

Parameter or statistic	Full CsfrRNAV	Empty CsfrRNAV
Data collection and processing		
Magnification	79,000	79,000
Voltage (kV)	300	300
Electron exposure (e ⁻ /Å ²)	30	30
Defocus range (μm)	-0.10 to -4.00	-0.10 to -4.00
Pixel size (Å)	1.1041	1.1041
Symmetry imposed	1	1
Initial micrographs (No.)	10,310	10,310
Final micrographs (No.)	9,947	9,772
Initial particle images (No.)	174,942	152,127
Final particle images (No.)	3,984	1,302
Map resolution (Å)	3.0	3.1
FSC threshold	0.143	0.143
Map resolution range (Å)	2.696 to 3.406	2.4 to 3.588
Refinement		
Map sharpening <i>B</i> factor (Å ²)	91.6	82.9
Model composition		
Non-hydrogen atoms	6171	5689
Protein residues	785	718
<i>B</i> factors (Å ²)		
Protein (min./max./mean)	17.34/122.72/55.06	20.66/131.81/64.13
RMSD*		
Bond lengths (Å)	0.003	0.005
Bond angles (°)	0.680	0.627
Validation		
MolProbity score	2.27	1.88
Clashscore	5.86	7.52
Poor rotamers (%)	3.72	0
Ramachandran plot		
Favored (%)	90.99	92.39
Allowed (%)	9.01	7.61
Disallowed (%)	0	0
CC (mask)	0.88	0.88
EMRinger score	3.21	3.25
Data deposition		
EMDB	EMD-15823	EMD-15830
PDB	8B38	8B3J

