Supporting Information

Reguera et al. 10.1073/pnas.1302298110

SI Materials and Methods

Plasmids. The La Crosse Orthobunyavirus (LACV) nucleoprotein (NP) coding sequence (UniProt accession code P04873) was optimized for recombinant expression in *Escherichia coli* and synthesized (Geneart). A histidine tag and a tobacco etch virus (TEV) cleavage site (MGHHHHHHDYDIPTTENLYFQG) were added to the amino terminus of the NP coding sequence, resulting in an addition G at the N terminus after TEV cleavage. The sequence was cloned into a plasmid for expression by T7 RNA polymerase (pET) 9a (Novagen) modified vector between NdeI 5′ and NotI 3′ sites for expression in *E. coli*. Site-directed mutagenesis was performed by PCR using overlapping oligonucleotides and *Pyrococcus furiosus (pfu)* DNA polymerase.

Protein Expression and Purification. NP was expressed in *E. coli* BL21 (DE3) cells grown in LB with 50 μg/mL of kanamycin at 25 °C overnight after induction with 0.5 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG). Labeled NP was expressed in M9 minimal medium and 50 mg/L of seleno-methionine. The cells were disrupted by sonication 3–5 min on ice in lysis buffer (20 mM Tris·HCl pH 8, 150 mM NaCl, 2.5 μM β-mercapto-ethanol) with EDTA free protease inhibitor complex (Roche) and 20 mM imidazol. The protein from the soluble fraction was loaded onto 3 mL Nickel column (chelating Sepharose, GE) washed with 15 volumes of lysis buffer with 50 mM Imidazol and eluted with 5 volumes of 300 mM imidazol. The eluted protein was cleaved with histidine-tagged TEV protease from 24 to 48 h at 4 °C in dialysis against lysis buffer, resulting in an almost complete cleavage. A second nickel column step was performed to remove unwanted material. The resulting proteins were concentrated in Amicon Ultra centrifugal filter devices and purified by gel filtration chromatography using SD 75 or SD 200 columns (Pharmacia) with lysis buffer. Homogeneous tetrameric NPs were purified by treating the NPs from the second nickel column step with 20 μg/mL of RNase A overnight at 4 °C previous to the gel filtration. Monomeric NP was purified by disrupting the tetramers by adding 10 μg/mL of RNase A and 0.4 M Na thiocyanate in the lysis buffer during sonication and incubating the clarified lysate overnight at 4 °C. The first Ni column is washed with 10 volumes of lysis buffer with 0.4 M thiocyanate and 50 mM imidazole and eluted with 5 volumes of 300 mM imidazole. The thiocyanate was removed by 24–48 h dialysis at 4 °C against lysis buffer in the presence of TEV protease. The resulting monomeric NP was then concentrated and purified free of TEV and remaining tetrameric protein by heparin affinity chromatography. After binding in the dialysis buffer, the column is washed with 10 volumes of lysis buffer with 250 mM NaCl and eluted with 600 mM NaCl. A final gel filtration using a SD75 or SD200 column and the final lysis buffer gave homogeneous, pure, and RNA-free monomeric protein, as judged by the 260/280 absorbance ratio of 0.57.

