PNAS
www.pnas.org

Supplementary Information for

A Novel Packing for A-form DNA in an Icosahedral Virus

Fengbin Wang, Ying Liu, Zhangli Su, Tomasz Osinski, Guilherme A.P. de Oliveira,

James F. Conway, Stefan Schouten, Mart Krupovic,

David Prangishvili and Edward H. Egelman

Edward H. Egelman
Email: egelman@virginia.edu

This PDF file includes:

Supplementary text
Figures S1 to S11
Table S1
SI References

www.pnas.org/cgi/doi/10.1073/pnas.1908242116

Supplementary Methods

SPV1 production and purification

The virus was produced as described previously (1). The infected cell culture was harvested, and cells were removed by centrifugation (Sorvall SLA 3000 roter, 9,000 rpm, 30 min). The supernatant (viral fraction) was collected and concentrated by flipflow filtration using Vivaflow-200 cassettes (Sartorius Stedim Biotech, France). The concentrated virions were suspended in the sample buffer (citric acid – Na_2HPO_4 buffer (pH3.6) containing 500 mM NaCl). For cryo-EM imaging with a coarser sampling of 3.71 Å per pixel used in SPV1 asymmetrical reconstruction, the virions were further purified in 0.45 g ml^{-1} CsCl (in the sample buffer) by isopycnic gradient centrifugation (Beckman SW60 Ti rotor, 215,000 × g, 18 h, 15°C). For lipid analysis, the virions were subjected to two rounds of purification, first by rate zonal centrifugation (Beckman SW32 Ti rotor, 82,667 × g, 20 min, 15°C) in 5% to 20% sucrose (in the sample buffer), second by isopycnic density-gradient centrifugation in 0.45 g ml^{-1} CsCl.

Cryo-electron microscopy and image processing

The SPV1 sample was applied to glow-discharged lacey carbon grids and vitrified using a Vitrobot Mark IV (Thermo Fisher) with back side blotting. Grids were first imaged at on a Titan Krios (Thermo Fisher) at the National Cryo-Electron Microscopy Facility (NCEF) at NCI. About 2,000 micrographs were collected. Due to low concentration of virus particles (~4,000 in total), the icosahedral reconstruction only reached a resolution of ~ 6.5 Å. Subsequently, grids with slightly higher particle concentration were imaged at

UVa in a Titan Krios at 300 keV and recorded with a Falcon 3ec (Thermo Fisher) direct electron detector at 1.4 Å per pixel. About 6,000 micrographs were collected using a defocus range of 1.5–2.50 µm, with a total exposure of 2.4 s (amounting to ~ 50 electrons/Å^2) distributed into 24 fractions. These images were not combined with the previous ones due to a different pixel size. Most of the image processing steps were done in RELION (2). First, all micrographs were motion corrected (ignoring the first fraction) using MotionCorr (3) and then the defocus per micrograph was estimated by the gCTF program (4). A total of 19,092 SPV1 virus particles were then manually boxed and extracted. After removing bad particles via 2D classification, a 4.2 Å icosahedral reconstruction was obtained in Refine3D using a sphere as the starting model. After movie-refinement, particle polishing steps, and Ewald sphere correction, the resolution of the SPV1 capsid region was further improved to 3.7 Å. The final volume of the capsid region was estimated to have a resolution of 3.7 Å based on the model:map FSC and d99 (5) (Supp. Fig. 2, Supplementary Table 1), and sharpened with a negative B-factor of -180 Å^2.

To obtain an asymmetrical reconstruction, more micrographs were recorded with a coarser sampling of 3.71 Å per pixel. Micrographs were processed in RELION in the same manner as described, and a total of 77,354 particles were manually selected after auto-picking. After 2D classification, a ~ 40 Å asymmetrical cryo-EM reconstruction (C1 symmetry) was obtained in Refine3D, using the icosahedral reconstruction as the starting reference. A soft mask with the solid icosahedral shape filtered to 60 Å was used in the reconstruction to: (1) get rid of the artifacts introducing by the carbon edge

close to some of the SPV1 particles; (2) limit the Euler angles of individual particles to be close to one of the 60 asymmetric units. A similar result was also obtained using the asymmetrical reconstruction algorithms in JSPR (6).

The extracted SPV1 genome sample (buffer: 100 mM sodium citrate pH 5.0) was shaken at 300 Hz for 24 hours at room temperature, 4 µl applied to freshly glow-discharged lacey carbon grids and then plunge-frozen using the Vitrobot. Frozen grids were imaged with the Falcon 3ec camera at 1.4 Å per pixel, and micrographs were collected using a defocus range of 1.5–2.5 µm, with a total exposure time of 2.4 s (amounting to ~ 52 electrons/Å^2) distributed into 24 fractions. The micrographs were first motion corrected using all the fractions by MotionCorr v2.1 and then used for defocus estimation by CTFFIND3 (7). Filament images were extracted using the e2helixboxer program within EMAN2 (8) from the dose-weighted fractions 2-10 (amounting to ~ 20 electrons/Å^2), after the images were corrected for phase reversals through multiplication by the theoretical CTF. A total of 6,431 overlapping 384-px long segments (with a shift of 32 pixels, which is ~ 1.5 times the axial rise per subunit) were generated. The helical symmetry was determined given the 1/(29 Å) meridional layer-line in the power spectra and the RELION 2D class averages which showed a VP1 dimer within the obvious 29 Å periodicity.