*RMSD, root mean square deviation

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

Genus	Species	Virus name	Accession number	Abbreviation	Average pLDDT (%)		VP1 sequence used for AlphaFold
Bacillarnavirus	<i>Chaetoceros socialis formidans RNA virus 1</i>	Chaetoceros socialis f. radians RNA virus 01	A8469874	CsfrRNAV	N/A	N/A	
Bacillarnavirus	<i>Chaetoceros tenuissimus RNA virus 01</i>	Chaetoceros tenuissimus RNA virus 01	A8375474	CtenRNAV01	82		GAQKLTEFVSDVSEVETPSWSDITRADALNPEADLGSFLSRPVQARINWNTSSYLQTSLSNVWSD FLTHPTIVKKLQNSYWSLGSLSHVKLQINGPPYFGKALASYTPYGFSAPTNINSNVYRNNSLARMSQR DHVILDPTCEKGGTIVCPFYPKPLSLDTADLASDGVGRVDIMSFALQNTSADTSTPINISVYAW MEDVRLSGPTTLRSEKIM
Bacillarnavirus	<i>Rhizosolenia setigera RNA virus 01</i>	Rhizosolenia setigera RNA virus 01	A8243297	RsRNAV01	80		DVSVNVGFNSEVPTHTTEIKSAFDRPLSDGWSKLELTSFLRRPVRIEYKTWTVGSFLRDDFFPWDAYL NTASIRKKLENYLRLGSLKLQIFNGTQHHYQKGIYATPFGAAQHPFPFGSAPVLDNVQVYSQKPHV LLDPTDNVGGEMLPVFPFHNVYRHHTQSDLLQQLGQ/DLSSFNKLGHSLGDVTPSVRIAFATYMDDD VALSGSTQFS
Kusarnavirus	<i>Astarnavirus</i>	Asterionellopsis glacialis RNA virus	A8973945	AglaRNAV	90		NSEPLTMSSEAPVEQQITAFADQDAGWKEIHGNYDATMDKVEYTSNDSLGSFLQRPGRSQVT VLINQRLVYKPNPWEFMEFPFIRNKIANEYLVRCKMHCIMVSGTFFHYGRALTSYNPLSGFDQVT VERNYLDQDLYQASQKPHFLNPTKNTGGEICMPFFWEKNYLSLSDKDYENMGELTIKSGNLLHA NNGNDPVTVTYTLWAEDVVLTMPTSLNP
Labyrnavirus	<i>Aurantiochytrium single-stranded RNA virus 01</i>	Aurantiochytrium single-stranded RNA virus 01	A8193726	ASSRNAV01	68		ESKEVGEMSMGGDIHTEHQNVQDHTGEASDRVTGSSIQETGDSGSGTQYLSNFFERPAIYDAT WDGSTEHVNVFEPWDLWSKDTSRKALANYSYKGTMTMHIKATITGPHHTGTIMASYQPYGVYNG NLKYEDEMNMATTGPGPTLPAYKCYLSQAPGAYIDYKKNQCPVLEIPFISHKPKRFLRNNSPDVTN ATSPVDFKEAGELRLVTNLKLAJAVANDFSDVSLNIYAWYTDVLEGNITATDMO
Locarnavirus	<i>Jericarnavirus B</i>	Marine RNA virus JP-B	EF198242	JP-B	84		QSSDIEMTPESTGNTMQEENVGFGQQDPLTIPNDLSRVQDQSQNVTLGNFLKRPVQIATQTW TVGSTITDGLVDNFDPWHLVFNHAPIRKLDNYLVRCLHLKFPVNVASPPYGTCLASYQLTNFAP GFTPTGTAGCERTATTQRPHWVYQDSGGGEMVLPFLYKXWLDATSAVDLTNMGTLFSQSGFSL ANANGATGDIIEIYYAWAEDIEVAGPTVALV
Locarnavirus	<i>Sanfarnavirus 1</i>	Marine RNA virus SF-1	JN661160	SF-1	90		VQADVIPSVEVTGVTGQQRQTVTFLDTPNDGHAGIERSDISMADQVTMAADLREFFSRPVRIASFTW LESDAGTSHTNPNWLVYDQRYKLNINFAQLAELEIKILYNASPPYQCLAGPLQDLPTSTI TNDGTGTYFIPQSQRPWHLEPTGKSGSLTFFYHKNWLNACQAGMDTMDGQLTFLNHTTLAS ANGVAGSGEVTVSFVASKNVRLSGPSVLAV
Locarnavirus	<i>Sanfarnavirus 2</i>	Marine RNA virus SF-2	KF412901	SF-2	84		VATNLPTTFDNDQVAGVTVSPTDNWAEADVGDANTLADLSGFLQRPVRIHTLWLESDPVGVLG NTIDPWSLFFSNAIAIKKIDINYAFRLATLKKIINSTFPNYGAARVYTPRLWTHGRIPTSNALNLVP QSQLPGAWIPQNSQSGCELSPPYKHENRIALTSDMADMGRLSFLYSALKSANGVTATGVSQIV
