Supporting Information

Li et al. 10.1073/pnas.1222552110

SI Materials and Methods

Protein Production. The gene of the full-length Bunyamwera virus (BUNV) nucleocapsid protein (NP) (residues 1–233) was cloned into the pQE-30 expression vector with a 6× His tag fused at the N terminus following a general protocol. The plasmid was transformed into the *Escherichia coli* strain BL21(DE3), and the transformed cells were cultured at 37 °C in LB media containing 100 mg/L ampicillin. After the OD_{600} reached 0.5, the culture was cooled to 289 K, and the expression of the recombinant protein was induced. After overnight induction, the cells were harvested by centrifugation. The cells were then resuspended in lysis buffer containing 20 mM Tris·HCl (pH 7.5) and 150 mM NaCl followed by homogenization using an ultra-high pressure cell disrupter (JNBIO) at 277 K. The insoluble material was removed by centrifugation at 20,000 × *g*. The fusion protein was first purified by Ni–nitrilotriacetic acid affinity chromatography and eluted with wash buffer containing 60 mM imidazole in 20 mM Tris·HCl (pH 7.5) and 150 mM NaCl. The bound target protein was pretreated with RNase at a concentration of 0.2 mg/mL and subsequently eluted with a buffer containing 300 mM imidazole in Tris·HCl (pH 7.5) and 150 mM NaCl after extensive washing with a buffer containing 20 mM Tris·HCl (pH 7.5), 150 mM NaCl, and 60 mM imidazole to remove the untagged RNase. The eluted protein was further purified by a Superdex 200 gel filtration column (GE Healthcare) and dialyzed in a buffer containing 20 mM Tris·HCl (pH 7.5) and 150 mM NaCl. The purified BUNV NP was >95% pure according to SDS/PAGE analysis and had an A_{260}/A_{280} ratio of 1.3. The purified BUNV NP–RNA complex was concentrated to 10 mg/mL and stored at 193 K.

Crystallization. Crystallization of the BUNV NP–RNA complex was performed at 291 K by the hanging drop vapor diffusion technique. Crystals were obtained by mixing 1 μL of protein solution with an equal volume of reservoir solution followed by equilibration against 200 μL of reservoir solution. The crystals for the native BUNV NP were obtained in a reservoir solution containing 0.02 M magnesium chloride hexahydrate, 0.1 M Hepes (pH 7.5), and 22% (wt/vol) polyacrylic acid sodium salt. Crystals with high diffraction quality grew to a final size of 50 × 50 × 100 μm within 7 d in the optimized solution containing 100 mM sodium malonate (pH 7.0) and 10% (wt/vol) PEG 3350. The selenomethionine derivatives of BUNV NP were purified following a general procedure (1) and then crystallized under conditions similar to those used for the native protein.

X-Ray Data Collection, Processing, and Structure Determination. All crystals were gradually transferred into a harvesting solution containing the respective precipitant solutions plus 5% (vol/vol) glycerol before being flash-frozen in liquid nitrogen for storage. Data were collected under cryogenic conditions at 100 K. The selenomethionine single-wavelength anomalous dispersion (SAD) dataset of the BUNV NP–RNA complex was collected at 4.0 Å using a wavelength corresponding to the Se peak at the Photon Factory beamline BL17A, and another native dataset was collected at 3.2 Å at this beamline. All datasets were processed using the HKL2000 package (2). The crystals belonged to the space group *I422* with cell parameters $a = b = 106.2$ Å, $c = 485.4$ Å, and $\alpha = \beta = \gamma = 90°$. Excluding the first selenium of each polypeptide, 24 of 28 selenium atoms in the asymmetric unit were located and refined, and the SAD data phases were calculated and substantially improved by solvent flattening using the PHENIX program (3). A model was manually built into the modified experimental electron density using Coot (4) and further refined in PHENIX. Model geometry was verified using the program MolProbity (5). The final refinement statistics are summarized in Table S1. Structural figures were drawn using the program PyMOL (6). The coordinates and structure factors were deposited in the Collaboratory for Structural Bioinformatics under the Protein Data Bank ID code 4IJS.