Model building of SPV1 MCPs

First, the density corresponding to a single subunit of VP4 and VP10 was segmented from the icosahedral reconstruction using Chimera (9). About 180 C_α atoms could be traced manually in Coot (10) for both the VP4 and VP10 segmented maps. The

initial models of VP4 and VP10 were built based on this C_α trace using RosettaCM protocols (11), and further edited in Coot and refined in real-space by Phenix (12). Since VP4 hexons do not have C6 symmetry, the density corresponding to a VP4 hexon was then segmented from the icosahedral reconstruction in Chimera. Six copies of VP4 were individually docked into this hexon map, and regions outside the cryo-EM density were adjusted manually in Coot. The VP4 hexon model was subsequently real-space refined against the hexon map without NCS restraints. Finally, one VP10 model and seven VP4 hexon models were docked into the icosahedral asymmetrical unit. The quality of the VP4 and VP10 models was evaluated with MolProbity (13). The refinement statistics are given in Supplementary Table 1.

VP1 purification and gel shift assay

The VP1 protein was ordered from Genscript. Briefly, the VP1 gene was cloned into *E. coli* expression vector pET-30a, and VP1 protein was expressed in 3 L Luria Broth (LB) culture. The VP1 protein was obtained from the supernatant of the *E. coli* cell lysate, followed by four-step purification by Ni column, TEV protease tag removal, 2^{nd} Ni column and SP Sepharose column. The final protein concentration as determined by Bradford protein assay was 0.92 mg/ml, and the purity was about 90% estimated by SDS-PAGE. The storage buffer was 50 mM Tris-HCL, 150 mM NaCl and 10% glycerol pH 8. Two gel shift assays were done. The first was a pH gradient assay, where 50 ng PCR product (114bp) amplified from SPV1 genomic DNA was incubated with 5 µl VP1 (0.92 mg/ml) under varieties of different pH, 10 mM $MgCl_2$ and $CaCl_2$. The incubation was

done at 75 °C for 1 hour, followed by room temperature overnight and then ran on a 1% agarose-TBE gel. The second gel shift was done using a similar setup but with different PCR product (250 bp, 70 ng), under pH 5.5 and several different temperatures.

Dissociation of viral particles

The purified viral sample was dialyzed against water, and then dissociated by freezing at -80°C and thawing at room temperature for 18 cycles.

Analysis of SPV1 and host cell lipids

The freeze-dried host-cell preparation and the twice purified virion preparation were directly acid hydrolyzed by refluxing with 5% HCl in methanol for 3 h to release GDGT lipids. GDGT lipids were analyzed by high-performance liquid chromatography/atmospheric pressure chemical ionization-mass spectrometry (14). The mass spectrometer was operated in single ion mode (SIM) to monitor archaeol and GDGTs with 0 to 8 cyclopentane moieties. Relative abundances of GDGTs were determined by integrating peak areas of the SIM signal.

