Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/6/36/eabc0023/DC1

Supplementary Materials for

Three-dimensional structure of 22 uncultured ssRNA bacteriophages: Flexibility of the coat protein fold and variations in particle shapes

Jānis Rūmnieks, Ilva Liekniņa, Gints Kalniņš, Mihails Šišovs, Ināra Akopjana, Jānis Bogans, Kaspars Tārs*

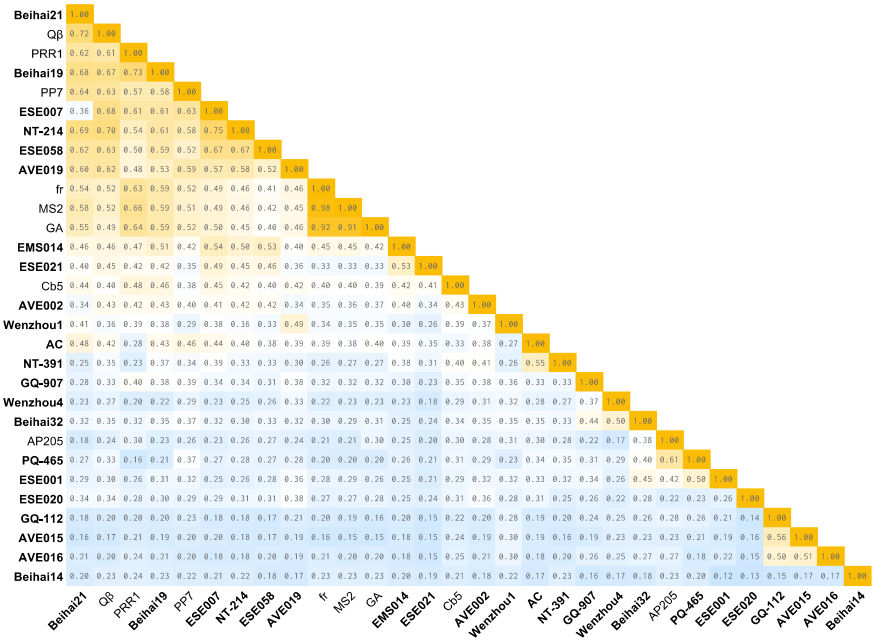
*Corresponding author. Email: kaspars@biomed.lu.lv

Published 2 September 2020, *Sci. Adv.* **6**, eabc0023 (2020)
DOI: 10.1126/sciadv.abc0023

This PDF file includes:

Figs. S1 to S5
Tables S1 and S2
References

A



B

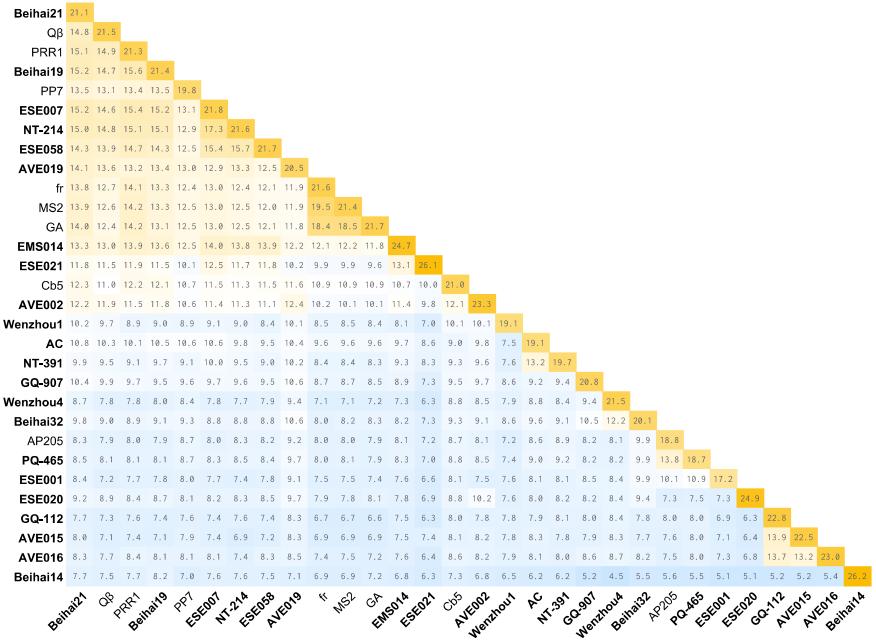
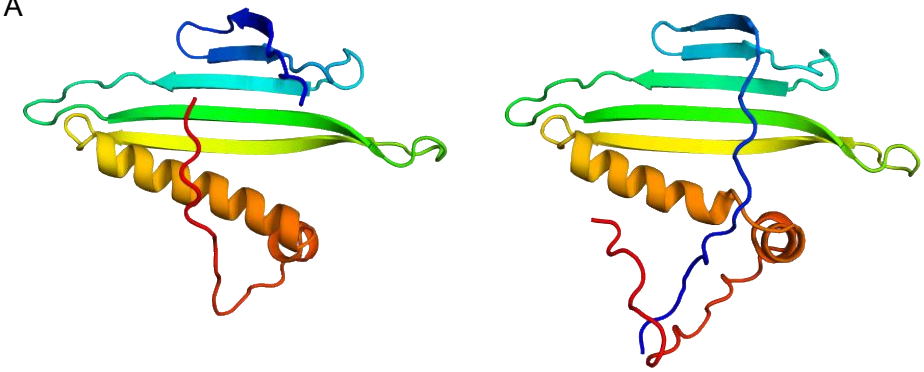
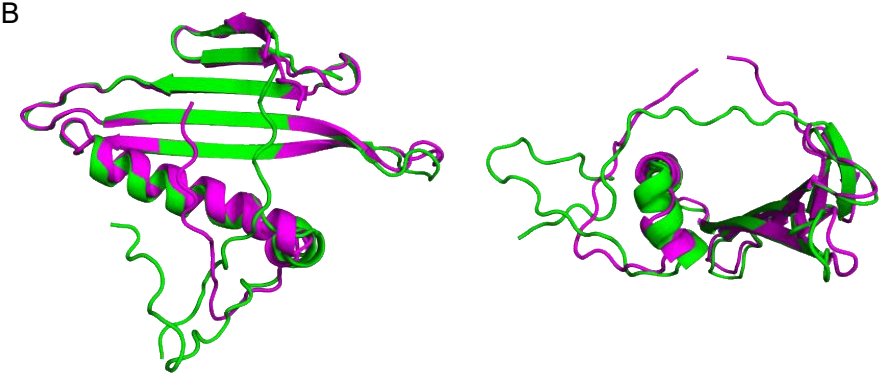


Figure S1. Structural similarity of ssRNA bacteriophage coat proteins. (A) Similarity analysis based on pairwise secondary structure matching using the program SUPERPOSE (49). The numerical values in the table represent the quality (Q) scores for each respective pairwise superposition. **(B)** All-against-all similarity analysis using the DALI server (50). The numerical values in the table correspond to the Z-scores for each pair. For both analyses the same dataset of quasi-equivalent C monomer coordinates from all currently available ssRNA phage structures was used. CP clustering was done manually. The CPs originating from the current study are shown in bold.