EM of the Recombinant BUNV NP–RNA Complex. Samples were prepared by negative staining with uranyl acetate. For specimens prepared by conventional negative staining, images were taken using an FEI Tecnai 20 electron microscope equipped with a tungsten filament and operated at an acceleration voltage of 200 kV. Images were recorded on a 2 k × 2 k Gatan CCD camera at a magnification of 50,000× using a defocus value of –1.5 μm, and the pixel size was 1.6 Å on the specimens. Two-dimensional class averages and 3D reconstructions were calculated using the Electron Micrograph Analysis (EMAN) software package (7).

For 2D analysis, 4,037 particles were interactively selected from 300 CCD images. After linearization, normalization, and centering, all of the particles were subjected to refine2d.py for the reference-free 2D classification. For 3D reconstruction, due to the preferred orientation of the proteins on the carbon films, 600 additional CCD images of 30° tilted samples and 50° tilted samples were collected. The 3D reconstruction with C4 symmetry was generated using 5,801 particles. According to the Fourier shell correlation = 0.5 criterion, the final density map had a resolution of 18.9 Å (Fig. S6).

EM of Natural BUNV Ribonucleoproteins. Bunyamwera virions released by infected BHK-21 cells were produced and purified as previously described (8). The purified virions were adsorbed to EM grids covered by Formvar and carbon and made hydrophilic by glow discharge. Adsorbed virions were incubated for 3 min with 1 M sucrose in TEN buffer [0.01 M Tris·HCl (pH 7.4) with 0.1 M NaCl and 1 mM EDTA] and washed by floating the grids on two drops of TEN buffer before negative staining with 2% uranyl acetate for 30 s. The sucrose treatment lysed the adsorbed viruses and provoked the release of the ribonucleoproteins. The samples were subsequently visualized using a JEOL JEM 1011 electron microscope operating at 100 kV.

1. Ren L, et al. (2011) Structural insight into substrate specificity of human intestinal maltase-glucoamylase. *Protein Cell* 2(10):827–836.
2. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Macromolecular Crystallography, Part A*, eds Carter, Jr CW, Sweet RM (Academic, New York), pp 307–326.
3. Adams PD, et al. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* 58(Pt 11):1948–1954.
4. Emsley P, K. Cowtan (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.
5. Lovell SC, et al. (2003) Structure validation by Calpha geometry: Phi,psi and Cbeta deviation. *Proteins* 50(3):437–450.
6. DeLano W (2002) The PyMOL molecular graphics system (DeLano Scientific, San Carlos, CA). Available at www.pymol.org.
7. Ludtke SJ, Baldwin PR, Chiu W (1999) EMAN: Semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* 128(1):82–97.
8. Novoa RR, Calderita G, Cabezas P, Elliott RM, Risco C (2005) Key Golgi factors for structural and functional maturation of bunyamwera virus. *J Virol* 79(17):10852–10863.

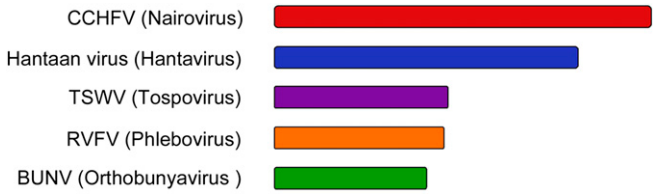

Fig. S1. Bunyavirus NPs differ from one another. The molecular weights of the prototypical members in each genus belonging to the *Bunyaviridae* family, that is, Crimean–Congo hemorrhagic fever virus (CCHFV; *Nairovirus*), Hantaan virus (*Hantavirus*), tomato spotted wilt virus (TSWV; *Tospovirus*), Rift Valley fever virus (RVFV; *Phlebovirus*), and Bunyamwera virus (BUNV; *Orthobunyavirus*), significantly differ from one another. The bars for each NP denote the molecular weight: CCHFV NP, 52 kDa; Hantaan virus NP, 42 kDa; TSWV, 29 kDa; RVFV, 27 kDa; BUNV, 25 kDa.