Crystallization. The first crystals were obtained with *E. coli* expressed tetrameric NP treated with RNase A before size exclusion purification. The crystals grew after 3 wk in sitting drops from a 1:1 ratio of 15–20 mg/mL protein in lysis buffer and reservoir composition of 0.2 M $MgCl_2$ 0.1 M Tris·HCl pH 8.5, 15% (wt/vol) PEG 4K, and 0.2 M Na-thiocyanate at 4 °C. Derivative selenomethionine (SeMet) crystals were reproduced under the same conditions but at 20 °C using crushed native crystals for seeding. The crystals were frozen in liquid nitrogen with reservoir buffer plus 15% (vol/vol) glycerol. Native crystals with better resolution were grown with a reservoir composition of 0.2 M $LiSO_4$, 0.1 M Tris·HCl pH 8.5, 15% (wt/vol) PEG 4K using tetrameric NP treated overnight at 4 °C with RNase A (2 μg/mL) in lysis buffer plus 0.4 M Na-thiocyanate before the size exclusion. From the same sample, a different crystal form ($P4_1$, a = b = 133.3, c = 53.9 Å, 7.4 Å resolution) was grown in 10 mM Mg-acetate, 50 mM Na-cacodylate, 1.3 M $LiSO_4$, and 0.1 $SrCl_2$. For the best NP–RNA complex crystals ($P2_1$, a = 82.9, b = 86.5, c = 90.2 Å, β = 106°, 3.4 Å resolution), the monomeric NP was mixed in a 1.2:4 RNA–protein molar ratio with 45 nucleotides (nts) long U-rich RNA (see below) and incubated overnight at 4 °C before concentrating to 15–20 mg/mL. Crystals grew as thin bars with a reservoir buffer of 0.1 M bicine pH 9 and 10% (vol/vol) PEG 5K monomethyl ether (MME). Similar cocrystallization with a 1.2:1 ratio of 11 nt RNA (see below) yielded two other crystal forms: a C2 form (a = 237.0, b = 121.0, c = 79.6 Å, β = 109.1°, 3.7 Å resolution) in hanging drops with 0.1 M citric acid pH 5.2 and 10% (vol/vol) PEG 5K MME, and a second $P2_1$ form (a = 78.7, b = 116.1, c = 220.2 Å, β = 90.3°, 4.2 Å resolution) in sitting drops with 0.1 M citric acid pH 5.2 and 12% (vol/vol) PEG 5K MME. Cocrystallization with a 1.5:1 ratio gives a $P2_12_12_1$ form (a = 78.6, b = 126.2, c = 364.8 Å, 4.0 Å resolution) with 0.1 M citric acid pH 4 and 5% (vol/vol) PEG 5K MME. More details of all these crystal forms are given in Table S2.

Crystallography. The RNA-free NP (apo-NP) crystals are of space-group $P2_12_12_1$ with one molecule per asymmetric unit. A selenomethionine derivative dataset was collected at 2.8 Å resolution on ID14-4 at the European Synchrotron Radiation Facility (ESRF) at the selenium edge (0.979 Å) for single-wavelength anomalous diffraction (SAD) phasing. Native apo-NP data were collected to 1.8 Å on ID23-2 at the ESRF at a wavelength of 0.872 Å. All data were processed using the XDS package (1) and subsequent analysis performed with the CCP4i package (Table S1). The apo-NP structure was solved by the SAD method using SHELX (2), which found six anomalous sites belonging to the six selenomethionines of the monomer. SHARP (3) was used for selenium site refinement, phasing, and density modification of the electron density maps. With the resulting map, an initial model was built using the BUCANEER (4) automated model-building program. This model was transferred to the 1.8 Å resolution apo-NP data by molecular replacement by PHASER (5) and refined using REFMAC (6). The NP–RNA (45 nts) crystals are of space-group $P2_1$, and a 3.4 Å resolution dataset was collected on ID23-2. The structure was solved by molecular replacement with the apo-NP structure and revealed four proteins in the asymmetric unit corresponding to a tetrameric protein–RNA complex. Extra density in the positively charged groove of each NP could be modeled as a continuous poly(U) ssRNA molecule of 44 nts. During refinement, map-sharpening via REFMAC was used to improve the electron density.

The $P4_1$ crystal dataset was collected on ESRF ID23-2–7.4 Å resolution and solved by molecular replacement using the arm-less core of the apo-structure. Two molecules were unambiguously placed in the asymmetric unit by PHASER, each one generating separate chains throughout the crystal mediated by *N*- and C-arm intermolecular interactions. One independent molecule generates, by virtue of the crystallographic screw axis, the 4_1 helix described in the text. The *C*2 and $P2_1$ crystal forms of the NP+11 nts RNA were collected on ESRF beamline ID23-1 at 3.6 and 4.2 Å resolution, respectively, and the $P2_12_12_1$ form at 4.0 Å resolution on ID23-2. All these crystal forms were solved by molecular replacement and revealed different numbers of RNA-bound tetramers and sometimes extra monomers in the asymmetric unit (Table S2).