A



B



C

```
Beihai32        ------------SKPIAIFKLRELSSDSTLFTLPGHSVTLPNTLGIVSHLPTPRKGNPGT    48
Wenzhou4        SIKYIFKKTDTLPRSVIGNVLRTTGPDTTVYSLPGHTPVNPFTLTAVSRLPVPRKGNAGT    60
                             : :    **  . *:*:::****: . * **  **:**.***** **

Beihai32        VKTMRNLRKTILLGAGTASERAVPIVIKTETSFPVGTTEEDRAEVLKQMASFLIEEVKNN    108
Wenzhou4        TKTTLSLRREVTINKGTDQEKIVPMIARIETSVPVGVSQDDFKAMIEGLACPLLLDEIHV    120
                .**  .**: : :. ** .*: **:: : ***.***.:::*   ::: :*. *: :  :

Beihai32        QELAYSGYVQDKYFIEDLVITE----    130
Wenzhou4        NDLFLSGLPIATTDVPDNEPLPPALL    146
                ::*  **    .  : *   :
```

Figure S2. Structural diversity among Beihai32-like coat proteins. (A) Three-dimensional structure of Beihai32 (left) and Wenzhou4 (right) CPs. The proteins are rainbow-colored blue (N-terminus) to red (C-terminus). **(B)** Secondary structure-based superposition of Beihai32 and Wenzhou4 CPs. The superposed Beihai32 (magenta) and Wenzhou4 (green) CPs are shown in two different orientations. **(C)** Sequence alignment of the Beihai32 and Wenzhou4 CPs. β-strands are highlighted in pink and α-helices in yellow. Identical residues are marked by asterisks and similar residues by colons.

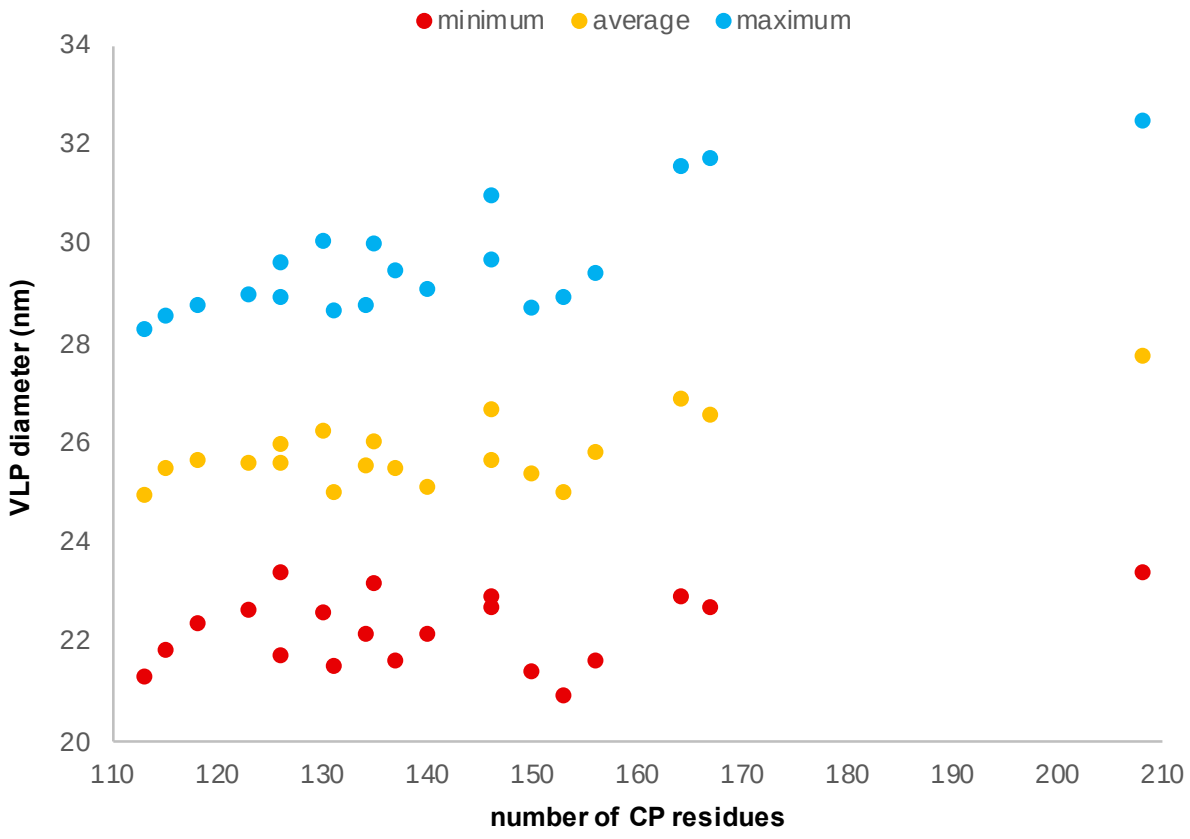


Figure S3. Relationship between coat protein length and particle size in the novel ssRNA phages. Data used for generating the graph are provided in Table S1.

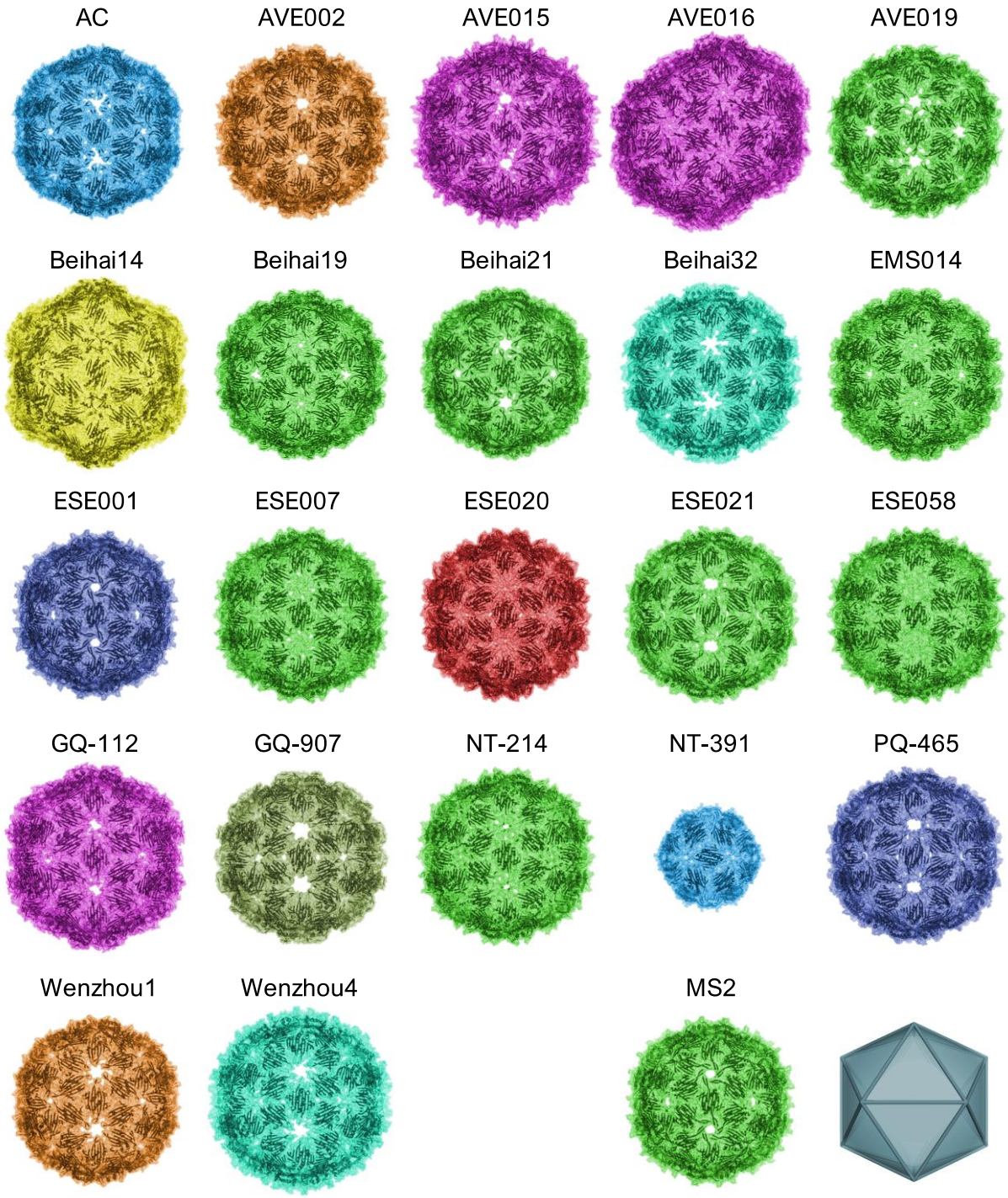
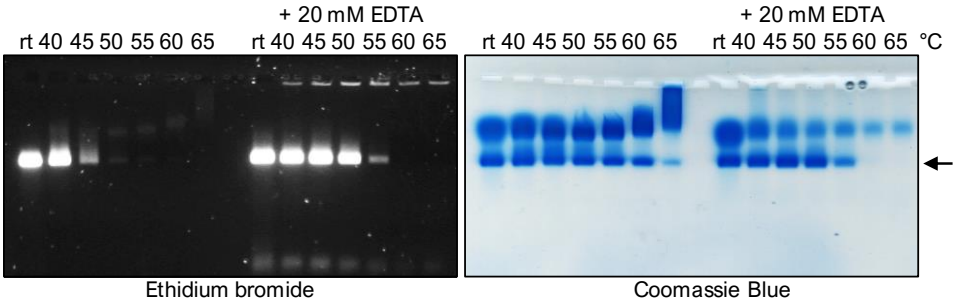