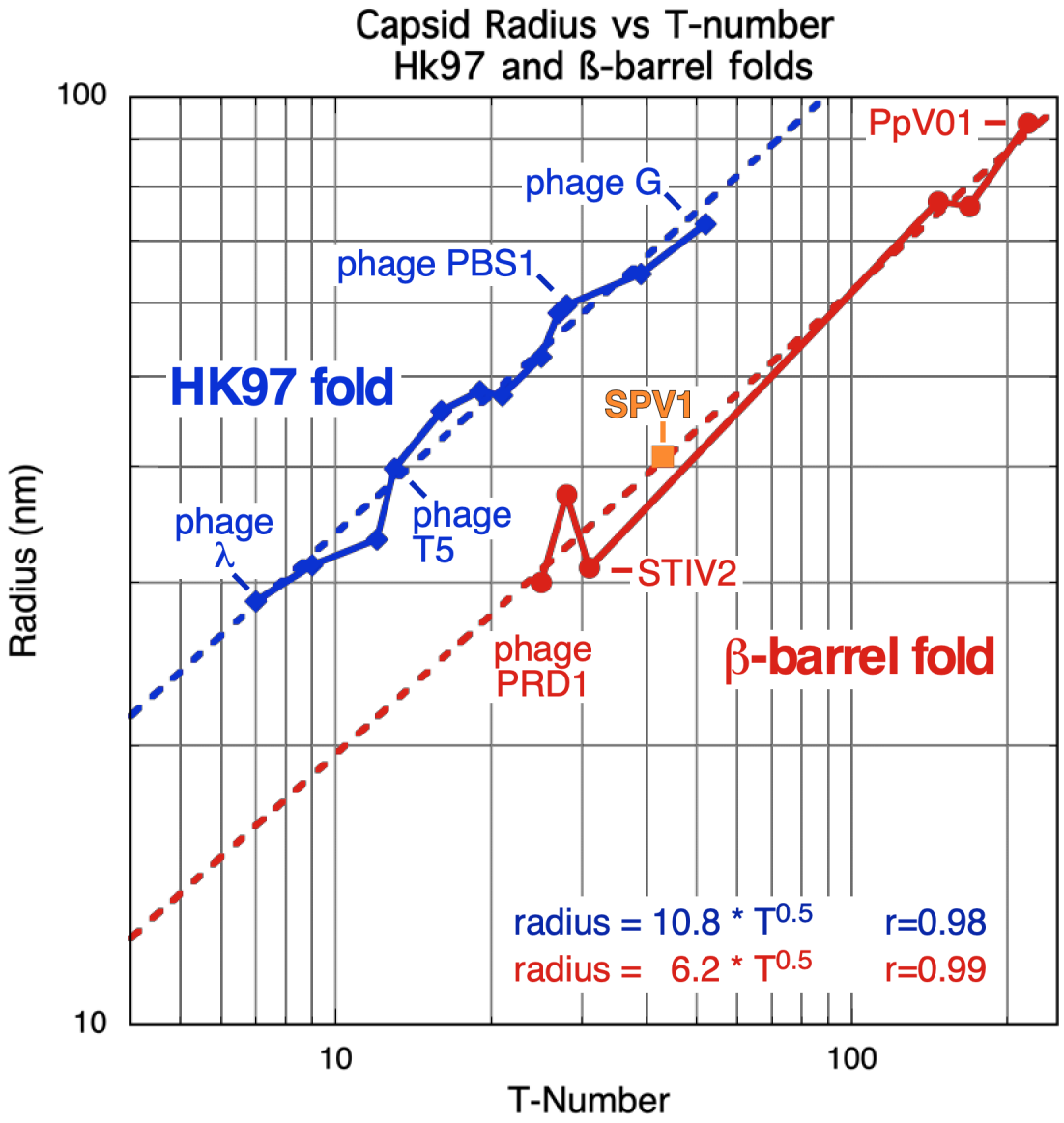

Fig. S1. Comparison of icosahedral capsid size vs Triangulation number for HK97 and ß-barrel fold families.

Radius measurements across the icosahedral 2-fold axis were used to measure capsid size as these were found to be most consistent with density averaged in 3D over the radii of the density maps, and could be measured from data deposited in the EMDataBank and from older published images where no EMDataBank deposition was made.

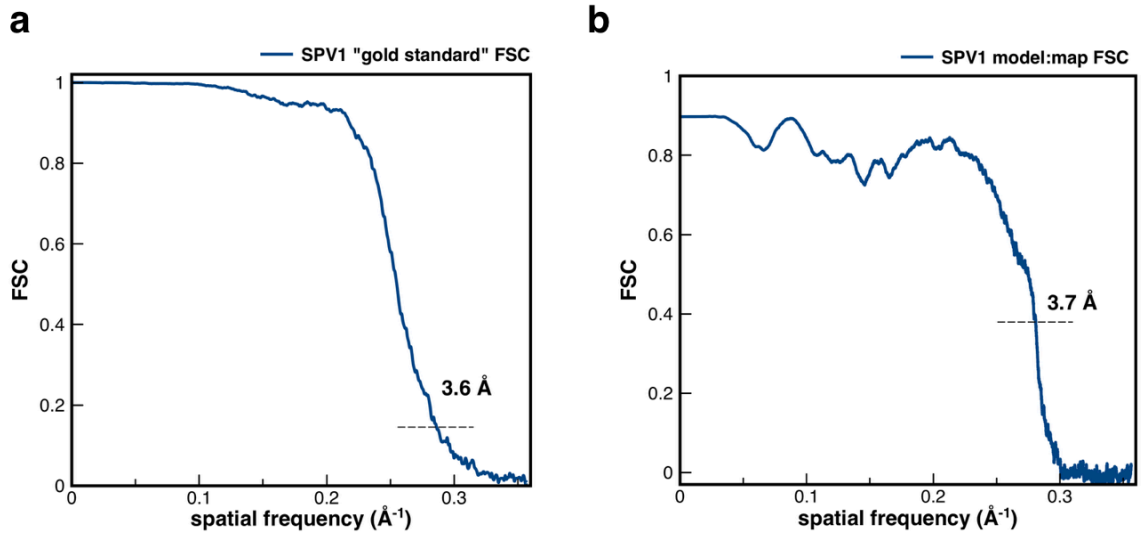

Fig. S2. Fourier Shell Correlation (FSC) calculations

a, The map:map "gold standard" FSC using the 0.143 criterion estimates the final reconstruction to have a resolution of 3.6 Å.

b, The model:map FSC calculation using a 0.38 criterion, which is sqrt (0.143), estimates the final reconstruction to have a resolution of 3.7 Å.