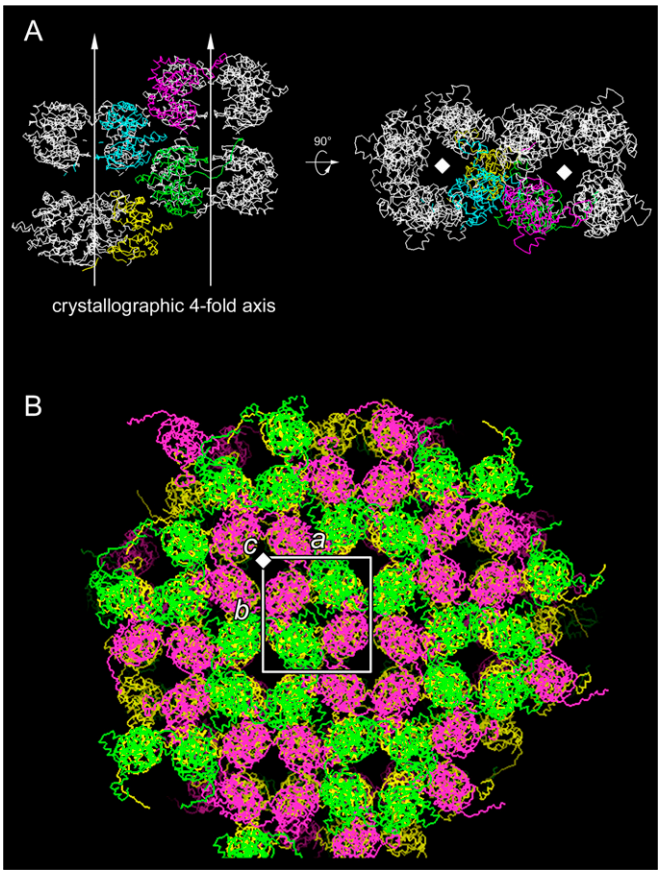

Fig. S2. Crystal packing of the BUNV NP–RNA complex. (*A*) Each BUNV NP-RNA molecule in one asymmetric unit forms the ring-shaped tetramer with three other molecules related by a crystallographic fourfold axis. The four molecules in one asymmetric unit are shown as colored ribbons, whereas the symmetric molecules are colored in gray. The crystallographic fourfold axes along the *c* axis of the *I422* space group are represented as arrows and solid squares in a side and top view, respectively. (*B*) Subunit packing in the crystal corresponding to the crystallographic [*a*,*b*] plane.

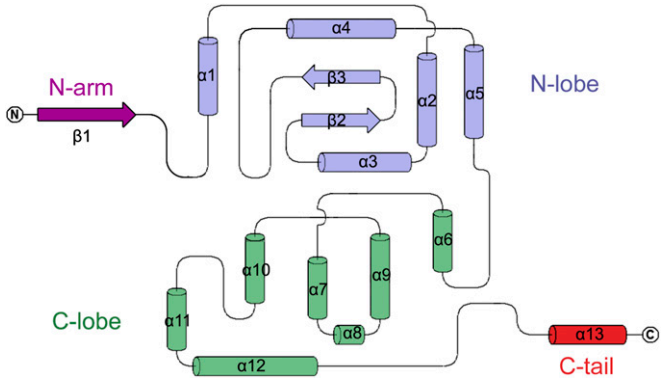

Fig. S3. Topology of the BUNV NP. The α-helices and β-strands are numbered. Residues at the boundary of the secondary structure elements are also numbered. The N-lobe region is colored in blue, and the C-lobe region is colored in green. The topology cartoon was drawn with TopDraw (1).