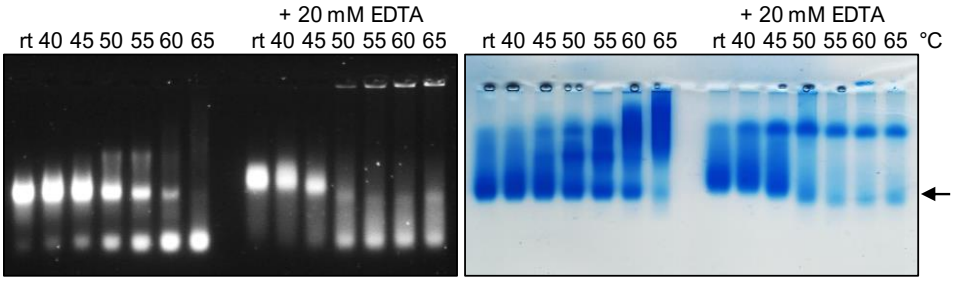
Figure S4. Structure of the novel ssRNA phage VLPs. Coat protein dimers are shown in cartoon representation inside semi-transparent VLP molecular surfaces and are differently colored as per different CP similarity groups. The back sides of the particles are clipped for clarity. All particles are shown on the same scale and in an orientation corresponding to the regular icosahedron on bottom right. Structure of the ssRNA bacteriophage MS2 is included for comparison.

