Structure, Volume *32*

Supplemental Information

Structural characterization of Thogoto Virus nucleoprotein provides insights into viral RNA encapsidation and RNP assembly

Alexej Dick, Vasilii Mikirtumov, Jonas Fuchs, Ferdinand Krupp, Daniel Olal, Elias Bendl, Thiemo Sprink, Christoph Diebolder, Mikhail Kudryashev, Georg Kochs, Yvette Roske, and Oliver Daumke

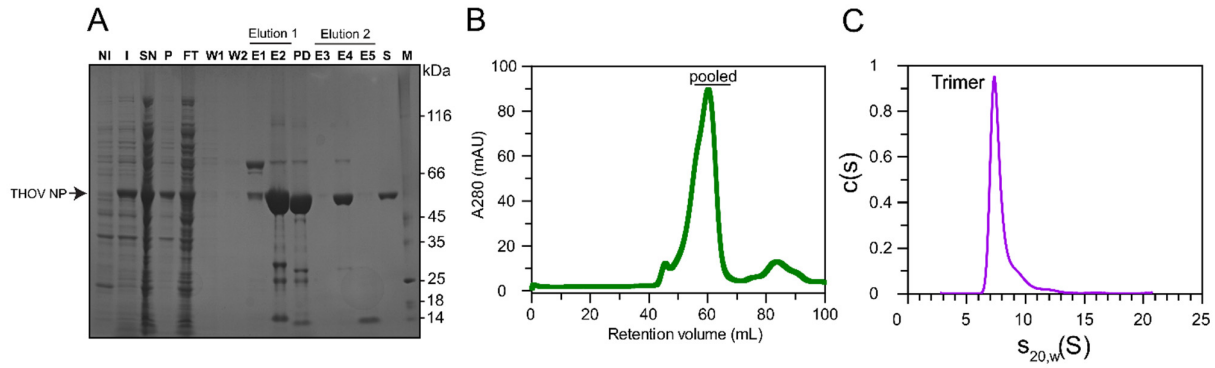


Figure S1: Purification and initial characterization of THOV NP (related to Fig. 1). (**A**) Purification of THOV NP, representative SDS-PAGE. NI, non-induced *E. coli* lysate; I, induced whole-cell *E. coli* lysate; SN, supernatant fraction of the cell lysate; P, pellet fraction; FT, flow-through of the soluble extract on Ni^{2+}-NTA column; W1, flow-through of a high-salt and ATP wash; W2, flow-through of a high-imidazole wash (wash 2); E1 and E2, elution 1 fractions; PD, post digestion with PreScission protease (visible at approx. 42 kDa); E3, E4 and E5, elution 2 fractions (removal of the GST-tagged protease via GSH-column); S, size exclusion chromatography fractions; M, protein molecular weight marker. (**B**) Preparative gel filtration run of THOV NP. Pooled fractions are indicated. (**C**) Analytical ultracentrifugation sedimentation velocity experiments of THOV NP at 0.5 mg/mL in gel filtration buffer containing 150 mM NaCl. The relative protein concentration c(s) is plotted as a function of the normalized sedimentation coefficient $s_{20,w}$.