1. Bond CS (2003) TopDraw: A sketchpad for protein structure topology cartoons. *Bioinformatics* 19(2):311–312.

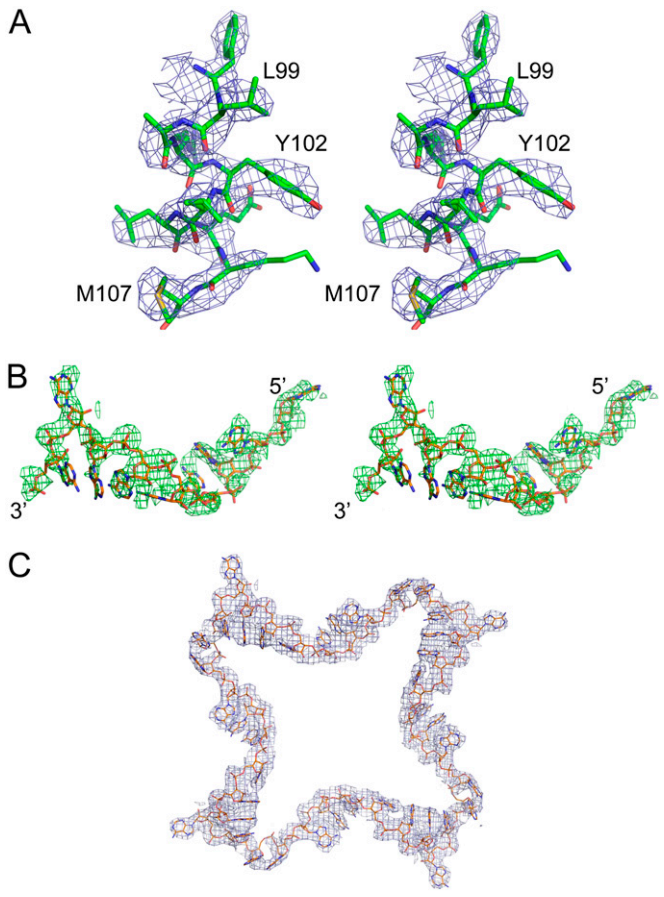

Fig. S4. Electron density of the BUNV NP polypeptide and bound RNA. (*A*) Omit map of a selected region of the BUNV NP polypeptide contoured at 1.0σ. Residues in this region (H98–M107) are shown as colored sticks. (*B*) Omit map of the bound RNA contoured at 1.0σ. The bound RNA molecule is represented as colored sticks. (*C*) Electron density ($2F_o$-F_c contoured at 1.3σ) of the nucleic acid bound to the BUNV NP tetramer.

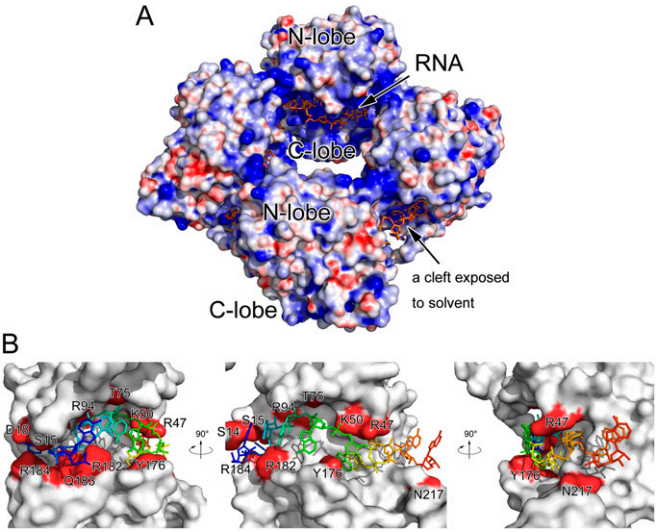

Fig. S5. RNA binding crevice. (*A*) The tetrameric nucleoprotein ring is depicted as an electrostatic potential surface (red and blue for acidic and basic, respectively). The RNA molecule is represented as colored sticks. The domains of the N subunit at the center are labeled (except for the N arm, which is buried and difficult to identify in this representation). (*B*) Enlarged views of the RNA binding crevice of the BUNV NP protomer. The color scheme of the RNA molecule is identical to that used in Fig. 4. The BUNV NP protomer is depicted in white, whereas the residues interacting with the RNA are colored in red.