The sequence alignment of Fig. 3 was drawn with CLUSTAL W (7) and ESPript (8).

RNA Production and Sequences. RNAs were in vitro transcribed by standard procedures using T7 polymerase. Radio-labeled RNA was transcribed with alpha-^{32}P-uridine triphosphate (^{32}P-UTP). RNAs were purified on 10–15% (wt/vol) acrylamide, 8 M Urea gels in tris-borate-EDTA (TBE) buffer. PH60 RNA corresponds to the LACV M segment RNA panhandle structure including the first and last 28 nts connected by a linker and with sequence 5′-AGUAGUGUACUACCAAGUAUAGAUAACGUUUAACA-UUAUUUAUACGUGGAGCACACUACU-3′. U-rich RNA was transcribed as described elsewhere (9). For crystallization, 45 and 11 nts RNAs were purchased from SIGMA, the sequences being 5′-GGCCAUCCUGUUUUUUUCCCUUUUUUUUUU-UCUUUUUUUUUUUUU-3′ and 5′-ACAUCGGUACA-3′, respectively.

Characterization of Protein–RNA Interactions. For the Electrophoresis Mobility Shift Assay (EMSA) experiments, monomeric NP was incubated at 12 μM with the ^{32}P-labeled RNA for 1 h at room temperature and then loaded into a native Tris-glycine 8% (wt/vol) acrylamide gel with 10 mM $MgSO_4$. We included 1:3 dilutions of 12 μM U-rich or PH60 into the mix for the competition assay.

For the Multi Angle Laser Light Scattering measurements, *N*-RNA complexes were loaded into a size exclusion column SD200 (Pharmacia) coupled to a multiangle light scattering detector [DAWN enhanced optical system (EOS), Wyatt Technology] as described (10).

Purification of wild-type LACV. Baby hamster kidney (BHK-21) cells grown in 32 T175 tissue culture flasks to 40% confluency were infected with LACV at an multiplicity of infection (MOI) of 0.001. After 3 d of incubation, cytopathic effects were visible and supernatants were harvested and cleared by centrifuging at 3,000 rpm in a JLA 16.250 rotor. Virus particles were concentrated by centrifuging the cleared supernatants in SW32 tubes through 5 mL of 30% (vol/vol) glycerol/PBS at 25,000 rpm in an SW32 rotor, 4 °C for 90 min. Virus pellets were resuspended and pooled in 2 mL of PBS and stored at 4 °C.

RNP purification. Purified LACV virions were incubated in 10 mM Tris pH 7.8, 200 mM NaCl, 0.1% Nonidet P-40 at 4 °C during 30 min to disrupt the virion membrane. After centrifugation 10 min at 10,000 × *g*, the disrupted virions were applied on top of a 20–40% (wt/vol) CsCl gradient containing 20 mM Tris pH 7.8 and 200 mM NaCl, and ultracentrifuged 2 h at 52,000 × *g* in a SW60 swing-out rotor (Beckman). Ribonucleoprotein particle (RNP)–containing fractions were identified on SDS/PAGE gel and dialyzed against a buffer containing 20 mM Tris pH 7.8 and 200 mM NaCl. Some samples were heated 1 h at 37 °C before EM observation.