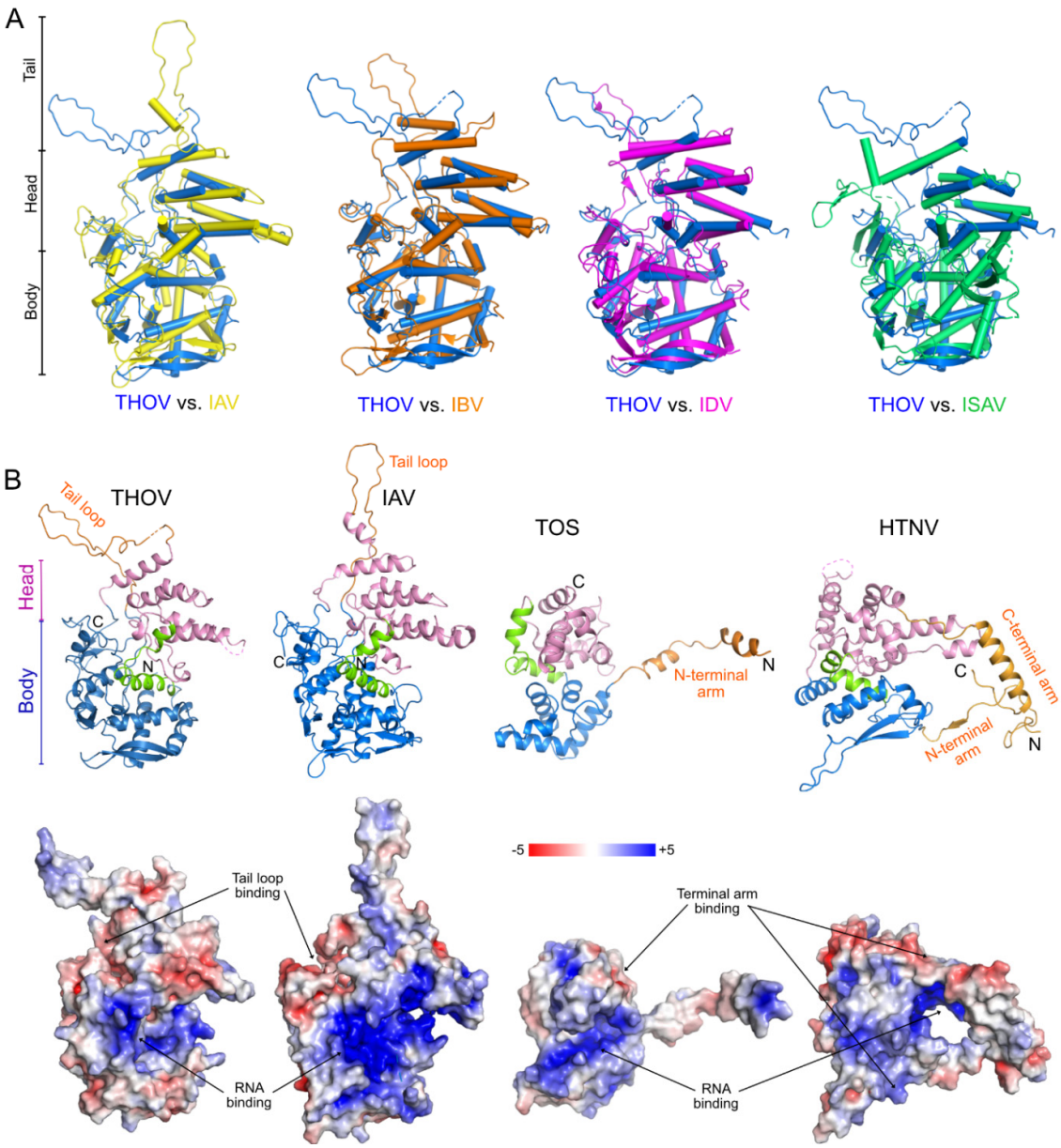
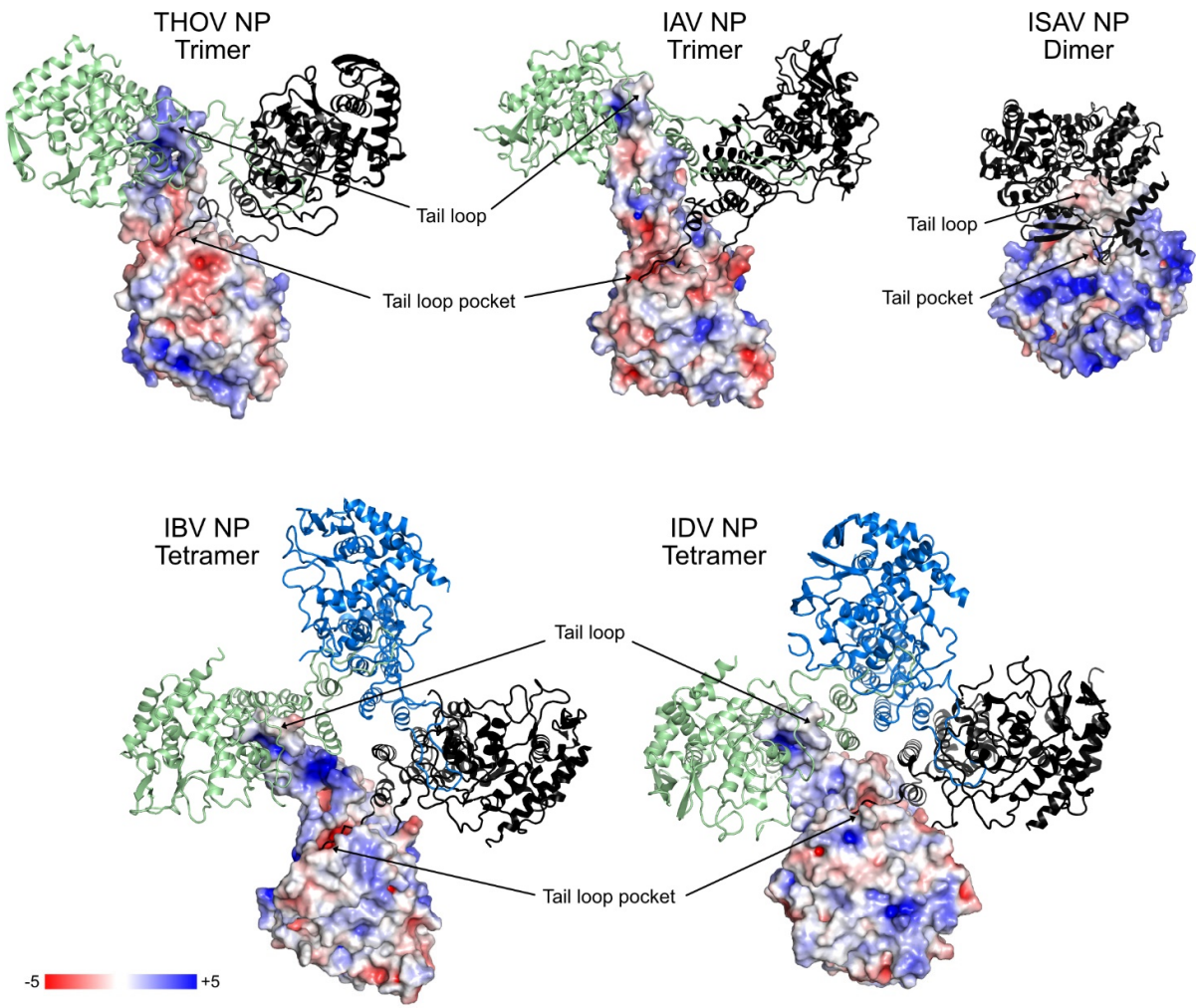