AVE015

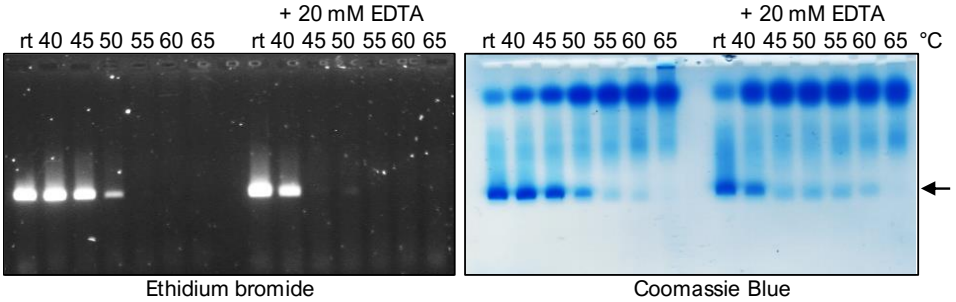


Ethidium bromide Coomassie Blue

AVE019

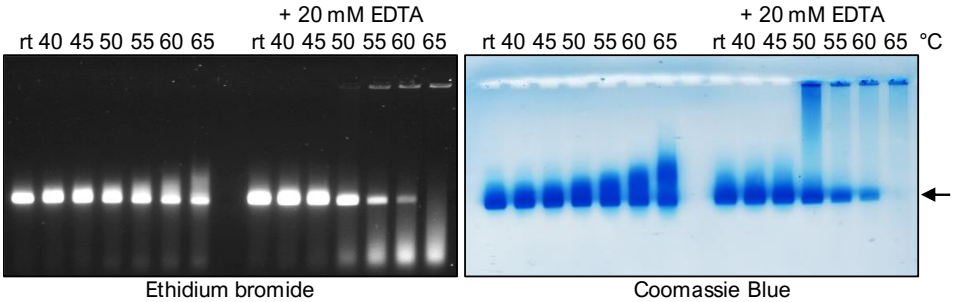


Beihai19



Ethidium bromide Coomassie Blue

Beihai32



Ethidium bromide Coomassie Blue

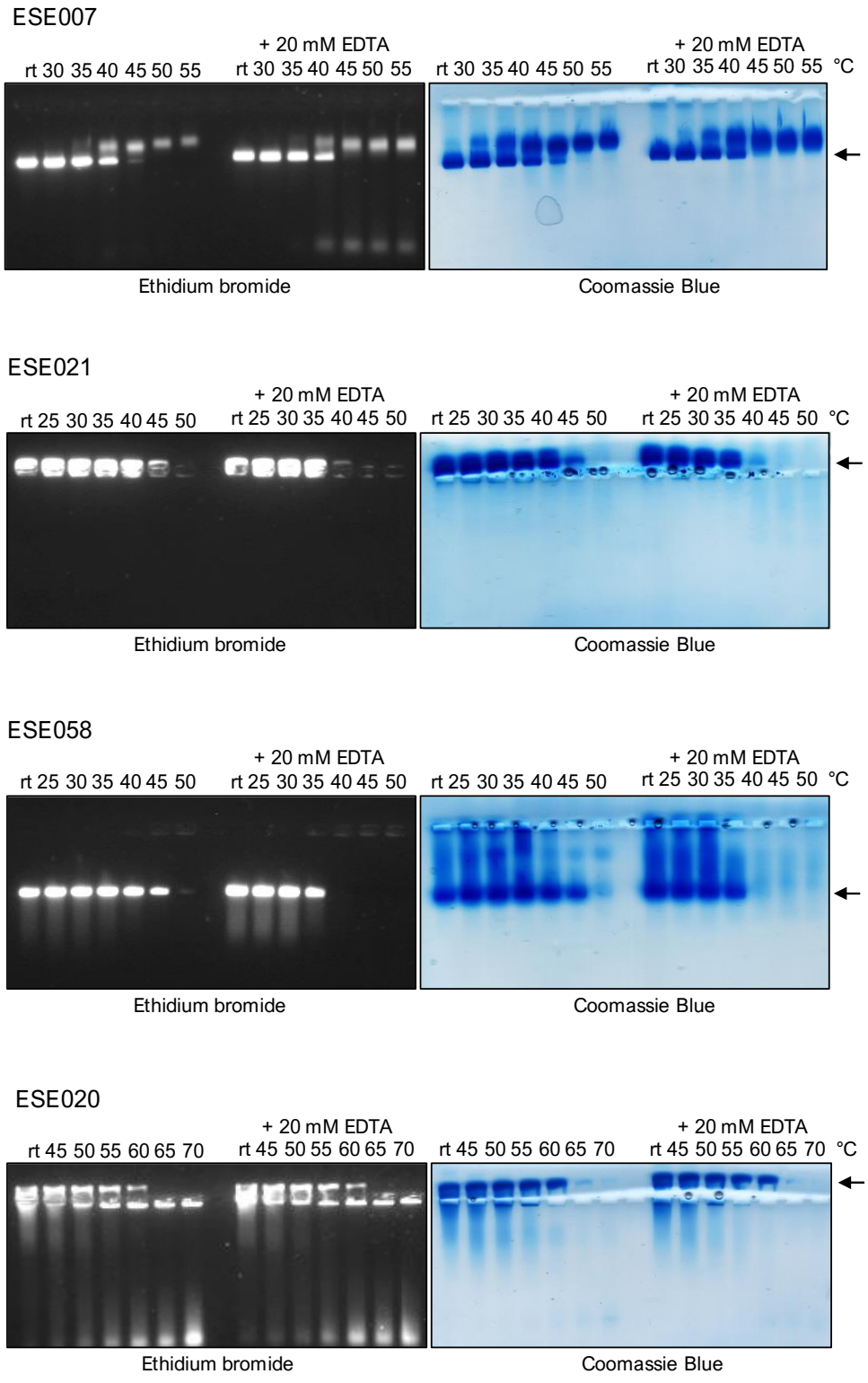


Figure S5. Stabilizing role of divalent metal ions in the novel ssRNA phage particles. Shown are agarose gels after electrophoresis of heated metal ion-containing virus-like particles in absence and in presence of EDTA. Temperature at which each sample was heated is provided above each track; rt: unheated control at room temperature. The arrow indicates the position of VLPs. ESE020 VLPs were used as a control for particles not containing bound metal ions.

name	GenBank accession	CP similarity group	CP length	VLP diameter (nm)			CP dimer interface area (Å^2)	
				minimum	average	maximum	AB	CC
AC	KF616864	AC	115	21.9	25.5	28.6	1872.7	1788.5
AVE002	KT462696	Cb5	140	22.2	25.1	29.1	2921.7	2722.1
AVE015		AVE015	167	22.7	26.6	31.8	3368.9	2689.9
AVE016		AVE015	166				4371.4	4114.0
AVE019		MS2	123	22.7	25.6	29.0	2353.9	2203.7
Beihai14	KX883502	Beihai14	208	23.4	27.8	32.5	5367.1	5293.4
Beihai19	KX883464	MS2	134	22.2	25.6	28.8	3093.1	3481.1
Beihai21	KX883506	MS2	126	21.7	25.6	29.0	2639.0	2458.2
Beihai32	KX883479	Beihai32	130	22.6	26.3	30.1	1997.6	1942.1
EMS014		MS2	156	21.6	25.9	29.4	3747.3	3315.9
ESE001	KT462710	AP205	118	22.4	25.7	28.8	2306.3	2486.9
ESE007		MS2	137	21.7	25.5	29.5	3971.7	3623.3
ESE020		ESE020	153	20.9	25.1	29.0	2692.3	2500.4
ESE021		MS2	150	21.4	25.4	28.7	4095.0	3629.4
ESE058		MS2	146	22.7	25.7	29.7	4669.7	4685.6
GQ-907	GALQ01034907	ESE017	131	22.9	26.9	31.6	2793.9	3026.7
GQ-112	GALQ01044112	AVE015	164	21.5	25.0	28.7	2634.4	2423.0
NT-214	NFYT01000214	MS2	135	23.2	26.0	30.0	3625.9	3468.0
NT-391	NFYT01000391	AC	123	12.1	15.2	18.3	1779.1	1778.0
PQ-465	PQDQ01001465	AP205	126	23.4	26.0	29.7	2582.3	2296.4
Wenzhou1	KX883612	Cb5	113	21.3	25.0	28.3	2342.2	2139.1
Wenzhou4	KX883621	Beihai32	146	22.9	26.7	31.0	1676.8	1927.0

Table S1. Properties of the ssRNA phage coat proteins and virus-like particles used in the current study. Accession numbers are provided for the respective genome sequences; entries without an accession number were sourced from Dataset S1 from (12). The VLP diameter was defined by calculating distances to all Cα atoms from the particle center and then finding the minimum, average and maximum values. The interface areas for CP dimers were calculated using PISA (51) and were defined as the buried surface between an AB or CC dimer and their four surrounding CP dimers in the particle.

	AC	AVE002	AVE015	AVE016
PDB ID	6YF7	6YF9	6YFA	6YFB
Crystallization				
Conditions	0.85 M ammonium sulfate	46% MPD, 0.1 M ammonium nitrate, 0.1 M MES pH 6.5	22% MPD, 0.1 M sodium nitrate, 0.1 M sodium acetate pH 4.6	30% MPD, 0.2 M sodium citrate, 0.1 M HEPES pH 7.5
Data collection				
Date	2015-11-05	2018-05-09	2018-05-09	2018-01-27
Site	MAX II	BESSY II	BESSY II	BESSY II
Beamline	I911-3	MX 14.1	MX 14.1	MX 14.1
Wavelength (Å)	0.97934	0.9184	0.9184	0.9184
Data reduction				
Space group	C 2	P 21 21 21	I 2 3	P 21
Cell parameters	a = 432.29 b = 307.87 c = 679.45 α = 90.00 β = 107.15 γ = 90.00	a = 278.09 b = 390.21 c = 554.47 α = 90.00 β = 90.00 γ = 90.00	a = 559.31 b = 559.31 c = 559.31 α = 90.00 β = 90.00 γ = 90.00	a = 325.47 b = 298.52 c = 417.12 α = 90.00 β = 100.00 γ = 90.00
Total number of observations	1069421	6155935	8135712	2073077
Number of unique reflections	591127	970502	430490	855258
Resolution (Å)	44.26 - 3.40	62.21 - 3.20	61.77 - 3.30	49.32 - 3.59
Highest resolution bin	3.58 - 3.40	3.37 - 3.20	3.48 - 3.30	3.81 - 3.59
Multiplicity	1.8 (1.7)	6.3 (5.9)	18.9 (18.6)	2.4 (2.4)
Completeness (%)	51.1 (47.3)	98.8 (97.0)	100.0 (100.0)	93.9 (92.4)
R-merge	0.198 (0.749)	0.672 (1.727)	0.702 (2.890)	0.471 (2.451)
Mean I/sig(I)	3.3 (1.0)	3.3 (1.1)	4.6 (1.1)	2.9 (0.4)
CC(1/2)	0.965 (0.455)	0.846 (0.190)	0.970 (0.412)	0.924 (0.189)
Wilson B-factor	32.2	38.5	66.5	75.7
Refinement				
Resolution	44.262 - 3.400	62.205 - 3.196	61.765 - 3.300	49.317 - 3.591
Highest resolution bin	3.521 - 3.400	3.310 - 3.196	3.418 - 3.300	3.720 - 3.591
Number of reflections				
work set	590873	969144	429748	845958
free set	29845	6518	9978	9894
R-work	0.2681 (0.4047)	0.2610 (0.3736)	0.2312 (0.3436)	0.2779 (0.3795)
R-free	0.2700 (0.4096)	0.2664 (0.3812)	0.2331 (0.3361)	0.2829 (0.3800)
Number of atoms				
protein	231390	184320	78960	275100
RNA				
water				
other			20	
Average B-factor				
protein	69.78	33.9	66.52	103.09
RNA				
water				
other			59.46	
rmsd from ideal geometry				
bonds (Å)	0.005	0.005	0.006	0.009
angles (°)	0.697	0.812	0.865	1.09
Ramachandran plot				
favored (%)	99.11	98.79	98.18	98.45
allowed (%)	100.00	100.00	100.00	100.00
Rotamer outliers (%)	0.34	0.00	0.00	0.71
Clashscore	10.67	8.47	8.65	15.62