Electron microscopy. Tetrameric recombinant LACV NP was applied to the clear side of carbon on a carbon-mica interface and stained with 1% (wt/vol) sodium silico-tungstate pH 7.5. Images were recorded under low-dose conditions with a JEOL 1200 EX II microscope at 100 kV and at a nominal 50,000× magnification. Negatives were digitized on a Zeiss scanner (Photoscan TD) with a pixel size of 2.8 Å at the object scale. A total of 5,000 LACV NP tetramers were semiautomatically selected with Boxer (11), contrast transfer function (CTF)-corrected with CTFFIND3 (12) and Bsoft (13), and subjected to multivariate statistical analysis and classification with IMAGIC-5 (14). Ab initio class averages were aligned with SPIDER (15) against projections of tetrameric LACV NP crystal structure filtered to 20 Å.

Disrupted LACV virions and purified LACV RNPs were applied for 30 s to a glow-discharged continuous carbon grid, blotted, stained for 30 s with 2% (wt/vol) uranyl acetate, and imaged as described above. Images were recorded on a GATAN Orius SC600 2k × 2k CCD camera at 52,970× magnification.

1. Kabsch W (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):133–144.
2. Schneider TR, Sheldrick GM (2002) Substructure solution with SHELXD. *Acta Crystallogr* 58(Pt 10 Pt 2):1772–1779.
3. de La Fortelle E, Bricogne G (1997) Heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods in Enzymology*, eds Carter CW Jr, Sweet RM (Academic, New York), Vol 276, pp 472–494.
4. Cowtan K (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr* 62(Pt 9):1002–1011.
5. Read RJ (2001) Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallogr D Biol Crystallogr* 57(Pt 10):1373–1382.
6. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53(Pt 3):240–255.
7. Thompson JD, Higgins DG, Gibson TJ (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(22):4673–4680.
8. Gouet P, Courcelle E, Stuart DI, Métoz F (1999) ESPript: Analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15(4):305–308.
9. Reguera J, Weber F, Cusack S (2010) Bunyaviridae RNA polymerases (L-protein) have an N-terminal, influenza-like endonuclease domain, essential for viral cap-dependent transcription. *PLoS Pathog* 6(9):e1001101.
10. Gerard FC, et al. (2007) Unphosphorylated rhabdoviridae phosphoproteins form elongated dimers in solution. *Biochemistry* 46(36):10328–10338.
11. Ludtke SJ, Baldwin PR, Chiu W (1999) EMAN: Semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* 128(1):82–97.
12. Mindell JA, Grigorieff N (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* 142(3):334–347.
13. Heymann JB (2001) Bsoft: Image and molecular processing in electron microscopy. *J Struct Biol* 133(2-3):156–169.
14. van Heel M, Harauz G, Orlova EV, Schmidt R, Schatz M (1996) A new generation of the IMAGIC image processing system. *J Struct Biol* 116(1):17–24.
15. Frank J, et al. (1996) SPIDER and WEB: Processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* 116(1):190–199.

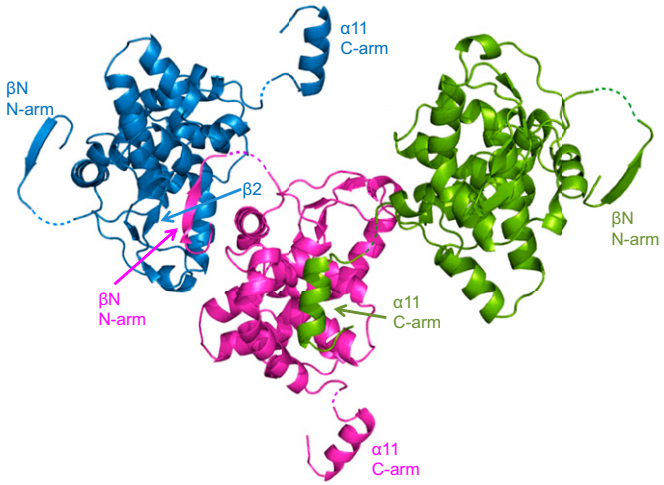

Fig. S1. Molecular packing in the RNA-free NP crystals. Crystal symmetry in the $P2_12_12_1$ space-group generates a 3D network throughout the crystal from the one molecule in the asymmetric unit. Three neighboring molecules, colored in magenta, green, and blue, are shown in ribbon representation with the *N*- and C-arms that mediate intermolecular interactions indicated.