Figure S2: Structural comparisons of THOV NP (related to Fig. 1). (**A**) Topology-based structural Cα atom alignment (side view) of the THOV NP structure with structures of the orthomyxovirus family members, together with RMSD values. The following RMSD values were obtained: Influenza A virus (IAV) (PDB code: 2Q06, sequence identity of 14.7% compared to THOV-NP): RMSD of 2.4 Å over 333 Cα-atoms. Influenza B virus (IBV) (PDB code: 3TJ0, sequence identity of 15.2%): RMSD of 2.6 Å over 348 Cα-atoms. Influenza D virus (IDV) (PDB code: 5N2U, sequence identity of 12.3%): RMSD of 3.0 Å over 308 Cα-atoms. Infectious Salmon Anemia virus (ISAV) (PDB code: 4EWC, sequence identity of 14.9%): RMSD of 2.7 Å over 315 Cα-atoms. (**B**) Comparison of the THOV NP structure to NP structures of other negative-strand RNA viruses. Shown from left to right are the NP structures of the Thogoto virus $NP^{\Delta 188\text{-}196}$, the Influenza A virus (IAV) NP (PDB code: 2Q06), the Toscana (TOS) virus NP (PDB code: 4CSG) and the Hantaan virus (HTNV) NP (PDB code: 6I2N). All NPs are composed of a head (pink) and a body domain (blue) connected via a helix-loop-helix motif (green). The tail regions are indicated in orange. Hantaan virus NP uses both termini for oligomerization. The electrostatic surface presentation for each NP structure is shown below the corresponding NP structure with identical orientation, highlighting the positively charged groove for RNA accommodation. The color range, calculated at 310 K, ranges from -5 kcal/mole*e (red) to +5 kcal/mole*e (blue).



Virus	Oligomer in solution/crystal	Interface area [Å^2]	Δ^iG [kcal/mol]	# HB	#SB	PDB ID
ISAV	Dimeric	3509	-47.7	36	2	4EWC
THOV	Trimeric	2160	-12.0	45	6	8CJW
IAV	Trimeric	2380	-18.1	33	10	2Q06
IBV	Tetrameric	2130	-28.7	24	5	3TJ0
IDV	Tetrameric	2260	-22.8	31	8	5N2U

Figure S3: Comparison of orthomyxovirus NP oligomer assembly (related to Fig. 3). Assembly of NPs from the orthomyxovirus family. One protomer is shown with an electrostatic surface, while the others are colored green, blue, and black in cartoon representation. IAV NP (PDB code: 2Q06, sequence identity of 14.7%), IBV NP (PDB code: 3TJ0, sequence identity of 15.2%), and IDV NP (PDB code: 5N2U, sequence identity of 12.3%), ISAV NP (PDB code: 4EWC, sequence identity of 14.9%) The electrostatic surface color range, calculated at 310 K and with an ionic strength of 150 mM, is from -5 kcal/mole*e (red) to +5 kcal/mole*e (blue). An interprotomer interface analysis table of represented orthomyxovirus NPs is shown below, with HB for hydrogen bonds and SB for salt bridges. Interface analysis was performed using the PDBe PISA server (https://www.ebi.ac.uk/pdbe/pisa/).[S1]