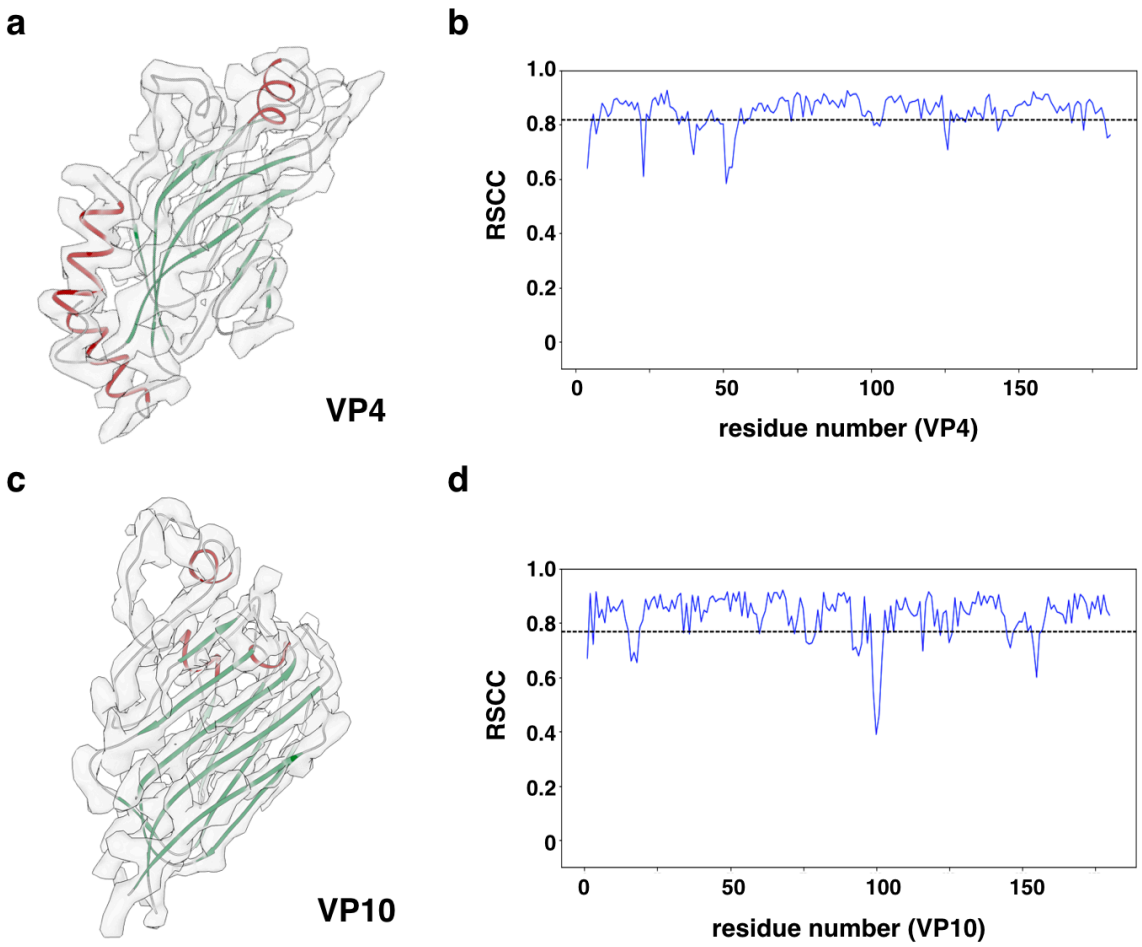

Fig. S3. *De novo* atomic model building of SPV1 major capsid proteins

a, The C_α trace of VP4 fit into the 3.7 Å cryo-EM map.

b, The per-residue real space correlation coefficient (RSCC) of the VP4 atomic model against the 3.7 Å cryo-EM map.

c, The C_α trace of VP10 fit into the 3.7 Å cryo-EM map.

d, The per-residue real space correlation coefficient (RSCC) of the VP10 atomic model against the 3.7 Å cryo-EM map.

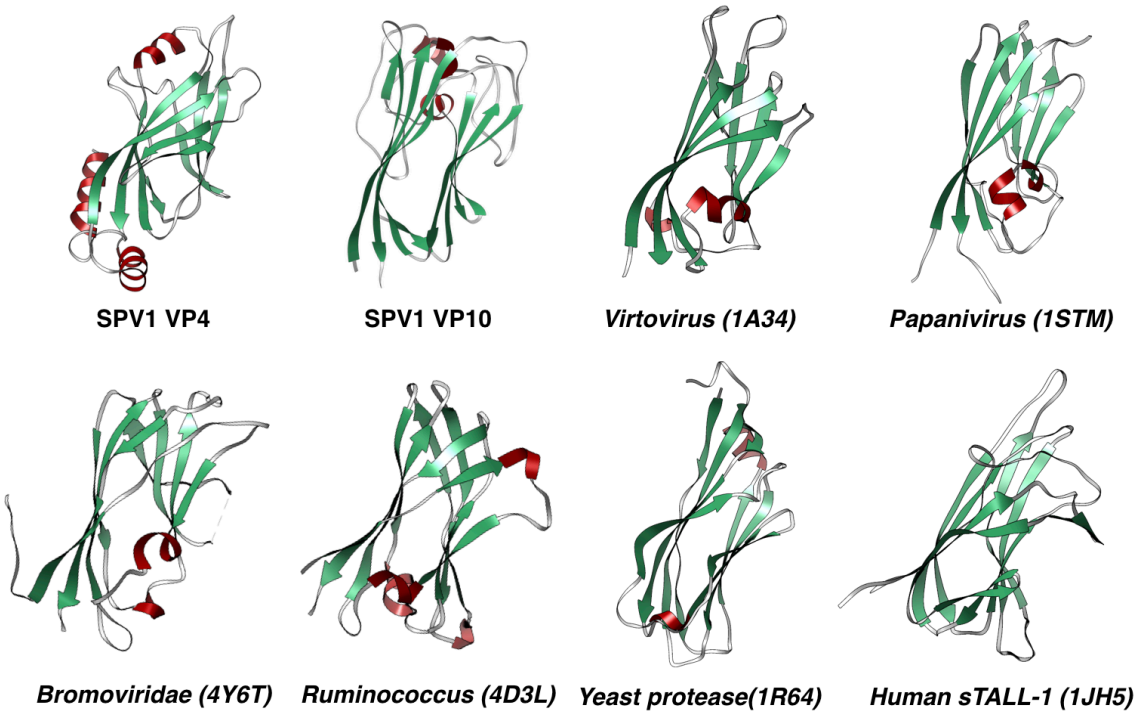

Fig. S4. The folds of VP4 and VP10

The folds of both VP4 and VP10 are single jelly-roll fold proteins, but with some modifications. A comparison with some other single jelly-roll proteins is shown.