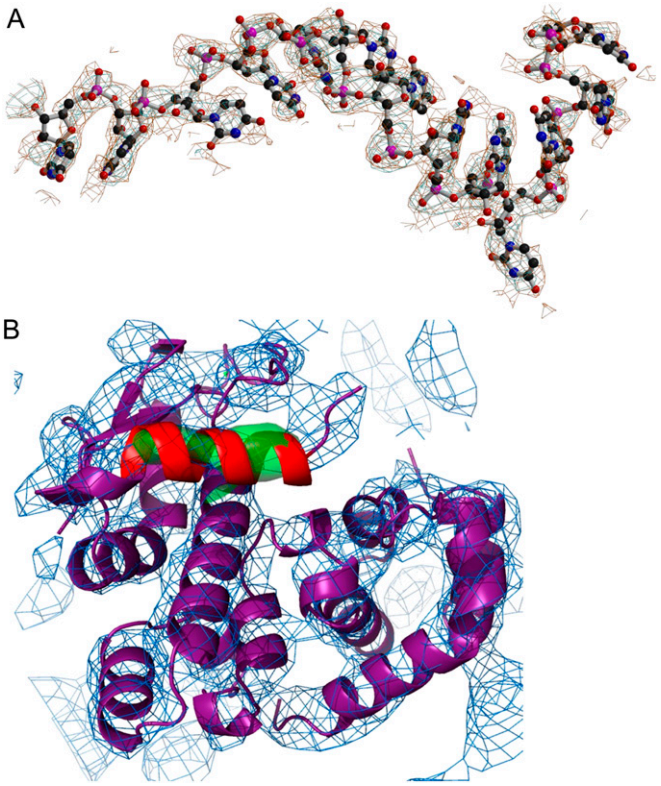

Fig. S2. Examples of electron density. (*A*) Electron density for the RNA in the LACV tetramer complex with 45 nts RNA. 2Fobs-Fcalc map contoured at 0.9 σ after refinement with the displayed 13 nts RNA omitted from model. (*B*) Electron density for one of the two independent molecules within the $P4_1$ crystal form of LACV NP after molecular replacement and without further refinement (blue, contour level 1.3 σ). PHASER gave a unique solution with log likelihood gain (LLG) of 197 and no clashes. When residues 19–30 of helix α1 (red) was omitted from the molecular replacement model, a positive difference density is observed at its position (green, contour level 2.5 σ).