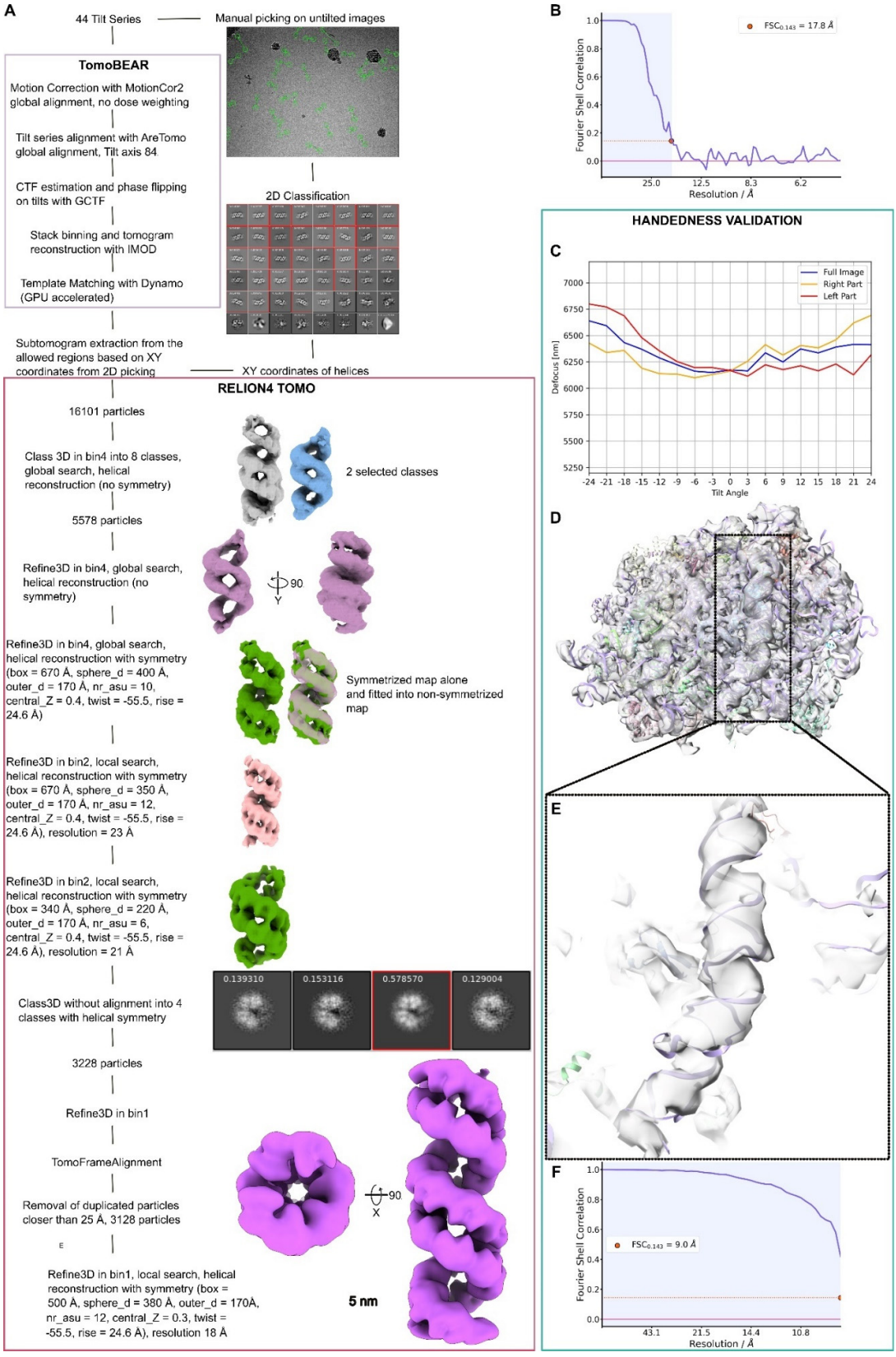


Figure S4: Structure and helical handedness determination of THOV RNP (related to Fig. 4). (**A**) Processing diagram of the RNP data set. (**B**) FSC curve for the RNP structure. (**C**) Plot of defocus values versus tilt angles for left and right sides of the aligned stack for the RNP data set. (**D**) Ribosome LSU map with fitted atomic model (PDB: 6LSS). (**E**) Close-up view on the RNA helix in the map clearly showing the correct handedness. (**F**) FSC curve for the ribosome LSU data set.