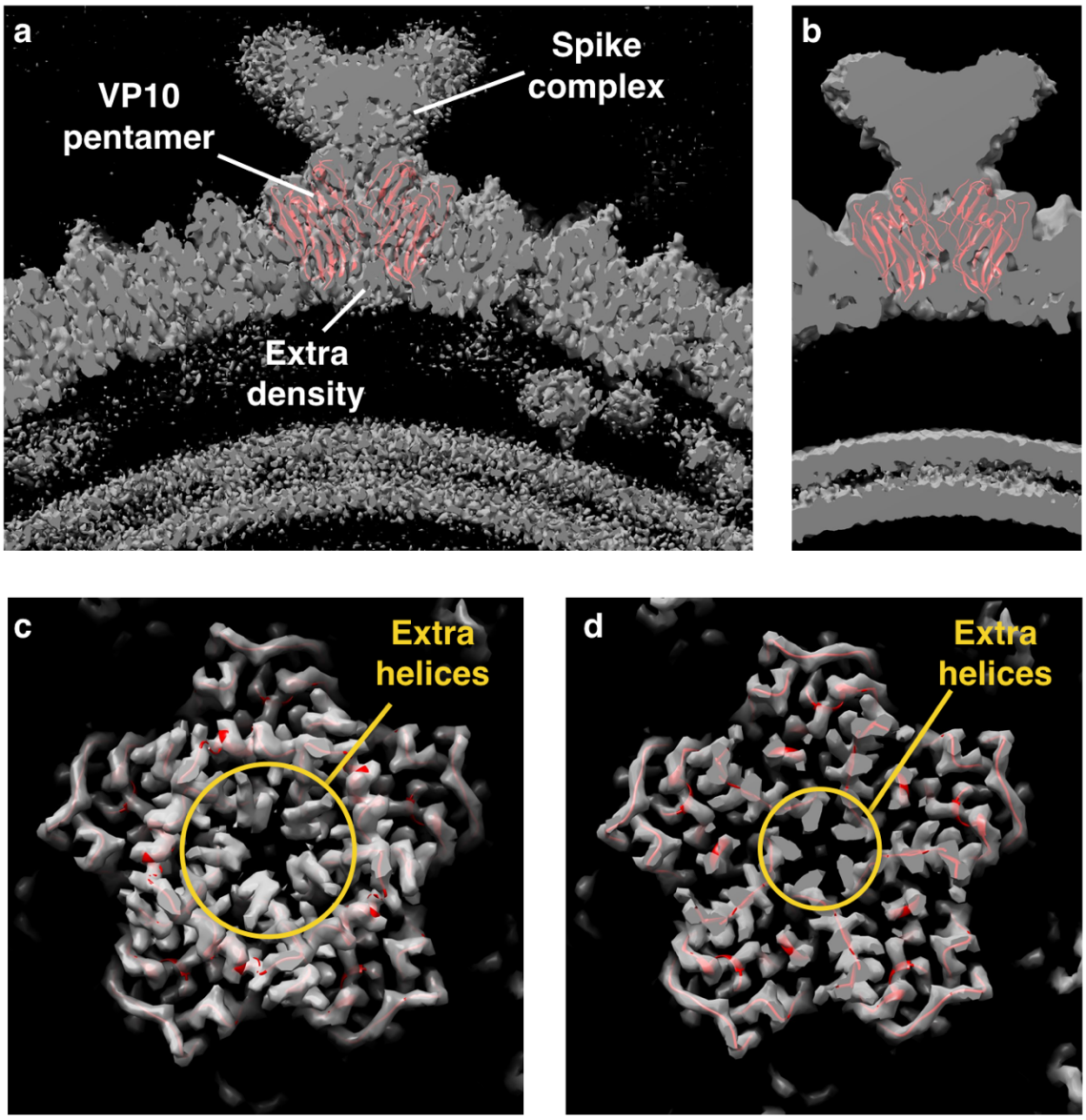

Fig. S5. Interactions at the vertices.

a, Five copies of the VP10 model fit into the SPV1 cryo-EM map, viewed perpendicular to the
icosahedral five-fold axis. The VP10 pentamer, the spike complex and the extra density beneath the pentamer are labeled.

b, Low pass filtered volume showing the same VP10 pentamer area in (a).

c-d, The top view (c) and the central portion (d) of the VP10 pentamer showing the extra helices.

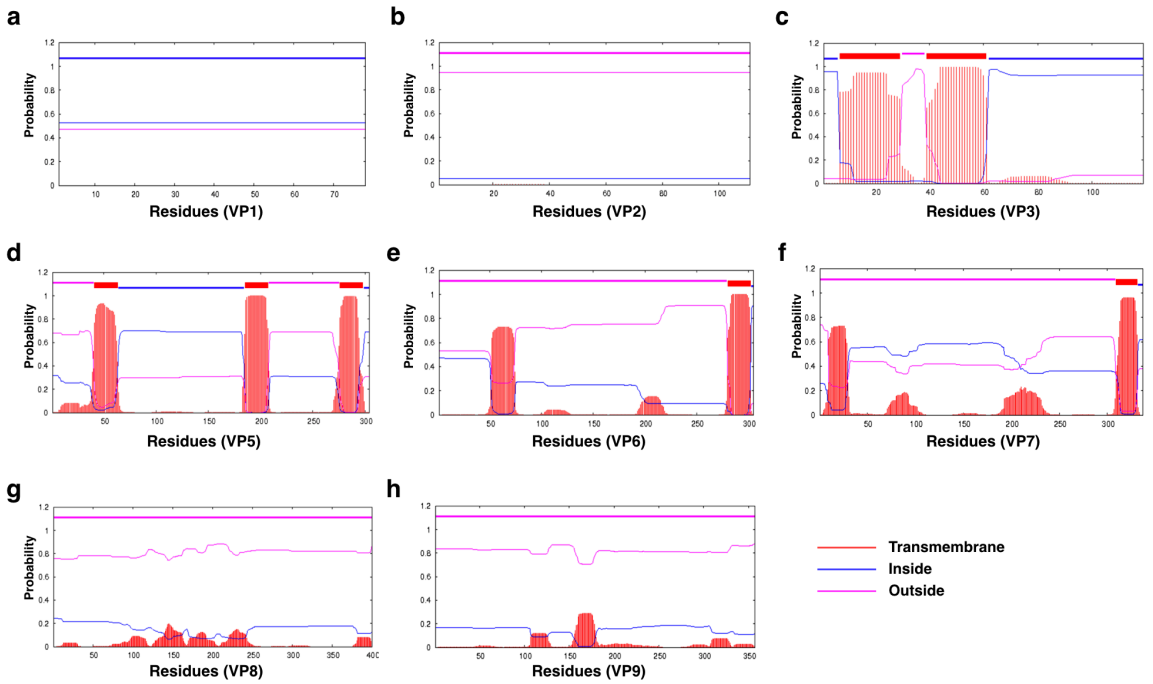

Fig. S6. Transmembrane helices in SPV1 proteins predicted by the TMHMM Server **a-h**, predicted transmembrane regions are indicated in red.