	AVE019	Beihai14	Beihai19	Beihai21
PDB ID	6YFC	6YFD	6YFE	6YFF
Crystallization				
Conditions	46% MPD, 0.1 M HEPES pH 7.5	5% PEG 4000, 0.1 M magnesium chloride, 0.1 M MES pH 6.5	66% MPD, 100 mM Tris-HCl pH 8.0	30% PEG 400, 0.2 M sodium citrate, 0.1 M Tris-HCl pH 8.5
Data collection				
Date	2019-01-24	2019-03-22	2017-09-01	2018-01-27
Site	BESSY II	BESSY II	BESSY II	BESSY II
Beamline	MX 14.1	MX 14.1	MX 14.1	MX 14.1
Wavelength (Å)	0.9184	0.9184	0.9184	0.9184
Data reduction				
Space group	P 1	P 21 21 2	P 21 21 2	I 2 2 2
Cell parameters	a = 296.42 b = 277.20 c = 277.44 α = 103.91 β = 117.39 γ = 106.96	a = 328.67 b = 373.71 c = 310.50 α = 90.00 β = 90.00 γ = 90.00	a = 298.11 b = 325.05 c = 346.02 α = 90.00 β = 90.00 γ = 90.00	a = 296.94 b = 306.88 c = 322.56 α = 90.00 β = 90.00 γ = 90.00
Total number of observations	3314889	2846288	1102377	1499728
Number of unique reflections	912012	536242	323252	264751
Resolution (Å)	49.27 - 3.25	49.34 - 3.30	49.44 - 3.79	48.94 - 3.09
Highest resolution bin	3.44 - 3.25	3.50 - 3.30	4.02 - 3.79	3.28 - 3.09
Multiplicity	3.6 (3.5)	5.3 (2.5)	3.4 (3.4)	5.7 (5.5)
Completeness (%)	94.0 (88.0)	94.4 (76.8)	98.0 (95.3)	99.1 (95.4)
R-merge	0.372 (2.509)	0.647 (2.255)	0.372 (1.495)	0.358 (2.785)
Mean I/sig(I)	4.3 (0.5)	2.3 (0.5)	3.6 (0.8)	5.6 (0.6)
CC(1/2)	0.988 (0.690)	0.937 (0.171)	0.965 (0.243)	0.985 (0.221)
Wilson B-factor	69.0	81.1	83.5	66.5
Refinement				
Resolution	49.267 - 3.246	49.332 - 3.300	41.392 - 3.794	48.938 - 3.089
Highest resolution bin	3.362 - 3.246	3.418 - 3.300	3.930 - 3.794	3.200 - 3.089
Number of reflections				
work set	898220	525682	322856	263402
free set	9890	9824	9968	9929
R-work	0.2413 (0.3928)	0.2649 (0.3861)	0.2609 (0.3706)	0.2661 (0.3845)
R-free	0.2393 (0.3869)	0.2670 (0.3925)	0.2668 (0.3807)	0.2710 (0.3790)
Number of atoms				
protein	169200	126960	89040	43650
RNA				
water				
other	60		30	
Average B-factor				
protein	97.39	92.15	107.17	80.85
RNA				
water				
other	64.63		53.34	
rmsd from ideal geometry				
bonds (Å)	0.005	0.006	0.005	0.005
angles (°)	0.786	0.848	0.74	0.794
Ramachandran plot				
favored (%)	98.88	98.06	98.43	98.12
allowed (%)	100.00	100.00	100.00	100.00
Rotamer outliers (%)	0.63	0.00	0.00	0.00
Clashscore	10.22	7.93	12.21	9.82

	Beihai32	EMS014 (VLP)	EMS014 (subunit)	ESE001
PDB ID	6YFG	6YFH	6YFI	6YFJ
Crystallization				
Conditions	26% MPD, 0.1-0.2 M sodium nitrate, 0.1 M HEPES pH 7.5	0.55 M sodium malonate, 0.25% Jeffamine ED-2003, 0.05 M HEPES pH 7.0	10% PEG 8000, 0.1 M sodium chloride, 0.05 M phosphate/citrate pH 4.2	30% MPD, 5 mM DTT, 0.1 M sodium acetate pH 4.6
Data collection				
Date	2018-10-19	2018-05-09	2018-05-09	2017-09-01
Site	MAX IV	BESSY II	BESSY II	BESSY II
Beamline	BioMAX	MX 14.1	MX 14.1	MX 14.1
Wavelength (Å)	0.97776	0.9184	0.9184	0.9184
Data reduction				
Space group	P 1	C 2 2 21	P 21 21 21	R 3
Cell parameters	a = 291.95 b = 292.53 c = 469.20 α = 75.79 β = 77.92 γ = 69.51	a = 287.11 b = 492.50 c = 553.71 α = 90.00 β = 90.00 γ = 90.00	a = 71.56 b = 92.70 c = 94.82 α = 90.00 β = 90.00 γ = 90.00	a = 283.35 b = 283.35 c = 666.04 α = 90.00 β = 90.00 γ = 120.00
Total number of observations	2265552	1622482	1109958	1123410
Number of unique reflections	1227697	352052	173303	312316
Resolution (Å)	49.84 - 3.90	49.41 - 3.89	48.63 - 1.25	48.95 - 3.23
Highest resolution bin	4.13 - 3.90	4.13 - 3.89	1.32 - 1.25	3.43 - 3.23
Multiplicity	1.8 (1.9)	4.6 (4.3)	6.4 (6.0)	3.6 (3.4)
Completeness (%)	96.2 (93.0)	99.1 (96.7)	99.1 (94.7)	97.9 (93.6)
R-merge	0.209 (1.559)	0.859 (3.144)	0.085 (1.054)	0.284 (1.738)
Mean I/sig(I)	3.0 (0.4)	2.3 (0.5)	8.5 (1.0)	4.4 (0.6)
CC(1/2)	0.985 (0.151)	0.836 (0.136)	0.998 (0.825)	0.984 (0.235)
Wilson B-factor	118.2	76.7	21.6	76.2
Refinement				
Resolution	49.833 - 3.897	49.357 - 3.893	42.210 - 1.248	48.944 - 3.233
Highest resolution bin	4.037 - 3.897	4.032 - 3.893	1.293 - 1.248	3.348 - 3.233
Number of reflections				
work set	1216193	346121	172802	310956
free set	19851	9855	8618	10007
R-work	0.2821 (0.3974)	0.2665 (0.3785)	0.1753 (0.3683)	0.2317 (0.3755)
R-free	0.2853 (0.4048)	0.2699 (0.3696)	0.1974 (0.3796)	0.2324 (0.3808)
Number of atoms				
protein	362880	104400	4450	52380
RNA				
water			1076	
other	120			
Average B-factor				
protein	170.05	111.97	22.13	97.97
RNA				
water			33.61	
other	148.62			
rmsd from ideal geometry				
bonds (Å)	0.006	0.004	0.005	0.006
angles (°)	0.870	0.697	0.762	0.826
Ramachandran plot				
favored (%)	98.18	99.13	98.09	99.14
allowed (%)	100.00	100.00	100.00	100.00
Rotamer outliers (%)	0.00	0.00	0.00	0.00
Clashscore	14.38	5.62	1.45	8.54