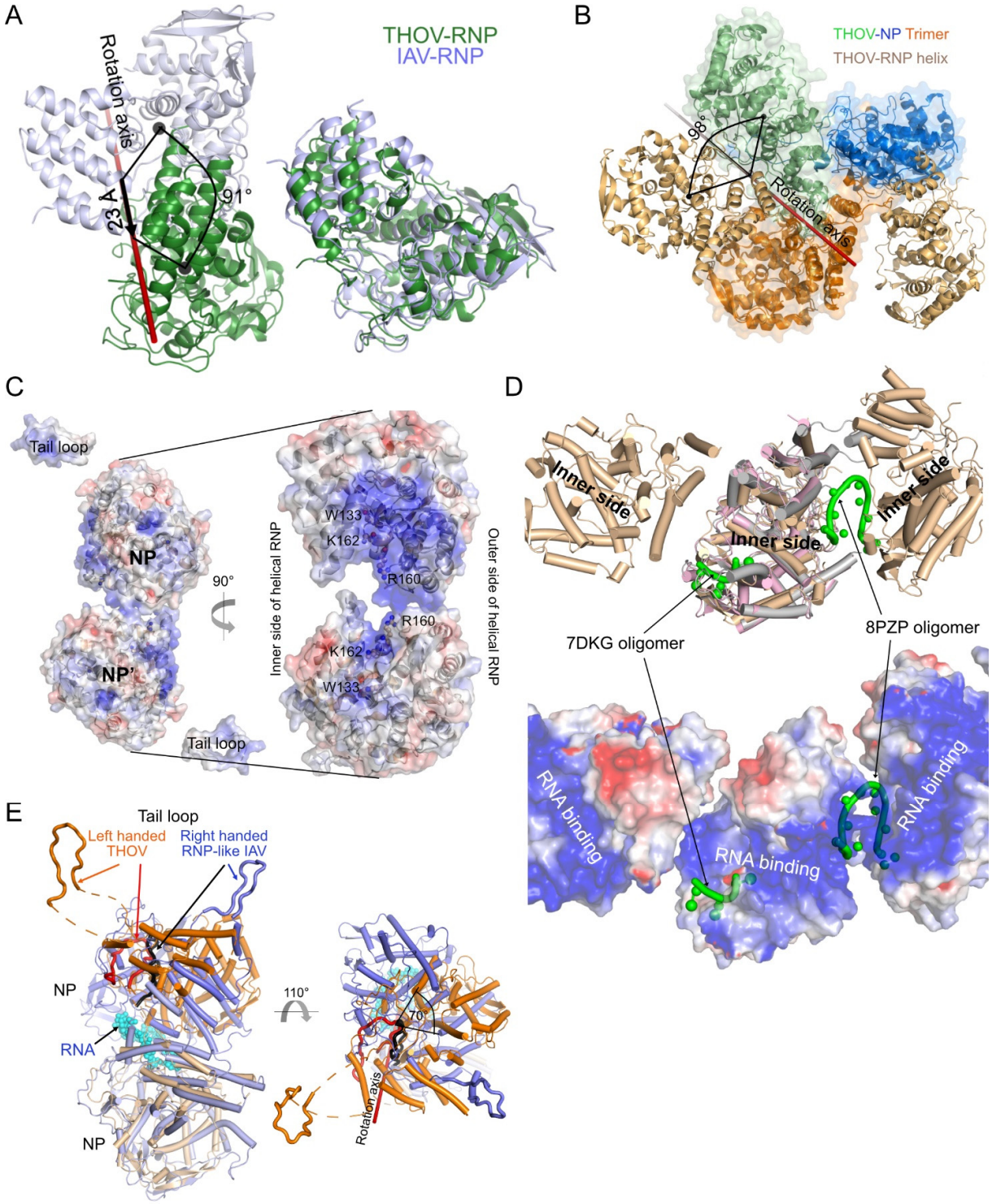