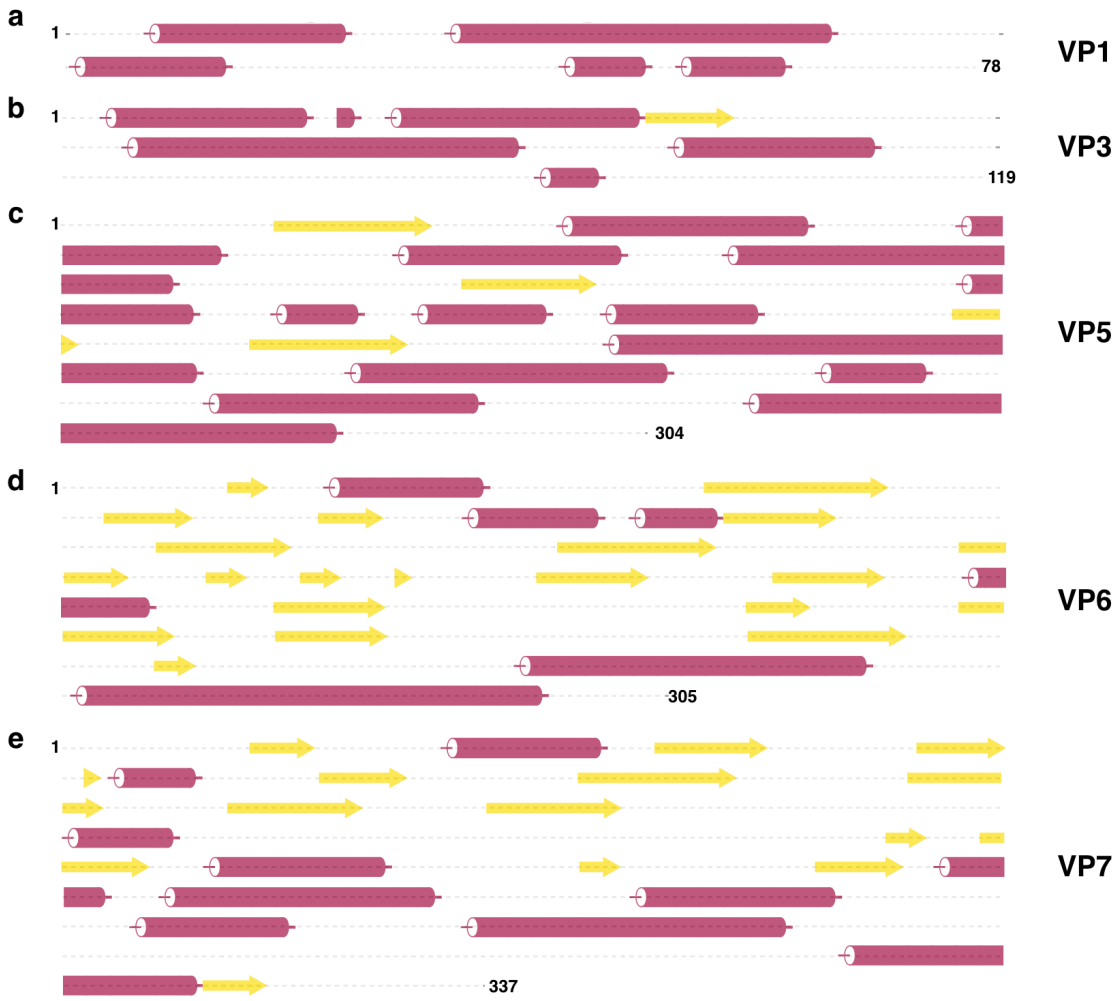

Fig. S7. Secondary structure predictions for SPV1 proteins (Jpred)

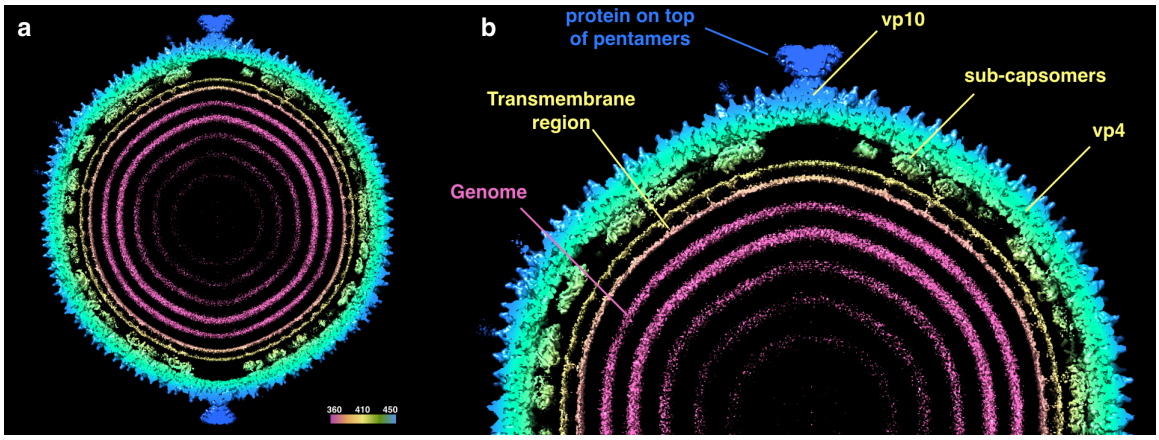

Fig. S8. Icosahedral reconstruction of SPV1

a, A central section of the icosahedral reconstruction.

b, A zoomed-in view of this reconstruction, with different components labeled.