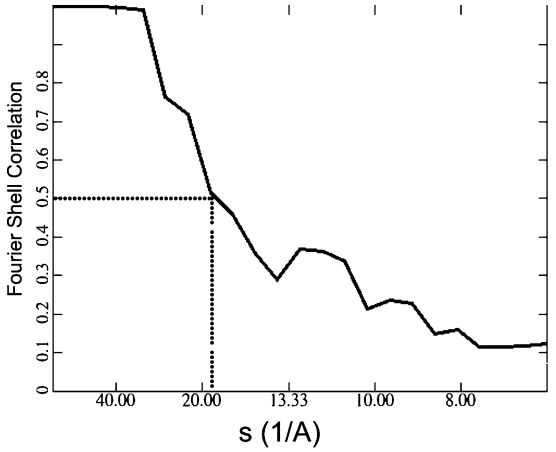

Fig. S6. Fourier shell correlation (FSC) curve of the 3D EM reconstruction of the tetrameric BUNV NP–RNA complex. The FSC curve suggests that the density map has a resolution of 18.9 Å according to the FSC = 0.5 criterion.

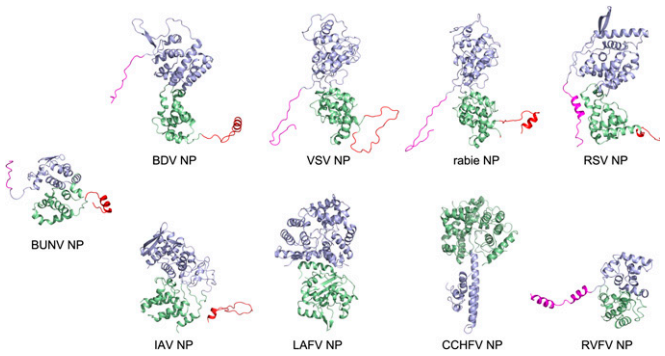

Fig. S7. Structures of the nucleocapsid proteins of negative-sense single-stranded RNA (–ssRNA) viruses. All molecules are shown as ribbon diagrams. The upper row presents the NPs of nonsegmented –ssRNA viruses; the lower row presents the NPs of segmented –ssRNA viruses. The molecules in the upper row are shown in an identical orientation, whereas the molecules in the lower row are turned to show the extensions involved in intermolecular interactions. All molecules are shown at an identical scale. A protomer of the BUNV NP is shown in the left panel. All bound RNAs were removed from these structures for clarity. The N- and C-terminal extensions for NP oligomerization are colored magenta and red; the N and C termini (or lobes) are colored blue and green, respectively. The Protein Data Bank codes for the structures are Borna disease virus (BDV), 1N93; vesicular stomatitis Indiana virus (VSV), 2GIC; rabies virus, 2GTT; respiratory syncytial virus (RSV), 2WJ8; influenza A virus, 2IQH; Lassa fever virus (LAFV), 3MX2; CCHFV, 3U3; and RVFV NP, 3OUO.

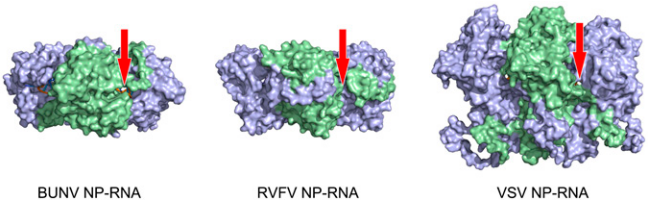

Fig. S8. Cleft between the protomer interface. NP–RNA complexes from BUNV, RVFV, and VSV are aligned and compared. The polypeptide for each complex is shown as a colored surface; the bound RNA molecule is shown as a colored cartoon diagram. The cleft between the protomer interface of the BUNV NP–RNA complex and identical positions in the RVFV and VSV are indicated by red arrows.