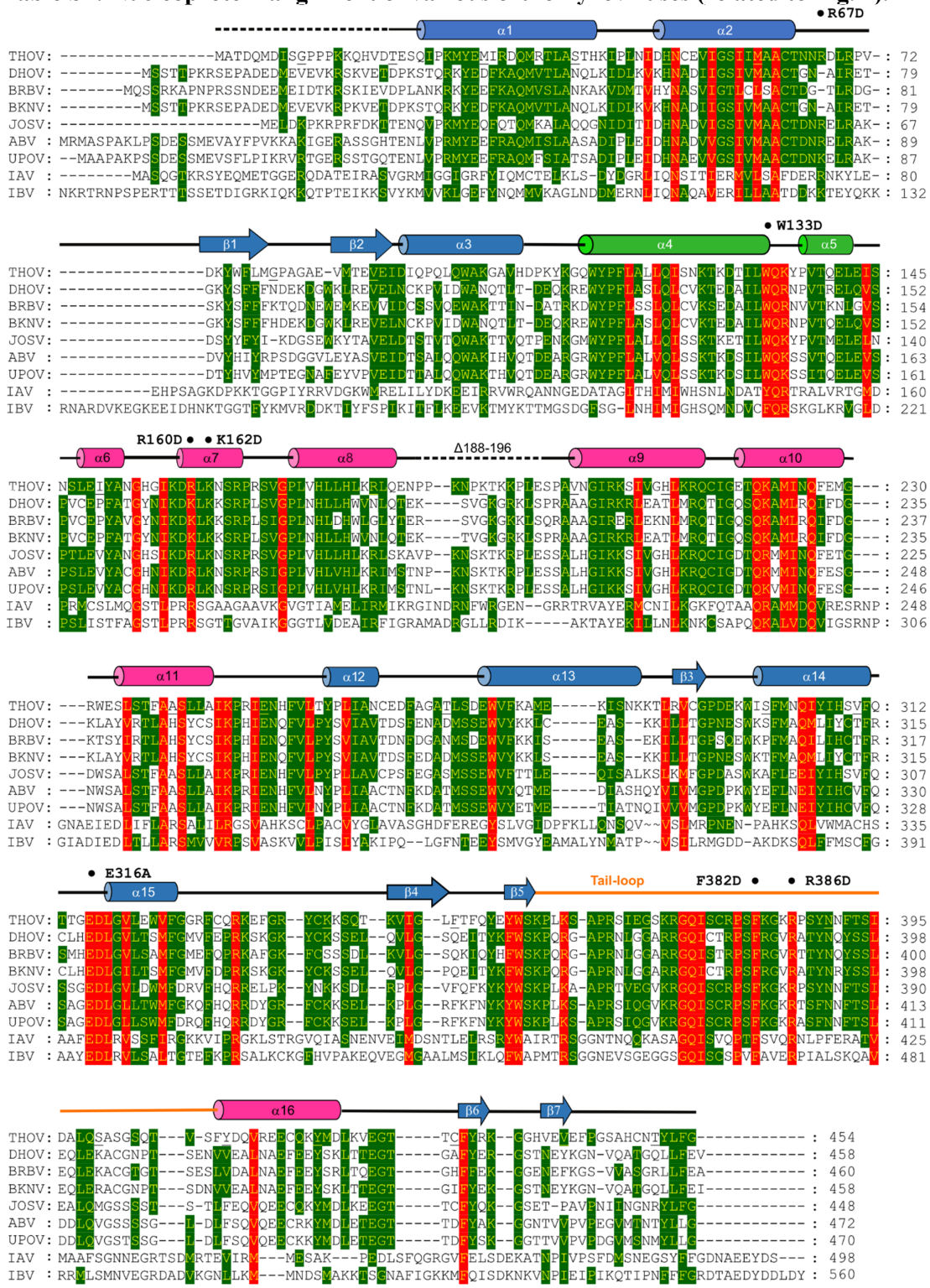