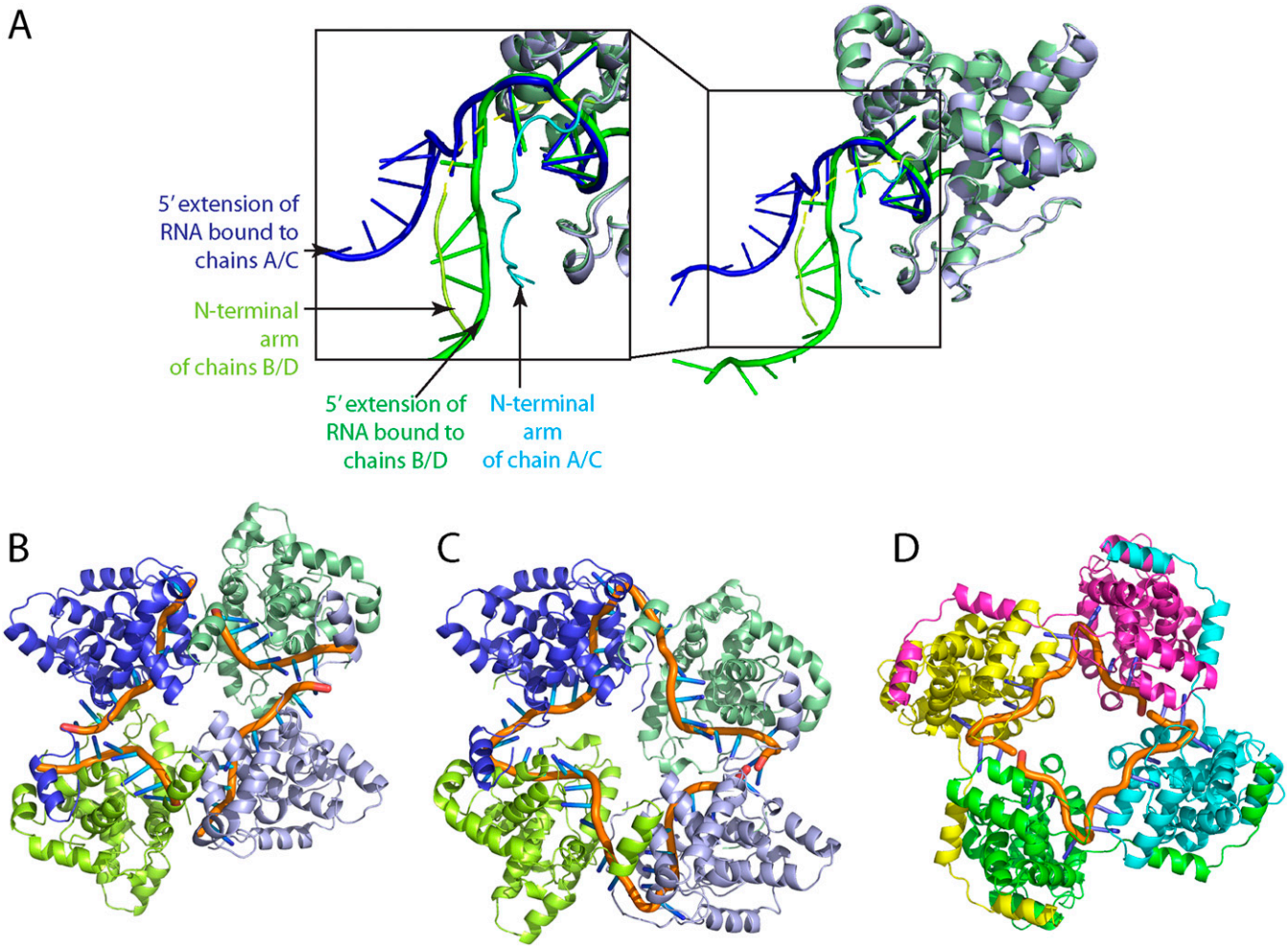

Fig. S3. Comparison of different LACV and Rift Valley Fever Phlebovirus (RVFV) NP–RNA tetramers. (*A*) Different *N*- and C-arm conformation and RNA bending at the two nonequivalent NP–NP interfaces of the LACV NP tetramer with 44 nts RNA. At the A/B interface, the bend is less sharp (65–70°), and there is room for an ordered chain A *N*-arm hinge. The three interface nucleotides U11′–U1–U2 are sandwiched between the chain A *N*-hinge (residues 9–13 contact the bases) and the chain B C-arm hinge (residues 217–219 contacts the RNA backbone). At the B/C interface, the RNA is bent more tightly (80–85°) and does not allow the same positioning of the chain B *N*-arm hinge, which remains disordered. This leaves the backbone of the interface nucleotides still interacting with the C-hinge of chain C (as well as Lys179 of chain B), but the bases are more solvent exposed. (*B*) LACV NP tetramer in complex with four separate RNA molecules as observed in the *C*2 cocrystal with 11 nts RNA. The chain in dark blue is shown in the same orientation as in *C*. (*C*) Cartoon representation of LACV NP tetramer in complex with 44 nts RNA. Protein colored as in Fig. 2, RNA shown as an orange tube. (*D*) RVFV NP–RNA tetramer (PDB 4H5P). The RVFV NP contacts only 7 nts (three of which are in the NP–NP interface) compared with 11 nts for LACV NP.