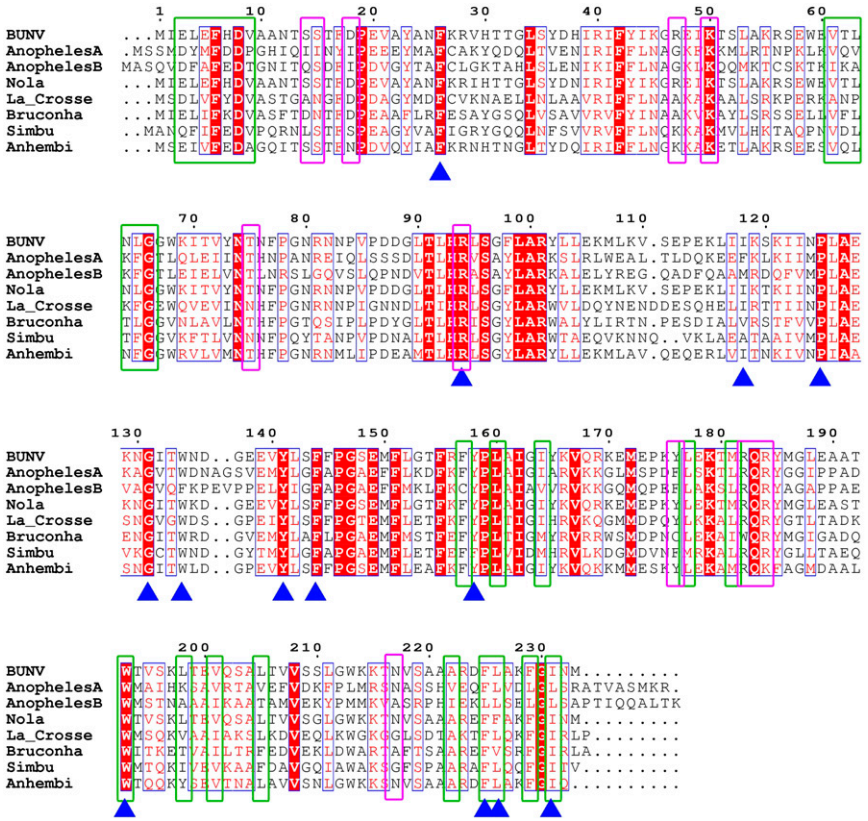

Fig. S9. Sequence alignment of the NP from the genus *Orthobunyavirus*. Variable residues are shown in black with white background, and conserved residues are shown in white with red background. The sequence alignment was generated by ClustalW (1) and plotted using ESPript (2). The green and magenta frames indicate the residues involved in oligomerization and those that directly interact with RNA, whereas the previously determined functionally important residues of the BUNV NP are indicated by blue triangles. The sequence alignment abbreviations are as follows: BUNV, Bunyamwera virus (Bunyamwera serotype); AnophelesA, Anopheles A virus (Anopheles A serogroup); AnophelesB, Anopheles B virus (Anopheles B serogroup); Nola, Nola virus (Bakau serogroup); La_Crosse, La Crosse virus (California serogroup); Bruconha, Bruconha virus (Group C serogroup); Simbu, Simbu virus (Simbu serogroup); and Anhembi, Anhembi virus strain SPAr2984 (Wyeomyia serogroup).

1. Larkin MA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947–2948.
2. Gouet P, Courcelle E, Stuart DI, Métoz F (1999) ESPript: Analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15(4):305–308.