Locarnavirus	<i>Sanfarnavirus 3</i>	Marine RNA virus SF-3	KF478836	SF-3	86		EKQSDSTPMETSSATNETTQVTSFMDGTTATGTAAAPPSSAPADALTGAELGDLGRPAQIAJF TWSQSDVVGTTTSSYVWQNFNNAIRKLENFAWLRCDLKVMVNASPPYGGAMLATYQPLPN FSPSTIVNDGTGTYFIPLTQRPHAWIPQNGEAMTLPFFWPKNWLSLTSNADFADMTLSLSVFT DLSANGATGSGATFTYAWAENVTLNGPTCGLIM
Marnavirus	<i>Heterosigma akashiwo virus</i>	Heterosigma akashiwo RNA virus	AY337486	HaRNAV	76		TCTPHDEKLLTETLCNTLKRYDTSGLRNESHOTMSNGTERRRPGEVDGLASLRPRTVFTDTPKPG TNFYQASNIWKSFLNDOTVDKISHFARLRGKMVRLINLNGNSMYGKLVHMHSYFQADVDVTVSS VPDPAEIQWIMQKPHSFDAITTGATMELPMLPNWDVLDLTDIEDMIARGLTYIHDLNLEHANA GTDELTVTLVAVWMEDEVLYLPTSTSEI
Salisharnaviruses	<i>Britannavirus 4</i>	Marine RNA virus BC-4	MH171300	BC-4	89		ADGTTNDSNVKISGAGKYENVCFSDQVDPYAYNVESIMDPTKRLQDGTGATLQNFSPRIKIEQVE WATTSQCLJFDVNPWEAFNGTRVANRLSNFLRLAKLHVKIVINGNGFQYGRALASYLPSIYDTLST NAGLJLRQDLVQASQKPHFLDPTTSQSGEMTLPPYNYWNYSSIPDTQWNELGILQFRSINDLKH GASDVTVTSVFAWAEDEVMVSLVSQDQ
Salisharnaviruses	<i>Palmarnavirus 473</i>	Marine RNA virus PAL473	K7727026	PAL473	87		FNNVINTDXTMMSSGNVTFKQDTPSYSVAAANDATFEVADRSDSLGEFFSRPVRIREFQWATT TLFESPNWHTDVTNPNWLNRTNHLRLKCLKPVINGNGFHYGRICISYPLRDDOFTVDRAPIE DIVGASQRPYHLDPPTSSGGTMCLPFWNYNCLVNEEWSKMGTMIMHTMQSLQHANGANDS ATISVFAWAEVDVLAVPTSSERG
Sogarnavirus	<i>Britannavirus 2</i>	Marine RNA virus BC-2	MG584188	BC-2	86		GTIQEQGVANFSEITDFNEQDAGWTTKCGSGMDATMNLGSSGDSLTGSLGRPTRIADFSWVVDQ PFKEIDPWSLFLNDRVAEKIANEYLRKLNHVMVISGTFHYGRALVSNPYSGYDQITVERNFLQ QDLVAASQKPHFLNPTNTNGQLDLPFFWPNYLSLSSSNRNLGEMFKSFLQHANEENDPV NITYVWASDVELTMPTSLTL
Sogarnavirus	<i>Britannavirus 3</i>	Marine RNA virus BC-3	MG584189	BC-3	88		EGSAAATEQITAFSDQAGWTTCKGGSDATMNLSSNSDGLGNFLERPIRIHEAPWIMGQPLTTF NPWVDFMSNPVKEIAHYELLRMNLHVKFVISGTFHYGRALVSNPYLFDITQRYLVLDVLQ ASQKPHFLNPTNAGSQKPHFLMDPTTSQGGEMLLPFWPENYLSLSSSNRNLGEMFKSFLQHANEENDPV WASDVLTMPTSLT
Sogarnavirus	<i>Chaetarnavirus 2</i>	Chaetoceros species RNA virus 02	A8639040	CspRNAV2	70		DSNSPSGGAYSVSKAAEQSTQNVHVDGDTPHWSYDISSPDVTTQLAGFSDAELGSFLGRPIKIEKQ WTPESTRLFEVNPWTEFNSADLEKINRYRNLRLRMKMLINGNSFYGRALVSNPYLNDNSV TLNRAFFEQDLVGASQKPHFLMDPTTSQGGEMLLPFWPENYLSLSSSNRNLGEMFKSFLQHANEENDPV HANGGTDPTVTYVFWAEDVLSVPTTVQAQ
Sogarnavirus	<i>Chaetenuissarnavirus II</i>	Chaetoceros tenuissimus RNA virus type II	A8971661	CtenRNAVII	N/A	N/A	
Sogarnavirus	<i>Jericarnavirus A</i>	Marine RNA virus JP-A	EF198241	JP-A	92		NSEGDTGFSVSKVDRESAAQNVHVDGDTPWWSYDIKANADETTKLAFTDAELGDLSPRIKIEKQ WTPGVALSAPRPNPWRREFNNSDVLQINRYRNLRLRMKMLVINGNSFYGRALVSNPYLNDNSV VTYVNTFREEDLQASQKPHFLDPTTSQGGEMLLPFWPENYLSLSSSNRNLGEMFKSFLQHANEENDPV QHANGGTDPSITFCWAENLTLAIPTAQVS
Sogarnavirus	<i>Palmarnavirus 156</i>	Marine RNA virus PAL156	K7727024	PAL156	89		GTSQKPTTNLTQETAFSDQAGWTTTGGSDPDTMDLAATVDSQDLGNFLGRPVRIAYNVWQVS GFERFNPWREFINDEVERKEATAHFLYRSLKHVMNVSGTFHYGRALVSNPYLGRDVOYTGQSF LQSDLVAAASQKPHFIMPTTNAGGQLDLPFFYDNYLSLRTGDAQEMGEITKSFDSLHANGGTGN VTYVYVWASDVLVAMPSTSVNLL