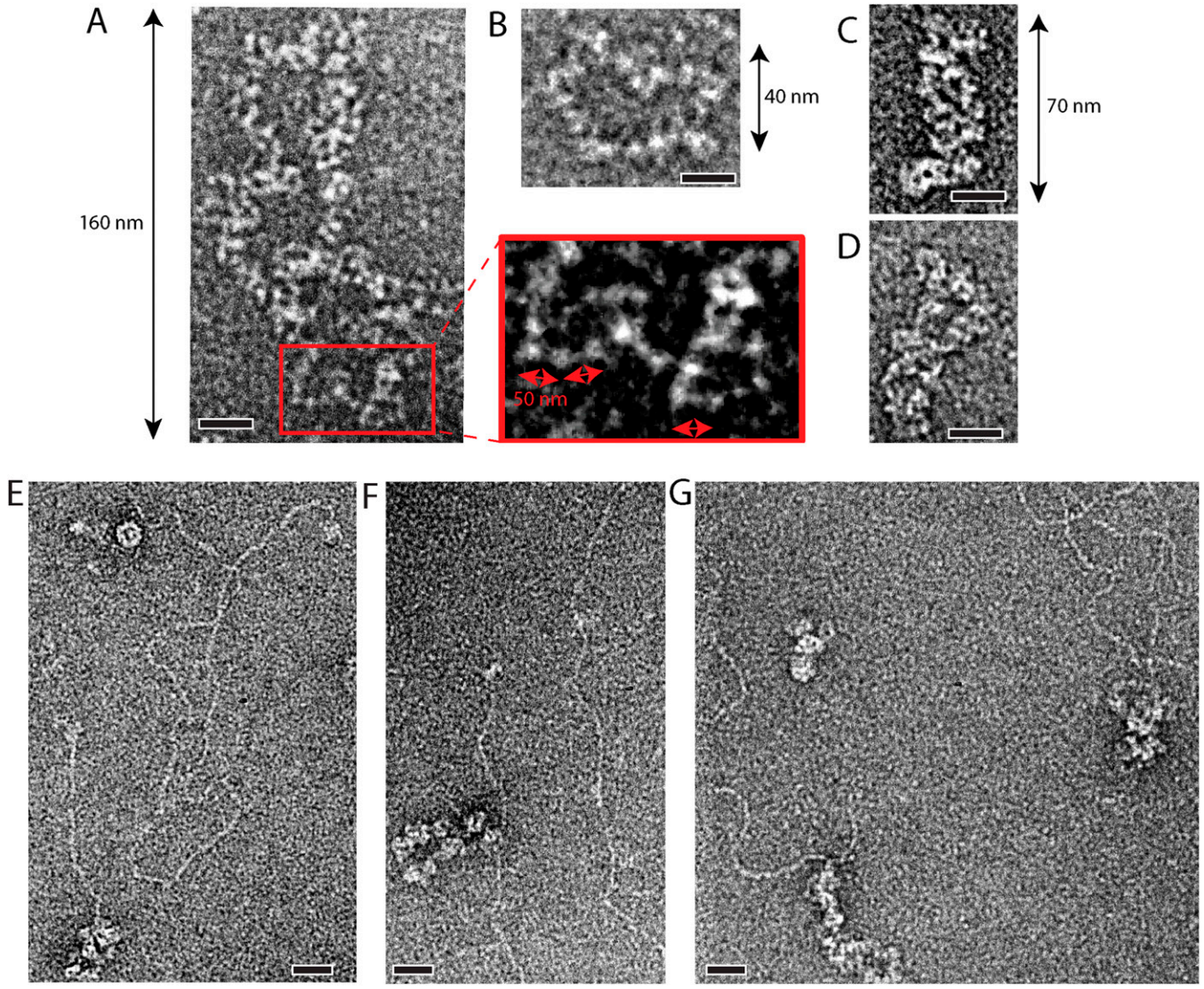

Fig. S4. EM of native LACV RNPs negatively stained with 2% uranyl-acetate. (*A*) RNP from membrane-disrupted LACV virion, likely to correspond to an M genome segment. Segments with individual NPs bound to RNA are highlighted and shown in a close-up view. (*B*) RNP from membrane disrupted LACV virion, likely corresponding to an S genome segment. (*C* and *D*) Purified LACV RNPs displaying a supercoiled form. (*E*–*G*) RNPs partially unwound after exposition incubation at 37 °C displaying long necklaces of NP bound to RNA emerging from more condensed regions. It was verified by SDS/Page gel analysis that the integrity of NP is not affected by this treatment.

Table S1. Crystallographic data collection and refinement statistics

	RNA-free LACV NP	RNA-free LACV NP SeMet	LACV NP tetramer with 45 nt RNA
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_1$
Cell dimensions, Å	a = 62.6, b = 70.1, c = 70.7	a = 62.6, b = 67.4, c = 70.4	a = 82.9, b = 86.5 ,c = 90.2, β = 106.6°
Solvent content, %			
Resolution range (last shell), Å	50–1.8 (1.89–1.80)	50–2.80 (2.90–2.80)	50–3.4 (3.5–3.4)
Beamline	ESRF ID23-2	ESRF ID14-4	ESRF ID23-2
Wavelength, Å	0.8726	0.9793	0.8726
Detector	MAR 225	ADSC Quantum Q315r	MAR 225
Completeness (last shell), %	100 (100)	99.9 (100)	99.6 (99.6)
Redundancy (last shell)	7.3 (7.4)	3.8 (3.8)	3.4 (3.4)