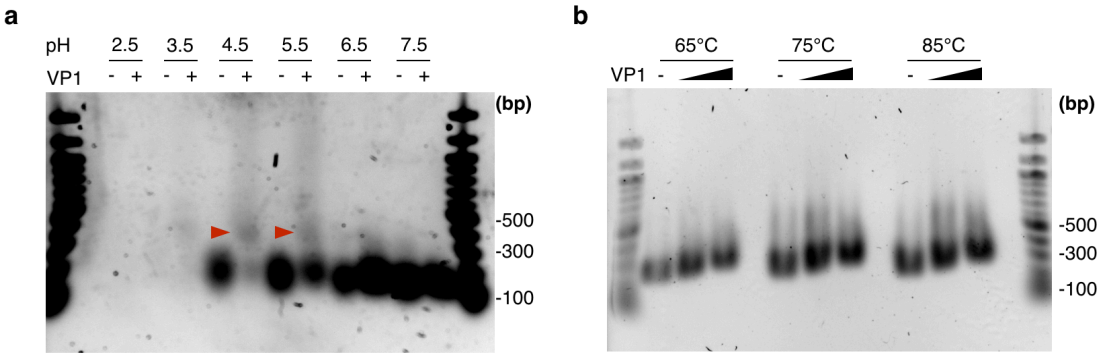

Fig. S9. Gel shift assay

a, pH gradient gel shift assay. The red arrowheads point to where a gel shift smear was observed.

b, gel shift assay done under different temperatures.

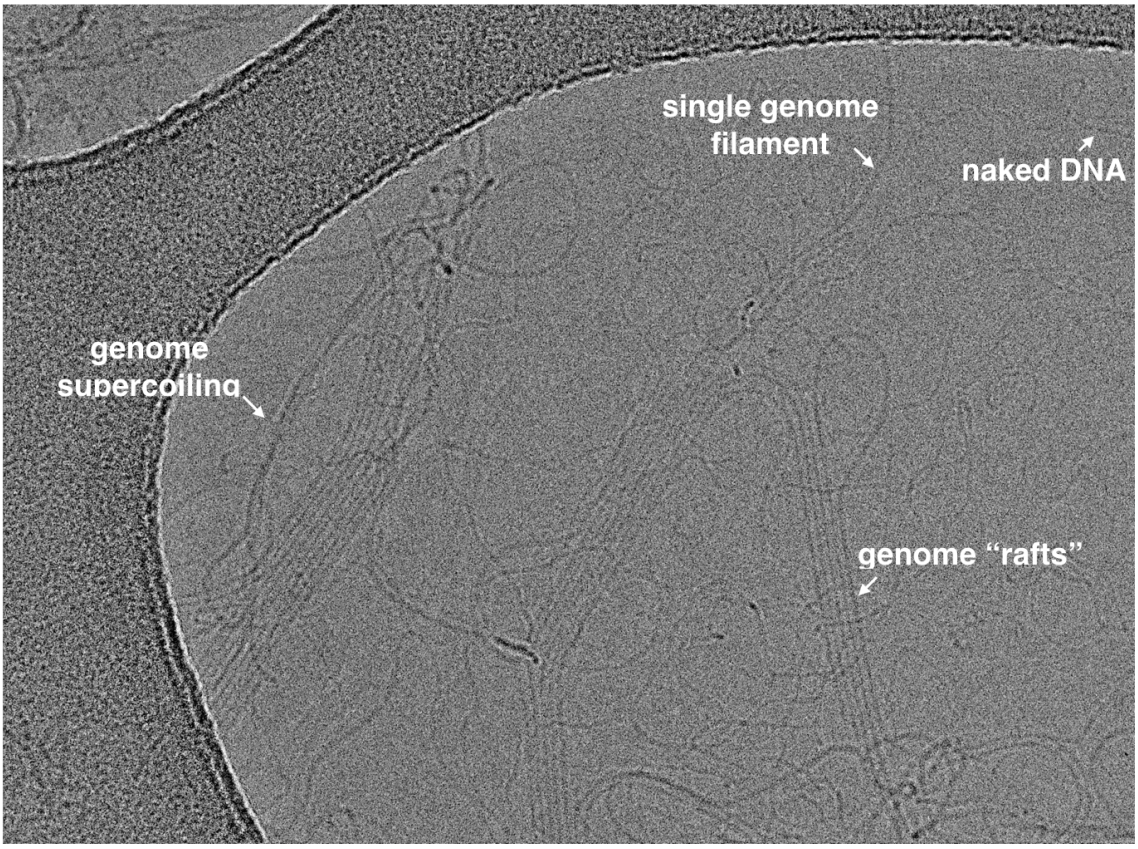

Fig. S10. Cryo-EM micrograph of the disrupted SPV1 genome with different types of filaments labeled.