	ESE007	ESE020	ESE021	ESE058
PDB ID	6YFK	6YFL	6YFM	6YFN
Crystallization				
Conditions	20 % PEG 10000, 0.1 M HEPES pH 7.5	36% PEG 300, 0.05 M Bicine pH 9.0	5% PEG 3000, 0.08 M zinc acetate, 0.05 M sodium acetate pH 4.6	12.5% PEG 1000, 7% PEG 8000
Data collection				
Date	2018-01-27	2017-11-04	2018-05-09	2017-11-04
Site	BESSY II	BESSY II	BESSY II	BESSY II
Beamline	MX 14.1	MX 14.1	MX 14.1	MX 14.1
Wavelength (Å)	0.9184	0.9184	0.9184	0.9184
Data reduction				
Space group	R 3	R 3 2	I 2 2 2	P 32 2 1
Cell parameters	a = 278.20 b = 278.20 c = 661.50 α = 90.00 β = 90.00 γ = 120.00	a = 419.73 b = 419.73 c = 761.14 α = 90.00 β = 90.00 γ = 120.00	a = 282.02 b = 302.81 c = 352.16 α = 90.00 β = 90.00 γ = 90.00	a = 500.89 b = 500.89 c = 287.05 α = 90.00 β = 90.00 γ = 120.00
Total number of observations	338161	1819911	2023525	6181589
Number of unique reflections	132293	380562	378707	679725
Resolution (Å)	38.01 - 3.70	99.46 - 3.30	60.08 - 2.76	49.26 - 3.19
Highest resolution bin	3.90 - 3.70	3.48 - 3.30	2.91 - 2.76	3.38 - 3.19
Multiplicity	2.6 (1.2)	4.8 (4.1)	5.3 (5.6)	9.1 (9.1)
Completeness (%)	64.9 (28.5)	99.5 (98.9)	99.4 (99.8)	99.8 (98.8)
R-merge	0.218 (0.817)	0.417 (2.056)	0.362 (1.532)	0.796 (3.730)
Mean I/sig(I)	3.2 (0.6)	4.3 (0.7)	4.7 (1.1)	3.7 (0.6)
CC(1/2)	0.968 (0.152)	0.952 (0.152)	0.957 (0.293)	0.958 (0.234)
Wilson B-factor	68.8	74.9	33.7	53.1
Refinement				
Resolution	38.003 - 3.700	96.967 - 3.300	60.079 - 2.762	49.258 - 3.189
Highest resolution bin	3.832 - 3.700	3.418 - 3.300	2.861 - 2.762	3.303 - 3.189
Number of reflections				
work set	131994	379266	378514	677292
free set	5027	9971	9967	9955
R-work	0.2351 (0.3812)	0.2477 (0.3936)	0.2445 (0.3695)	0.2308 (0.3770)
R-free	0.2375 (0.3643)	0.2506 (0.3825)	0.2480 (0.3645)	0.2342 (0.3802)
Number of atoms				
protein	62160	63520	50130	101460
RNA				
water				
other	20		60	30
Average B-factor				
protein	84.62	71.17	36.12	60.45
RNA				
water				
other	51.49		154.19	78.52
rmsd from ideal geometry				
bonds (Å)	0.006	0.005	0.004	0.005
angles (°)	0.845	0.803	0.589	0.764
Ramachandran plot				
favored (%)	98.02	98.25	98.87	99.53
allowed (%)	100.00	100.00	100.00	100.00
Rotamer outliers (%)	0.30	0.58	0.27	0.00
Clashscore	13.43	11.98	4.02	7.66

	GQ-907	GQ-112	NT-214	NT-391
PDB ID	6YFO	6YFP	6YFQ	6YFR
Crystallization				
Conditions	12.5 % PEG 3350, 0.05 M Bis-tris pH 5.5	0.615 M sodium succinate pH 7.0	30% MPD, 0.02 M calcium chloride, 0.1 M sodium acetate pH 4.6	34% MPD, 0.1 M sodium nitrate, 0.1M sodium acetate pH 4.6
Data collection				
Date	2019-01-24	2019-03-22	2018-10-19	2019-01-24
Site	BESSY II	BESSY II	MAX IV	BESSY II
Beamline	MX 14.1	MX 14.1	BioMAX	MX 14.1
Wavelength (Å)	0.9184	0.9184	0.97776	0.9184
Data reduction				
Space group	C 2	C 2	R 3 2	P 1
Cell parameters	a = 415.01 b = 335.00 c = 291.78 α = 90.00 β = 134.66 γ = 90.00	a = 469.42 b = 332.41 c = 293.53 α = 90.00 β = 127.87 γ = 90.00	a = 341.90 b = 341.90 c = 1253.60 α = 90.00 β = 90.00 γ = 120.00	a = 161.92 b = 162.07 c = 162.98 α = 71.16 β = 66.67 γ = 66.83
Total number of observations	1226838	1699944	5672505	690098
Number of unique reflections	353773	508294	275564	294001
Resolution (Å)	49.14 - 3.48	57.21 - 3.10	49.97 - 3.80	51.89 - 2.60
Highest resolution bin	3.69 - 3.48	3.27 - 3.10	4.01 - 3.80	2.74 - 2.60
Multiplicity	3.5 (3.5)	3.3 (2.0)	20.6 (19.2)	2.3 (2.1)
Completeness (%)	97.9 (94.2)	79.5 (35.7)	98.3 (90.0)	69.7 (56.2)
R-merge	0.449 (2.461)	0.310 (1.139)	0.597 (2.764)	0.239 (1.134)
Mean I/sig(I)	3.2 (0.5)	3.4 (0.7)	4.7 (1.0)	2.5 (0.5)
CC(1/2)	0.966 (0.180)	0.753 (0.254)	0.993 (0.467)	0.927 (0.102)
Wilson B-factor	83.2	46.4	96.6	46.7
Refinement				
Resolution	49.140 - 3.484	57.200 - 3.100	49.955 - 3.800	51.880 - 2.600
Highest resolution bin	3.609 - 3.484	3.211 - 3.100	3.936 - 3.800	2.693 - 2.600
Number of reflections				
work set	347398	502524	275362	286598
free set	9832	9844	4997	9703
R-work	0.2685 (0.3989)	0.2500 (0.3705)	0.2199 (0.3775)	0.3175 (0.4009)
R-free	0.2710 (0.3925)	0.2537 (0.3828)	0.2220 (0.3979)	0.3231 (0.3985)
Number of atoms				
protein	87480	116370	61200	56340
RNA				
water				
other				
Average B-factor				
protein	110.95	43.41	135.8	50.25
RNA				
water				
other				
rmsd from ideal geometry				
bonds (Å)	0.005	0.006	0.007	0.008
angles (°)	0.87	0.905	0.817	0.685
Ramachandran plot				
favored (%)	99.22	98.97	98.50	100.00
allowed (%)	100.00	100.00	100.00	100.00
Rotamer outliers (%)	0.00	0.00	0.00	0.00
Clashscore	11.77	7.84	12.14	2.73