Figure S5: Structural comparison of NP molecule arrangement in THOV and IAV RNP (related to Fig. 4). (**A**) Different angular arrangement of THOV NPs and IAV NPs in left-handed RNP structures. Shown are two adjacent intra-strand NP protomers along one RNP filament in the THOV-RNP and IAV-RNP (PDB: 4BBL) [S2] shown as green and blue cartoon, respectively. To demonstrate the different arrangement of the protomers to each other, one NP protomer of THOV and IAV is superimposed. Based on their centres of mass (black spheres), a displacement angle of 91° and a displacement of 23 Å along the rotation axis was determined for the adjacent IAV NP with respect to the THOV NP protomer. (**B**) One THOV NP protomer (orange) of the trimeric crystal structure arrangement shown as cartoon with transparent surface representation in orange, blue and green is superimposed with one NP protomer of the helical arrangement in THOV-RNP shown as wheat

colored cartoon (only the direct adjacent protomers are shown). From the trimeric to the helical arrangement a THOV-NP protomer has to rotate 98°. (**C**) View onto the electrostatic surface of two inter-strand connected NP molecules of THOV RNP to demonstrate each NP's highly positively charged inner side cavity forming a tunnel at each NP strand for possible RNA binding. The color ranges from -5 kcal/mol*e in red to +5 kcal/mol*e in blue. Depicted are the residues that were mutated and showed decreased RNA binding as analyzed in **Fig. 2E.** (**D, upper panel**) Section of a THOV RNP filament strand showing three intra-strand connected NP molecules (wheat color) with a view to the inner side of the RNP helix. Two influenza NP molecules (light pink: Influenza H5N1 (PDB: 7DKG), gray: Influenza A (PDB:8PZP)) were superimposed on the central THOV NP, with bound RNA oligomers (green color). (**D, lower panel**) Same THOV NP filament section but as electrostatic surface presentation with the RNA oligomers (green) from the influenza NP structures, demonstrating a possible RNA binding path. The color ranges from -5 kcal/mol*e in red to +5 kcal/mol*e in blue. (**E**) Superimposition of a THOV NP dimer from the left-handed THOV RNP structure in wheat-orange color onto the right-handed RNP-like IAV helical structure (blue) bound to an RNA oligo (cyan spheres). The corresponding tail loops are highlighted as thick lines in wheat, red in THOV-RNP, and blue and black for the IAV RNP-like structure. The neighbored NP molecule of the two superimposed molecules of the right-handed RNP-like structure is rotated 70° compared to the position in the left-handed THOV-RNP structure.

Table S1: Nucleoprotein alignment of various orthomyxoviruses (related to Fig. 1).



NP amino acid sequences of THOV (Swiss-Prot accession A0A0B6VKB5), DHOV (D6PT87), Bourbon virus (BRBV) (A0A0C5I425), Batken virus (BKNV)[3], Jos virus (JOSV) (H6SW55), Aransas Bay virus (ABV) (X2CUE7), Upolu virus (UPOV) (X2CWU1), IAV H1N1 (2009) (C4RU37) and IBV (P04665) were aligned using CLUSTAL Omega [S4] and manually adjusted. 100% type-conserved residues are boxed in red, and residues with >=50% conservation are boxed in green. α-Helices are shown as cylinders and β-strands as arrows in colors according to their domain contribution in blue, pink, and green for the body and head domain and helix-loop-helix motif.

Table S2: Data collection and refinement statistics (related to Fig. 1)

	THOV (SeMet)	THOV (pdb 8CJW) (native)
Data collection		
Beamline	PETRA III	BESSY 14.1
Wavelength (Å)	0.9800	0.91841
Space group	R3	R3
Cell dimensions:		
a, b, c (Å)	119.4, 119.4, 85.3	119.1, 119.1, 91.6
α, β, γ (°)	90.0, 90.0, 120.0	90.0, 90.0, 120.0
Resolution (Å)[a]	59.7 – 3.5 (3.59 – 3.5)	44.9 – 2.01 (2.70 – 2.54)
R_{meas}	8.6 (81.9)	5.6 (197.5)
I / σ(I)	7.55 (1.44)	15.30 (0.74)
CC1/2	0.998 (0.734)	0.999 (0.329)
Completeness (%)	97.7 (92.6)	99.5 (99.1)
Redundancy	2.4	4.3
Refinement		
Resolution (Å)		44.9-2.01
No. unique reflections		32,179 (5,199)
R_{work} / R_{free} (%)		18.5 / 23.4
No. atoms:		
Protein		3,420
Water		209
Mean B factor (Å^2):		78.71
Protein		78.74
Water		75.40
Molecules /asymmetric unit		1
Model geometry		
R.m.s deviations:		
Bond lengths (Å)		0.008
Bond angles (°)		0.896
Ramachandran analysis:		
Ramachandran:		
Favored (%)		97.11
Allowed (%)		2.65
Outliers (%)		0.24

[a] Values in highest resolution shell are indicated in parenthesis.

Table S3: Thermodynamic parameters of THOV NP mutants to a 24mer polyU RNA derived from ITC analysis (related to Fig. 2)

Protein	Binding affinity - K_D [µM]	Binding number n	Fold decrease
Wild type	0.34 ± 0.05*	0.640 ± 0.001	
R67D	3.4 ± 0.6	1.22 ± 0.08	10
W133D	1.0 ± 0.5	0.72 ± 0.04	3
R160D	20 ± 20 (not saturated)	0.6 ± 0.5	47
K162D	5.2 ± 1.2	0.59 ± 0.02	15

*Note that these experiments, due to sample restraints, were performed with an iTC200 device, in contrast to all other measurements reported in this manuscript performed on a VP-ITC. Slight deviations of affinities may result from using different devices.