Table S2. Related *Marnaviridae* virus species used for AlphaFold2 and the average pLDDT scores.

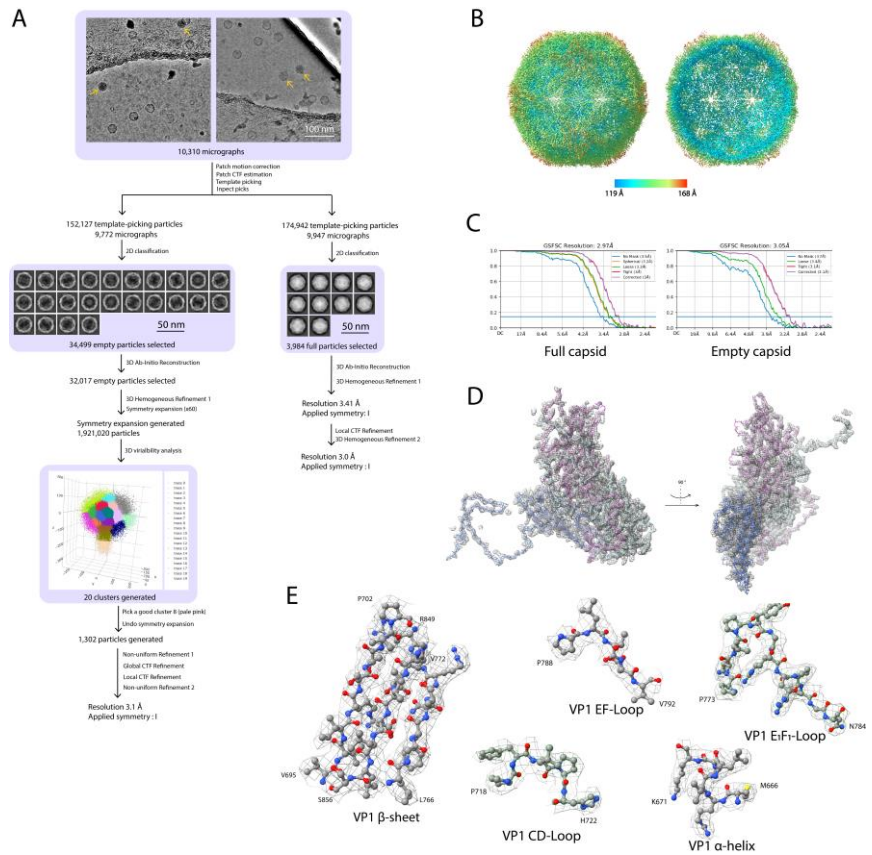


Figure S1. Summary of the cryo-EM data analysis. A) Flowchart of the CsfrRNAV capsid data analyses of the cryo-EM micrographs using cryoSPARC. Yellow arrows point out full CsfrRNAV particles in the cryo-EM micrographs. B) Overall 3D reconstruction of the full CsfrRNAV capsid. Left: The overall capsid map. Right: The back half of the capsid shows the interior features. Capsid is colored from blue to red according to its calculated radius. C) The gold standard FSC resolution curves of both full and empty particles in icosahedral (I) symmetry. The resolutions of these cryo-EM models are estimated as 3.0 (full) and 3.1 (empty) Å (FSC cutoff = 0.143). D) The cryo-EM map and the backbone fitting with the atomic models of the capsid VP1 (pink), VP2 (blue), and VP3 (light green). E) Representative β -sheet, loops and α -helix show the refined side chains in full capsid of CsfrRNAV.

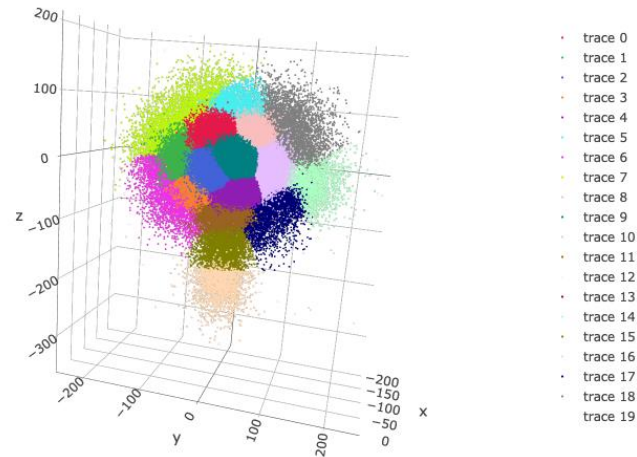


Figure S2. 3D variability display. The 3D variability display performed in cryoSPARC 3.3.2. The symmetry expanded empty particles are sorted into 20 clusters in total. Only the map in cluster 8 (pale pink) shows clear secondary structure and therefore is used for further 3D refinement.