	PQ-465	Wenzhou1	Wenzhou4
PDB ID	6YFS	6YFT	6YFU
Crystallization			
Conditions	38% MPD, 0.1 M sodium acetate pH 4.6	10 % v/v ethanol, 0.1 M Tris-HCl pH 8.5	30 % MPD, 0.1 M MES pH 6.5
Data collection			
Date	2019-06-08	2019-01-24	2018-10-19
Site	BESSY II	BESSY II	MAX IV
Beamline	MX 14.1	MX 14.1	BioMAX
Wavelength (Å)	0.9184	0.9184	0.97776
Data reduction			
Space group	C 2	P 1	I 2 3
Cell parameters	a = 453.15 b = 309.76 c = 294.60 α = 90.00 β = 130.08 γ = 90.00	a = 277.47 b = 396.64 c = 399.87 α = 69.66 β = 83.65 γ = 83.49	a = 551.80 b = 551.80 c = 551.80 α = 90.00 β = 90.00 γ = 90.00
Total number of observations	1350175	3182922	9650681
Number of unique reflections	382128	1749831	229056
Resolution (Å)	49.35 - 3.50	61.16 - 3.50	49.96 - 4.02
Highest resolution bin	3.71 - 3.50	3.69 - 3.50	4.24 - 4.00
Multiplicity	3.5 (3.5)	1.8 (1.8)	42.1 (40.0)
Completeness (%)	97.5 (96.3)	87.6 (89.9)	98.3 (90.1)
R-merge	0.473 (4.064)	0.307 (1.481)	0.882 (6.402)
Mean I/sig(I)	3.5 (0.3)	2.6 (0.5)	5.2 (0.6)
CC(1/2)	0.970 (0.110)	0.920 (0.117)	0.994 (0.201)
Wilson B-factor	98.8	71.1	127.6
Refinement			
Resolution	49.343 - 3.498	61.150 - 3.500	49.958 - 4.018
Highest resolution bin	3.623 - 3.498	3.625 - 3.500	4.162 - 4.018
Number of reflections			
work set	357164	1741933	227552
free set	9293	19893	4953
R-work	0.2522 (0.3874)	0.2612 (0.3894)	0.2424 (0.3710)
R-free	0.2593 (0.3856)	0.2606 (0.3897)	0.2498 (0.3832)
Number of atoms			
protein	84900	302040	66540
RNA		25200	
water			
other			
Average B-factor			
protein	120.8	93.43	161.39
RNA		186.53	
water			
other			
rmsd from ideal geometry			
bonds (Å)	0.005	0.008	0.005
angles (°)	0.755	0.955	0.994
Ramachandran plot			
favored (%)	98.09	98.50	98.61
allowed (%)	100.00	100.00	100.00
Rotamer outliers (%)	0.00	0.00	0.00
Clashscore	11.85	4.57	15.19

Table S2. Crystallographic data collection, reduction and refinement statistics. Values in parentheses correspond to the highest resolution bin.

REFERENCES AND NOTES

1. P. Pumpens, *Single-Stranded RNA Phages: From Molecular Biology to Nanotechnology* (CRC Press, 2020).

2. Y. I. Wolf, D. Kazlauskas, J. Iranzo, A. Lucía-Sanz, J. H. Kuhn, M. Krupovic, V. V. Dolja, E. V. Koonin, Origins and evolution of the global RNA virome. *mBio* **9**, e02329-18 (2018).

3. K. Valegård, L. Liljas, K. Fridborg, T. Unge, The three-dimensional structure of the bacterial virus MS2. *Nature* **345**, 36–41 (1990).

4. X. Dai, Z. Li, M. Lai, S. Shu, Y. Du, Z. H. Zhou, R. Sun, In situ structures of the genome and genome-delivery apparatus in a single-stranded RNA virus. *Nature* **541**, 112–116 (2017).

5. L. Liljas, K. Fridborg, K. Valegård, M. Bundule, P. Pumpens, Crystal structure of bacteriophage fr capsids at 3.5 Å resolution. *J. Mol. Biol.* **244**, 279–290 (1994).

6. R. Golmohammadi, K. Fridborg, M. Bundule, K. Valegård, L. Liljas, The crystal structure of bacteriophage Qβ at 3.5 Å resolution. *Structure* **4**, 543–554 (1996).

7. K. Tars, M. Bundule, K. Fridborg, L. Liljas, The crystal structure of bacteriophage GA and a comparison of bacteriophages belonging to the major groups of *Escherichia coli* leviviruses. *J. Mol. Biol.* **271**, 759–773 (1997).

8. K. Tars, K. Fridborg, M. Bundule, L. Liljas, The three-dimensional structure of bacteriophage PP7 from *Pseudomonas aeruginosa* at 3.7-Å resolution. *Virology* **272**, 331–337 (2000).

9. M. Persson, K. Tars, L. Liljas, The capsid of the small RNA phage PRR1 is stabilized by metal ions. *J. Mol. Biol.* **383**, 914–922 (2008).

10. P. Plevka, A. Kazaks, T. Voronkova, S. Kotelovica, A. Dishlers, L. Liljas, K. Tars, The structure of bacteriophage φCb5 reveals a role of the RNA genome and metal ions in particle stability and assembly. *J. Mol. Biol.* **391**, 635–647 (2009).

11. M. Shishovs, J. Rumnieks, C. Diebolder, K. Jaudzems, L. B. Andreas, J. Stanek, A. Kazaks, S. Kotelovica, I. Akopjana, G. Pintacuda, R. I. Koning, K. Tars, Structure of AP205 coat protein reveals circular permutation in ssRNA bacteriophages. *J. Mol. Biol.* **428**, 4267–4279 (2016).

12. S. R. Krishnamurthy, A. B. Janowski, G. Zhao, D. Barouch, D. Wang, Hyperexpansion of RNA bacteriophage diversity. *PLOS Biol.* **14**, e1002409 (2016).

