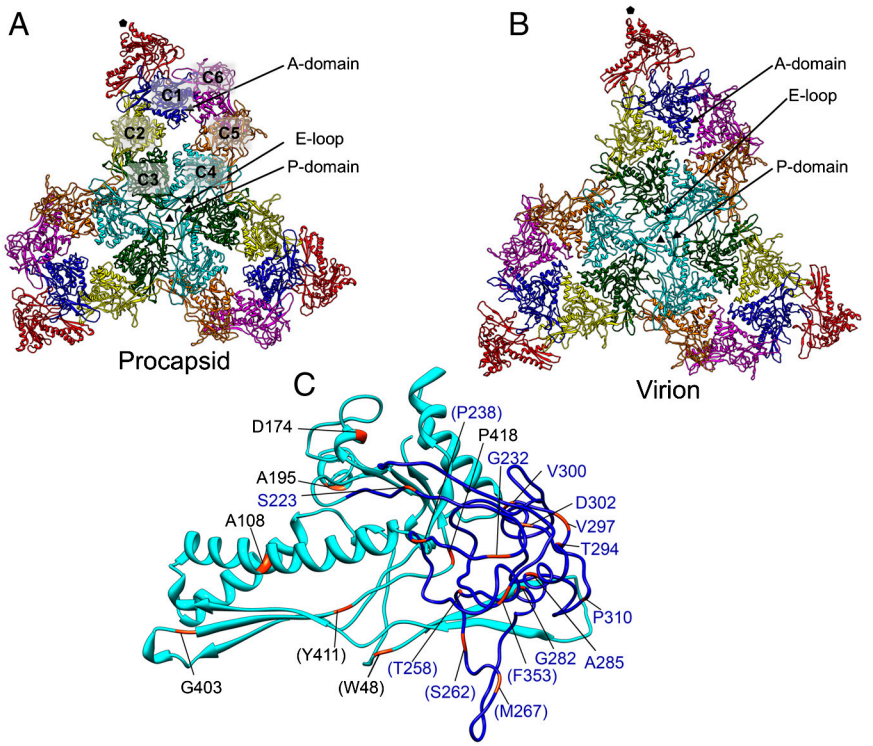


Fig. S3. The inter- and intracapsomere interactions. (*A* and *B*) Three asymmetric units from one icosahedral facet of procapsid (*A*) and virion models (*B*). Six coat protein subunits (C1–C6) are labeled in *A*. Lateral interactions between helix 2 of one subunit and helix 3 of an adjacent subunit vary within the skewed hexamers of procaspid and are uniformly tight in the virion. (*C*) Temperature-sensitive mutations mapped in one procapsid subunit. Fourteen mutations (orange) are located in ED domain (blue) with their residue numbers labeled in blue. The seven mutations involved in the local twofold interactions are labeled in parentheses.

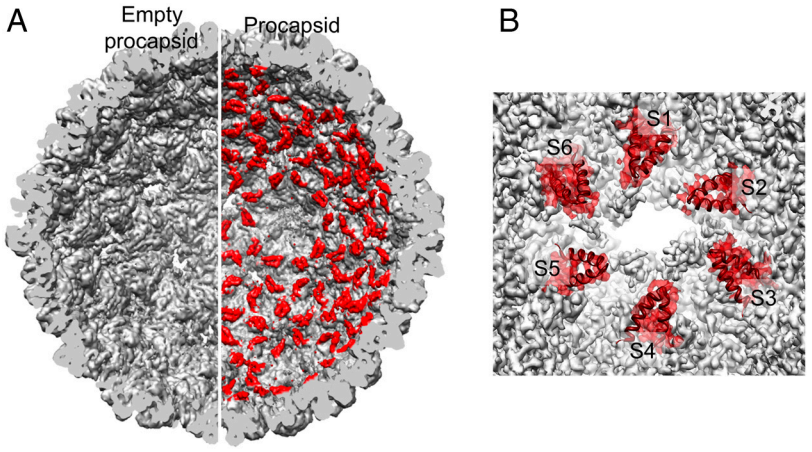


Fig. S4. The scaffolding protein. (*A*) Half of the interior surface of the empty P22 procapsid (*Left*) and the scaffolding-containing P22 procapsid (*Right*) maps reconstructed with imposed icosahedral symmetry. The procapsid map with scaffolding proteins (red) is from the 3.8-Å icosahedral reconstruction filtered to 7.0-Å resolution. The procapsid density beyond the scaffolding protein density (red) on the right half was colored in gray. (*B*) The fit of the NMR structure of the C-terminal helix-loop-helix motif of the scaffolding protein, 2GP8 (red ribbon) into each V-shaped density (S1–S6) at one hexamer in the 3.8-Å resolution procapsid map.