R_{merge} (last shell)	5.4 (74.2)	9.3 (57.7)	17.6 (78.7)
I/σI (last shell)	22.97 (2.72)	13.38 (2.51)	6.89 (1.89)
No. of reflections used in refinement (free reflections)	27,988 (1500)	8,646 (437)	17,489 (944)
R_{factor} (last shell)	0.182 (0.250)	—	0.205 (0.329)
R_{free} (last shell)	0.206 (0.226)	—	0.258 (0.397)
No.of nonhydrogen			
Atoms	2,121		8,159
Protein	1,870		7,282 (chains A, B, C, D)
Ligand	16 (1 glycerol, 2 SO4)		877 [44 nts poly(U), chain U)]
Water	240		—
Average B-value, Å^2	28.8		84.4
Ramachandran plot*			
Favored regions	98.3		90.7
Allowed regions	100		99.1
Disallowed regions	—		8 outliers in hinge regions
Deviations from ideal geometry			
Bond distances, Å	0.009		0.007
Angles, °	1.28		1.12

*Molprobity; available at http://molprobity.biochem.duke.edu.

Table S2. Summary characteristics of the different LACV NP crystal datasets

Space-group	Unit cell	Resolution, Å	Contents of asymmetric unit
$P2_12_12_1$	a = 62.6, b = 70.1, c = 70.7 Å	1.8	1 NP
$P4_1$	a = b = 133.3, c = 53.9 Å	7.4	2 NP
$P2_1$	a = 78.7, b = 116.1, c = 220.2 Å, β = 90.3°	4.2	2 tetramers + 4 NP
			10 NPs with 11 nts RNA
$C2$	a = 237.0, b = 121.0, c = 79.6 Å, β = 109.1°	3.7	1 tetramers + 2 NP
			5 NPs with 11 nts RNA
$P2_12_12_1$	a = 78.6, b = 126.2, c = 364.8 Å	4.0	3 tetramers
			Each NP with 11 nt RNA
$P2_1$	a = 82.9, b = 86.5, c = 90.2 Å, β = 106°	3.4	1 tetramer +45 nts RNA