Table S4: Cryo-EM data collection, refinement, and validation statistics (related to Fig. 4)

	THOV RNP **EMDB:** EMD-19599 **PDB:** 8RYT	**Ribosome LSU**
Data Collection		
Magnification	105,000	19,500
Defocus range (μm)	-3 to -6	-3 to -7
Voltage (kV)	300	300
Microscope	G3i Titan Krios	G3i Titan Krios
Camera	Gatan K3 BioQUantum GIF	Gatan K3 BioQUantum GIF
Fraction exposure time (s)	0.12	0.2
No. of movie fractions	10	10
Total electron dose ($e^-/Å^{-2}$)	115.5	102
Pixel size (Å)	2.6	4.486
angle range (°)	-48° to +48°	-48° to +48°
increment (°)	3	3
No. of tilt series collected	44	9
No. of tilt series used	33	9
Refinement		
Box size (pixel)	192	96
Inter-box distance (Å)	50	-
No. of segments extracted	16,101	30,928
No. of segments for final map	3,128	5,451
Resolution	17.8	9
Map sharpening B-factor ($Å^2$)	0	-200
binned pixel size (Å)	2.6	4.486
Imposed symmetry	Helical (twist = -55.5°, rise = 24.6 Å)	C1
Model composition		
No. atoms Protein	52,473	
R.m.s. deviation		
Bond length (Å)	0.003	
Bond angles (°)	0.538	
Ramachandran plot		
Favored (%)	96.6	
Allowed (%)	3.4	
Disallowed (%)	0	

Table S5: Primer used in this study (related to STAR Key resource Table)

OLIGONUCLEOTIDES	SOURCE	IDENTIFIER
pSKB-LNB THOV NP R386A fw: 5'CAGACCTTCTTTCAAAGGAAAGGCTCCCTCTTACAACAATTTCACC 3'	This Paper	This Paper
pSKB-LNB THOV NP R386A rev: 5' GGTGAAATTGTTGTAAGAGGGAGCCTTTCCTTTGAAAGAAGGTCTG 3'	This Paper	This Paper
pSKB-LNB THOV NP E316A fw: 5' CAGTGTATTCCAGACCACAGGCGCTGATCTGGGAGTTTTAGAGT 3'	This Paper	This Paper
pSKB-LNB THOV NP E316A rev: 5' ACTCTAAAACTCCCAGATCAGCGCCTGTGGTCTGGAATACACTG 3'	This Paper	This Paper
pSKB-LNB THOV NP F382D fw: 5' CAGATCTCGTGCAGACCTTCTGATAAAGGAAAGCGCCCCTCTTAC 3'	This Paper	This Paper
pSKB-LNB THOV NP F382D rev: 5' GTAAGAGGGGCGCTTTCCTTTATCAGAAGGTCTGCACGAGATCTG 3'	This Paper	This Paper
pSKB-LNB THOV NP R67D fw: 5' TGGCTGCCTGTACAAACAACGATGACCTGAGACCAGTGGAT 3'	This Paper	This Paper
pSKB-LNB THOV NP R67D rev: 5' ATCCACTGGTCTCAGGTCATCGTTGTTTGTACAGGCAGCCA 3'	This Paper	This Paper
pSKB-LNB THOV NP W133D fw: 5' CAAGACCAAAGATACCATCTTGGATCAGAAATACCCCGTCACCC 3'	This Paper	This Paper
pSKB-LNB THOV NP W133D rev: 5' GGGTGACGGGGTATTTCTGATCCAAGATGGTATCTTTGGTCTTG 3'	This Paper	This Paper
pSKB-LNB THOV NP R160D fw: 5' CGGTCACGGCATCAAAGACGATCTTAAGAACAGCAGGCCCAGGT 3'	This Paper	This Paper
pSKB-LNB THOV NP R160D rev: 5' ACCTGGGCCTGCTGTTCTTAAGATCGTCTTTGATGCCGTGACCG 3'	This Paper	This Paper
pSKB-LNB THOV NP K162D fw: 5' CACGGCATCAAAGACCGGCTTGATAACAGCAGGCCCAGGTCTGTC 3'	This Paper	This Paper
pSKB-LNB THOV NP K162D rev: 5' GACAGACCTGGGCCTGCTGTTATCAAGCCGGTCTTTGATGCCGTG 3'	This Paper	This Paper

Supplemental References

S1. Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J Mol Biol *372*, 774-797. https://doi.org/10.1016/j.jmb.2007.05.022.

S2. Arranz, R., Coloma, R., Chichón, F.J., Conesa, J.J., Carrascosa, J.L., Valpuesta, J.M., Ortín, J., and Martín-Benito, J. (2012). The structure of native influenza virion ribonucleoproteins. Science *338*, 1634-1637. https://doi.org/10.1126/science.1228172.

S3. Fuchs, J., Lamkiewicz, K., Kolesnikova, L., Hölzer, M., Marz, M., and Kochs, G. (2022). Comparative Study of Ten Thogotovirus Isolates and Their Distinct In Vivo Characteristics. J Virol *96*, e0155621. https://doi.org/10.1128/jvi.01556-21.

S4. Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res *22*, 4673-4680. https://doi.org/10.1093/nar/22.22.4673.