13. M. Shi, X.-D. Lin, J.-H. Tian, L.-J. Chen, X. Chen, C.-X. Li, X.-C. Qin, J. Li, J.-P. Cao, J.-S. Eden, J. Buchmann, W. Wang, J. Xu, E.C. Holmes, Y.-Z. Zhang, Redefining the invertebrate RNA virosphere. *Nature* **540**, 539–543 (2016).

14. E. P. Starr, E. E. Nuccio, J. Pett-Ridge, J. F. Banfield, M. K. Firestone, Metatranscriptomic reconstruction reveals RNA viruses with the potential to shape carbon cycling in soil. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 25900–25908 (2019).

15. J. Callanan, S. R. Stockdale, A. Shkoporov, L. A. Draper, R. P. Ross, C. Hill, Expansion of known ssRNA phage genomes: From tens to over a thousand. *Sci. Adv.* **6**, eaay5981 (2020).

16. I. Liekniņa, G. Kalniņš, I. Akopjana, J. Bogans, M. Šišovs, J. Jansons, J. Rūmnieks, K. Tārs, Production and characterization of novel ssRNA bacteriophage virus-like particles from metagenomic sequencing data. *J. Nanobiotech.* **17**, 61 (2019).

17. P. Pumpens, R. Renhofa, A. Dishlers, T. Kozlovska, V. Ose, P. Pushko, K. Tars, E. Grens, M.F. Bachmann, The true story and advantages of RNA phage capsids as nanotools. *Intervirology* **59**, 74–110 (2016).

18. S. C. Harrison, A. J. Olson, C. E. Schutt, F. K. Winkler, G. Bricogne, Tomato bushy stunt virus at 2.9 Å resolution. *Nature* **276**, 368–373 (1978).

19. K. Tars, A. Zeltins, L. Liljas, The three-dimensional structure of cocksfoot mottle virus at 2.7 Å resolution. *Virology* **310**, 287–297 (2003).

20. P. Plevka, K. Tars, A. Zeltins, I. Balke, E. Truve, L. Liljas, The three-dimensional structure of ryegrass mottle virus at 2.9 A resolution. *Virology* **369**, 364–374 (2007).

21. M. A. Asensio, N. M. Morella, C. M. Jakobson, E. C. Hartman, J. E. Glasgow, B. Sankaran, P. H. Zwart, D. Tullman-Ercek, A selection for assembly reveals that a single amino acid mutant of the bacteriophage MS2 coat protein forms a smaller virus-like particle. *Nano Lett.* **16**, 5944–5950 (2016).

22. A. Luque, D. Reguera, The structure of elongated viral capsids. *Biophys. J.* **98**, 2993–3003 (2010).

23. I. Cielens, V. Ose, I. Petrovskis, A. Strelnikova, R. Renhofa, T. Kozlovska, P. Pumpens, Mutilation of RNA phage Qβ virus-like particles: From icosahedrons to rods. *FEBS Lett.* **482**, 261–264 (2000).

24. J. Rumnieks, V. Ose, K. Tars, A. Dislers, A. Strods, I. Cielens, R. Renhofa, Assembly of mixed rod-like and spherical particles from group I and II RNA bacteriophage coat proteins. *Virology* **391**, 187–194 (2009).

25. M. F. Moody, The shape of the T-even bacteriophage head. *Virology* **26**, 567–576 (1965).

26. A. Sicard, Y. Michalakis, S. Gutiérrez, S. Blanc, The strange lifestyle of multipartite viruses. *PLOS Pathog.* **12**, e1005819 (2016).

27. K. Valegård, J. B. Murray, P. G. Stockley, N. J. Stonehouse, L. Liljas, Crystal structure of an RNA bacteriophage coat protein–operator complex. *Nature* **371**, 623–626 (1994).

28. M. Persson, K. Tars, L. Liljas, PRR1 coat protein binding to its RNA translational operator. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 367–372 (2013).

29. J. Rumnieks, K. Tars, Crystal structure of the bacteriophage Qβ coat protein in complex with the RNA operator of the replicase gene. *J. Mol. Biol.* **426**, 1039–1049 (2014).

30. J. A. Chao, Y. Patskovsky, S. C. Almo, R. H. Singer, Structural basis for the coevolution of a viral RNA–protein complex. *Nat. Struct. Mol. Biol.* **15**, 103–105 (2008).

31. K. V. Gorzelnik, Z. Cui, C. A. Reed, J. Jakana, R. Young, J. Zhang, Asymmetric cryo-EM structure of the canonical *Allolevivirus* Qβ reveals a single maturation protein and the genomic ssRNA in situ. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 11519–11524 (2016).

32. Ó. Rolfsson, S. Middleton, I. W. Manfield, S. J. White, B. Fan, R. Vaughan, N. A. Ranson, E. Dykeman, R. Twarock, J. Ford, C. C. Kao, P. G. Stockley, Direct evidence for packaging signal-mediated assembly of bacteriophage MS2. *J. Mol. Biol.* **428**, 431–448 (2016).

33. D. J. Klein, P. B. Moore, T. A. Steitz, The roles of ribosomal proteins in the structure assembly, and evolution of the large ribosomal subunit. *J. Mol. Biol.* **340**, 141–177 (2004).

34. A. L. N. Rao, Genome packaging by spherical plant RNA viruses. *Annu. Rev. Phytopathol.* **44**, 61–87 (2006).

35. M. Krupovic, E. V. Koonin, Multiple origins of viral capsid proteins from cellular ancestors. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E2401–E2410 (2017).

36. T. G. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell, A. G. W. Leslie, *iMOSFLM*: A new graphical interface for diffraction-image processing with *MOSFLM*. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 271–281 (2011).

37. P. R. Evans, Scala. Joint CCP4 + ESF-EAMCB. *Newslett. Protein Crystallogr.* **33**, 22–24 (1997).

38. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin, K. S. Wilson, Overview of the *CCP4* suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).

39. W. Kabsch, XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).

40. K. M. Sparta, M. Krug, U. Heinemann, U. Mueller, M. S. Weiss, XDSAPP2.0. *J. Appl. Crystallogr.* **49**, 1085–1092 (2016).

41. L. Tong, M. G. Rossmann, Rotation function calculations with GLRF program. *Methods Enzymol.* **276**, 594–611 (1997).

42. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *Phaser crystallographic* software. *J. Appl. Cryst.* **40**, 658–674 (2007).

43. G. J. Kleywegt, T. A. Jones, Halloween ... masks and bones in *From First Map to Final model. Proceedings of the CCP4 Study Weekend,* S. Bailey, R. Hubbard, D. Waller, Eds. (SERC Daresbury Laboratory, Daresbury, 1994), pp. 59–66.

44. R. J. Read, Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. A* **42**, 140–149 (1986).

45. G. J. Kleywegt, R. J. Read, Not your average density. *Structure* **5**, 1557–1569 (1997).

46. P. Emsley, K. Cowtan, *Coot*: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).

47. D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L. W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams, P. D. Adams, Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in *Phenix*. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877 (2019).

48. V. B. Chen, W. B. Arendall III, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *MolProbity*: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).

49. E. Krissinel, K. Henrick, Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2256–2268 (2004).

50. L. Holm, Benchmarking fold detection by DaliLite v.5. *Bioinformatics* **35**, 5326–5327 (2019).

51. E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).