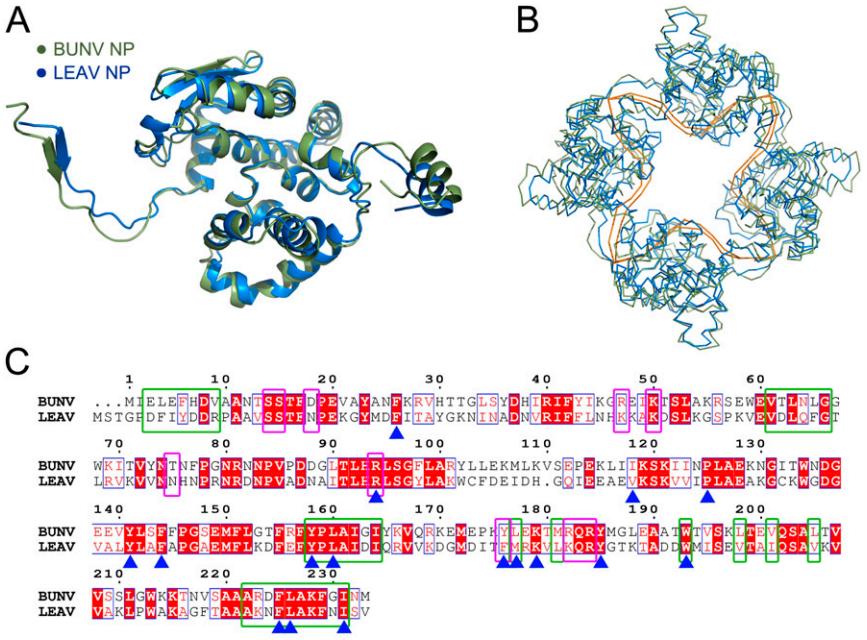

Fig. S10. Comparison of the NPs from BUNV and LEAV. (*A*) Superposition of the BUNV NP and LEAV NP monomers. Structures of the BUNV and LEAV NPs are represented as green and blue cartoons. (*B*) Comparison of the tetrameric rings formed by the BUNV and LEAV NP–RNA complexes. The polypeptides of the BUNV NP and LEAV NP are shown as green and blue ribbons, whereas the bound RNA molecules are presented as orange lines. (*C*) Primary sequence alignment of the BUNV NP and LEAV NP. The green and magenta frames indicate the residues involved in oligomerization and those that directly interact with RNA, whereas the previously determined functionally important residues of the BUNV NP are indicated by blue triangles.

Table S1. Data collection and refinement statistics

Parameters	Native	Se singlewavelength anomolous dispersion peak
Data collection statistics		
Cell parameters		
a, Å	106.2	106.0
b, Å	106.2	106.0
c, Å	485.7	486.6
α, β, γ,°	90.0, 90, 90.0	90.0, 90.0, 90.0
Space group	*I422*	*I422*
Wavelength used, Å	0.9798	0.9789
Resolution, Å	50.0–3.2 (3.26–3.20)‡	50.0–4.0 (4.06–4.00)
No. of all reflections	258,362 (12,040)	179,943 (6,255)
No. of unique reflections	23,467 (1,158)	24,235 (1,046)
Completeness, %	98.9 (99.9)	100.0 (100.0)
Average I/σ (I)	17.6 (7.2)	13.2 (11.0)
R_{merge},* %	7.8 (53.6)	9.1 (23.6)
Refinement statistics		
No. of reflections used [σ(F) > 0]	23,406	
R_{work},† %	22.0	
R_{free},† %	27.6	
Rmsd bond distance, Å	0.011	
Rmsd bond angle, °	1.548	
Avg B value, Å^2		
Avg B value for protein atoms	66.1	
Avg B value for ligand atoms	86.2	
Avg B value for solvent atoms	46.7	
No. of atoms		
No. of protein atoms	7,382	
No. of ligand atoms	880	
No. of solvent atoms	24	
Ramachandran plot		
Residues in favored regions, %	91.0	
Residues in allowed regions, %	8.8	
Residues in outlier regions, %	0.2	

*$R_{merge} = \Sigma_h\Sigma_I \mid I_{ih} - \langle I_h \rangle \mid / \Sigma_h\Sigma_I \langle I_h \rangle$, where $\langle I_h \rangle$ is the mean of the observations I_{ih} of reflection h.

†$R_{work} = \Sigma(||F_p(obs)| - |F_p(calc)||)/ \Sigma|F_p(obs)|$; R_{free} is an R factor for a preselected subset (5%) of reflections that was not included in refinement.

‡Numbers in parentheses are corresponding values for the highest resolution shell.