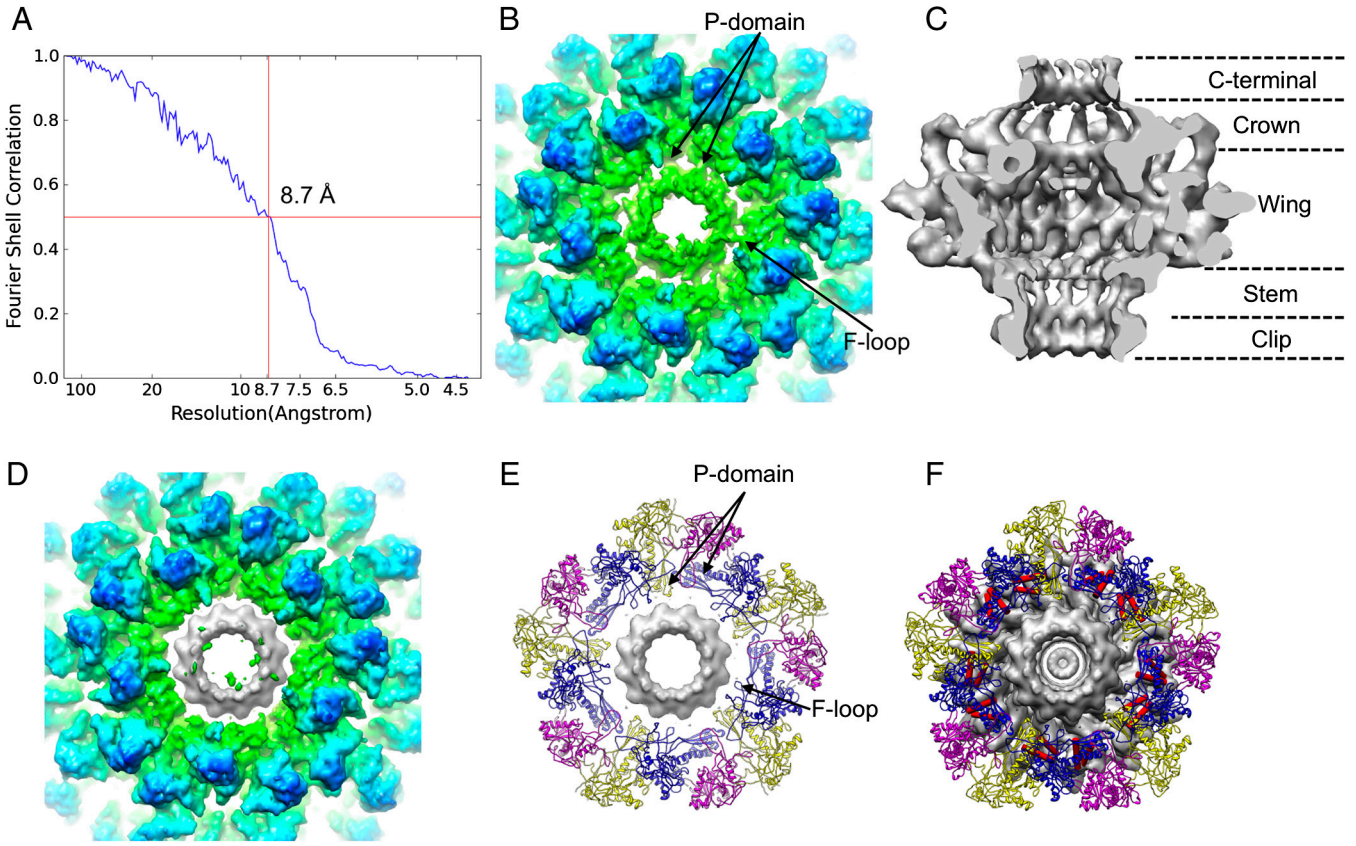
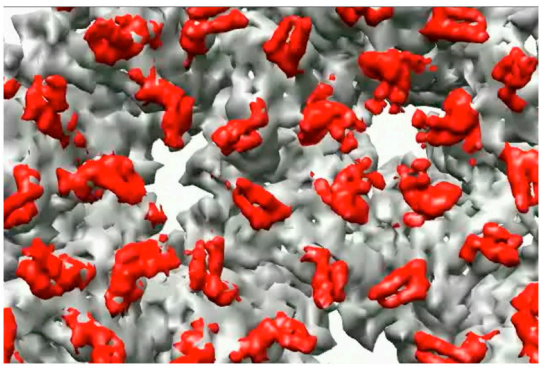
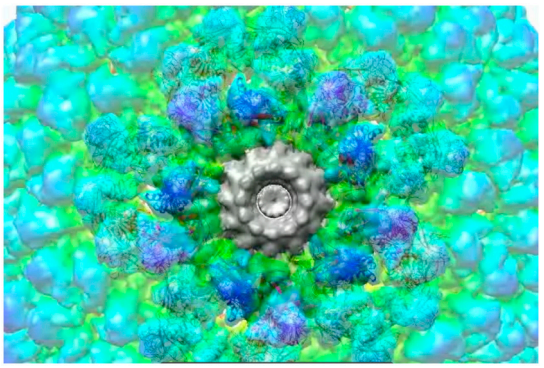