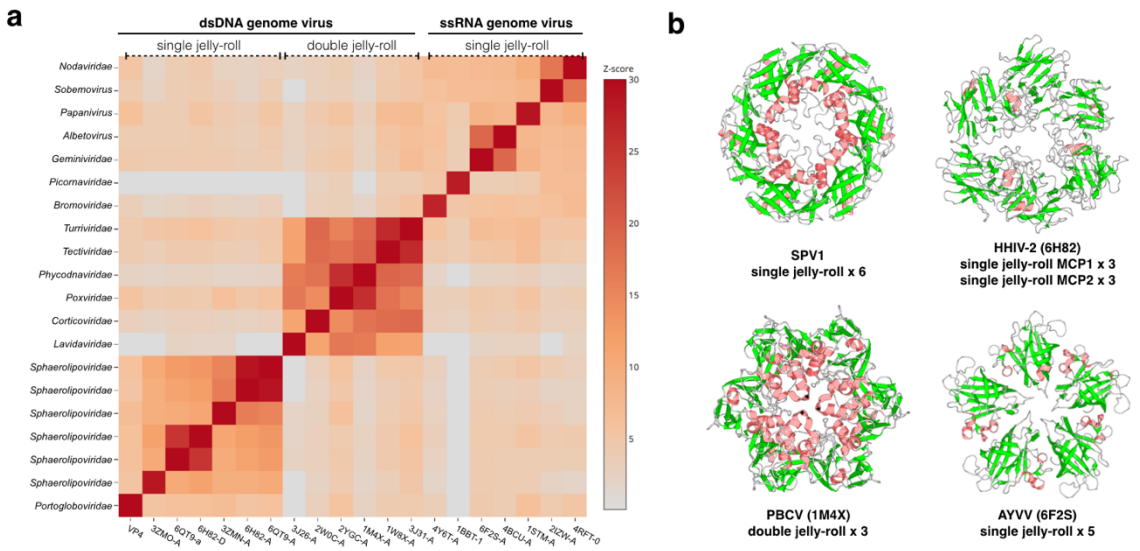

Fig. S11. Comparison of VP4 with other jelly-roll MCPs from other viruses.
a, All-against-all comparison of the VP4 and other major capsid proteins with single or double jelly-roll fold. The matrix is based on the pairwise z-score comparisons calculated using the Dali server. The color scale indicates the corresponding z-scores. Based on their functions and origins, the proteins are partitioned into three clusters. Virus classification is labeled on the left, and the PDB ID and chain used in this calculation are labeled at the bottom of the matrix.
b, Four representative capsomers from the DALI alignment: *Portogloboviridae* SPV1, *Sphaerolipoviridae* HHIV-2, *Phycodnaviridae* PBCV, and *Geminiviridae* AYVV. The capsomer structures are colored by secondary structure (β-sheet in green, α-helix in pink).

Table S1. Cryo-EM data collection, refinement and validation statistics

	Virus SPV1 model (EMDB-20083) (PDB 6OJ0)
Data collection and processing	
Magnification	59,000
Voltage (kV)	300
Electron exposure (e–/Å^2)	20
Defocus range (μm)	-1.5 to -2.5
Pixel size (Å)	1.4
Symmetry imposed	Icosahedral (I4)
Initial particle images (no.)	19,092
Final particle images (no.)	12,377
Map resolution (Å)	
model:map FSC (0.38)	3.7
"gold-standard" map:map FSC (0.143)	3.6
d_{99}	3.7
Refinement	
Initial model used	*de novo* C_α trace
Map sharpening *B* factor (Å^2)	-180
Real space correlation coefficient	0.83
Model composition	
Non-hydrogen atoms	57,765
Protein residues	7662
Ligands	N.A.
B factors (Å^2)	
Protein	43.7
Ligand	N.A.
R.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	0.956
Validation	
MolProbity score	1.98
Clashscore	3.5
Poor rotamers (%)	0.9
Ramachandran plot	
Favored (%)	90.0
Allowed (%)	9.9
Disallowed (%)	0.1

References

1. Liu Y, Ishino S, Ishino Y, Pehau-Arnaudet G, Krupovic M, & Prangishvili D (2017) A Novel Type of Polyhedral Viruses Infecting Hyperthermophilic Archaea. *J Virol* 91(13).
2. Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, & Scheres SH (2018) New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* 7.
3. Li X, Mooney P, Zheng S, Booth CR, Braunfeld MB, Gubbens S, Agard DA, & Cheng Y (2013) Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods* 10(6):584-590.
4. Zhang K (2016) Gctf: Real-time CTF determination and correction. *J Struct Biol* 193(1):1-12.
5. Afonine PV, Klaholz BP, Moriarty NW, Poon BK, Sobolev OV, Terwilliger TC, Adams PD, & Urzhumtsev A (2018) New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr D Struct Biol* 74(Pt 9):814-840.
6. Guo F & Jiang W (2014) Single particle cryo-electron microscopy and 3-D reconstruction of viruses. *Methods Mol Biol* 1117:401-443.
7. Mindell JA & Grigorieff N (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* 142(3):334-347.
8. Tang G, Peng L, Baldwin PR, Mann DS, Jiang W, Rees I, & Ludtke SJ (2007) EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol* 157(1):38-46.
9. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, & Ferrin TE (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605-1612.
10. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126-2132.
11. Wang RY, Kudryashev M, Li X, Egelman EH, Basler M, Cheng Y, Baker D, & DiMaio F (2015) De novo protein structure determination from near-atomic-resolution cryo-EM maps. *Nat Methods* 12(4):335-338.
12. Afonine PV, Poon BK, Read RJ, Sobolev OV, Terwilliger TC, Urzhumtsev A, & Adams PD (2018) Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol* 74(Pt 6):531-544.
13. Williams CJ*, et al.* (2018) MolProbity: More and better reference data for improved all-atom structure validation. *Protein Science* 27(1):293-315.
14. Schouten S, Huguet C, Hopmans EC, Kienhuis MV, & Damste JS (2007) Analytical methodology for TEX86 paleothermometry by high-performance liquid chromatography/atmospheric pressure chemical ionization-mass spectrometry. *Anal Chem* 79(7):2940-2944.