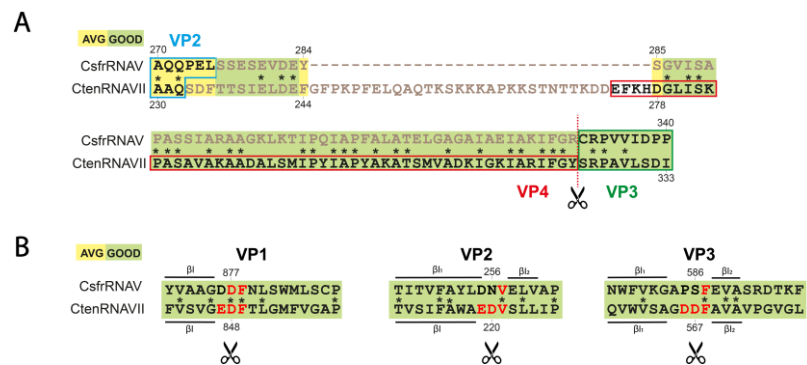


Figure S3. Self-cleavage sites and partial sequence alignments of CsfrrNAV and CtenRNAVII. AVG: average sequence identities. GOOD: good sequence identities. * Identical residues. A) Partial sequence alignment result and the putative VP4 VP3 proteolytic site in CsfrrNAV and CtenRNAVII. Blue, green, and red boxes represent parts of VP2, VP3, and VP4, respectively. Red dashed line represents the proteolytic site between VP4 and VP3. Modeled residues are in black. Unmodeled residues are in gray. B) Putative self-cleavage sites in the VP1, VP2, and VP3 of CsfrrNAV and CtenRNAVII. Residues that are potentially involved in self-cleavage in each VP are colored red. Sequences are aligned using the T-COFFEE (version 11.00) server [59].

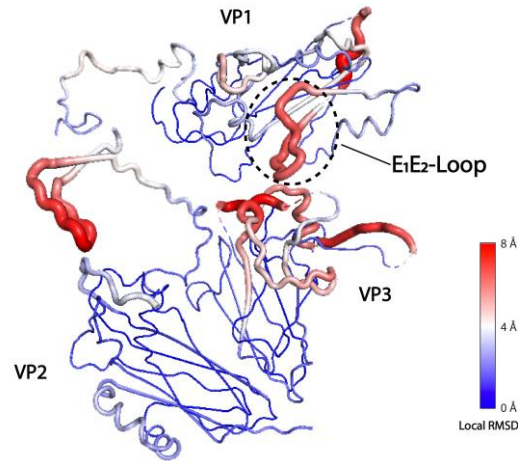


Figure S4. RMSD per residue for VP1–3 (CsfrRNAV vs CtenRNAVII). Residues are highlighted from blue to red, showing low to high RMSD values. VP2 does not show apparent structural differences apart from the N-terminal arm. VP3 shows some structural differences in the surface loops adjacent to the VP1 E₁E₂-Loop.

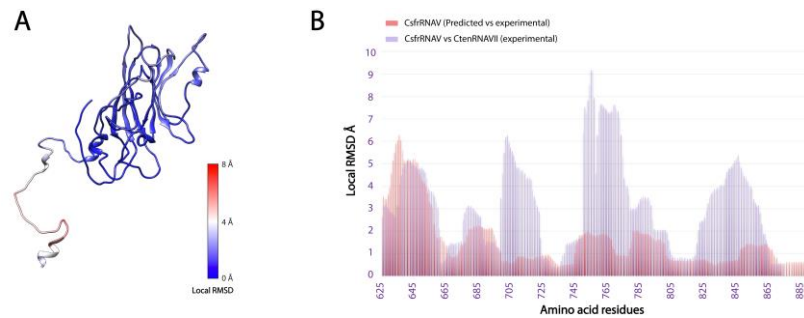


Figure S5. AlphaFold2 structural prediction error (predicted vs. experimental) and VP1 structural differences between two representative *Marnaviridae* viruses (CsfrRNAV vs. CtenRNAVII). A) The structural prediction error per residue between the experimental and predicted CsfrRNAV VP1s. The residues are colored from low to high deviations (blue to red) by the attribute of their local RMSD. B) The overall comparison between experimental CsfrRNAV VP1 and CtenRNAVII VP1 structural differences (purple) and the AlphaFold2 prediction errors (red), which were calculated using CsfrRNAV VP1 atomic model as a reference.

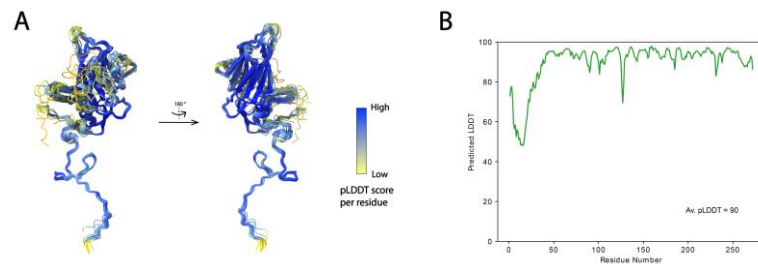


Figure S6. Superimposition of the predicted *Marnaviridae* viruses VP1s and the representative residue pLDDT score plot. A) Superimposition of VP1s from all 16 predicted *Marnaviridae* viruses. All models are colored according to their pLDDT scores by residue from yellow (backbone modeled well) to blue (high accuracy).

B) Predicted VP1 model of a locarnavirus (JN661160) is shown as an example plot of pLDDT score per residue.