Fig. S5. Portal in the asymmetric P22 procapsid reconstruction. (*A*) Resolution for the asymmetric reconstruction of the procapsid is 8.7 Å. (*B*) Slice of the density map at the unique portal vertex viewed from outside of the procapsid colored radially as in Fig. 1*D*. Some density connections are seen between the clip domain of the portal (central green ring density) and the coat protein subunits. The scaffolding protein densities are not visible in this display because they are underneath the coat proteins. (*C*) The segmented and 12-fold averaged portal with the domains labeled according to the nomenclature previously introduced (1). The putative C-terminal domain of the portal protein is pointing toward the interior of the procapsid. The domain arrangement of this in situ portal is similar to that of biochemically purified portal (2), the structural differences may be due to their different states or associated with the possible binding of the yet to be identified ejection proteins. (*D*) Same as *B*, except with the replacement of the clip domain (gray) of the segmented and 12-fold averaged portal density. (*E*) Same as *D*, except with the replacement of the models of 15 coat protein subunits. (*F*) View of the unique portal vertex from outside of the procapsid including the entire portal, 10 scaffolding protein C termini (red cylinders), and 15 coat protein models fit to the corresponding densities. This display is an end-on view of that shown in Fig. 2*D*.

1 Lebedev AA, et al. (2007) Structural framework for DNA translocation via the viral portal protein. *EMBO J* 26:1984–1994.
2 Lander GC, et al. (2009) The P22 tail machine at subnanometer resolution reveals the architecture of an infection conduit. *Structure* 17:789–799.

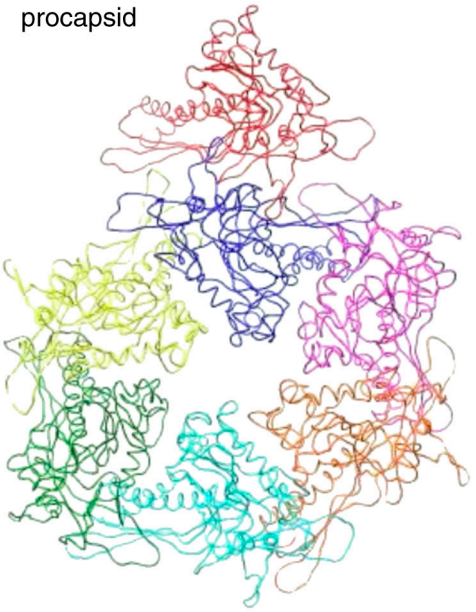


Movie S1. Scaffolding proteins on the inner surface of P22 procapsid shell from the icosahedral cryo-EM map at 7.0-Å resolution. The 7.0-Å map was low-pass filtered from the original 3.8-Å one. The V-shaped scaffolding protein C termini densities in red are located at the hexamers and pentamers. Each scaffolding protein cryo-EM density fits well with the ribbon (PDB ID: 2GP8) of C-terminal coat-binding domain of P22 scaffolding protein solved by NMR.
Movie S1 (MP4)



Movie S2. Interactions among the portal, scaffolding proteins and coat proteins around the portal vertex in the 8.7-Å resolution asymmetric map of the P22 procapsid. The scaffolding protein densities in red, similar to those found in the icosahedral procapsid reconstruction and representing the C termini of the scaffolding proteins, are also observed on the inner surface of capsid shell in the asymmetric map.

Movie S2 (MP4)



Movie S3. Conformational transition of P22 from procapsid to virion. This represents a simple morphing between two states of P22, but may not be the actual conformational trajectory during P22 capsid maturation. The overall capsid conformation changes from a round procapsid to a polyhedral virion with a thinner capsid shell and more symmetric hexamers in the virion. Coat protein conformational changes include the flexing of the A-domain tip, the unkinking of the long helix, an N-arm rotation, and movements of D and E loops.

Movie S3 (MP4)