

# The Crystal Structure and RNA-Binding of an Orthomyxovirus Nucleoprotein

Wenjie Zheng<sup>1</sup>, John Olson<sup>1</sup>, Vikram Vakharia<sup>2</sup>, Yizhi Jane Tao<sup>1\*</sup><sup>1</sup> Department of Biochemistry and Cell Biology, Rice University, Houston, Texas, United States of America, <sup>2</sup> Department of Marine Biotechnology, University of Maryland Baltimore County, Institute of Marine and Environmental Technology, Baltimore, Maryland, United States of America

## Abstract

Genome packaging for viruses with segmented genomes is often a complex problem. This is particularly true for influenza viruses and other orthomyxoviruses, whose genome consists of multiple negative-sense RNAs encapsidated as ribonucleoprotein (RNP) complexes. To better understand the structural features of orthomyxovirus RNPs that allow them to be packaged, we determined the crystal structure of the nucleoprotein (NP) of a fish orthomyxovirus, the infectious salmon anemia virus (ISAV) (genus *Isavirus*). As the major protein component of the RNPs, ISAV-NP possesses a bi-lobular structure similar to the influenza virus NP. Because both RNA-free and RNA-bound ISAV NP forms stable dimers in solution, we were able to measure the NP RNA binding affinity as well as the stoichiometry using recombinant proteins and synthetic oligos. Our RNA binding analysis revealed that each ISAV-NP binds ~12 nts of RNA, shorter than the 24–28 nts originally estimated for the influenza A virus NP based on population average. The 12-nt stoichiometry was further confirmed by results from electron microscopy and dynamic light scattering. Considering that RNPs of ISAV and the influenza viruses have similar morphologies and dimensions, our findings suggest that NP-free RNA may exist on orthomyxovirus RNPs, and selective RNP packaging may be accomplished through direct RNA-RNA interactions.

**Citation:** Zheng W, Olson J, Vakharia V, Tao YJ (2013) The Crystal Structure and RNA-Binding of an Orthomyxovirus Nucleoprotein. *PLoS Pathog* 9(9): e1003624. doi:10.1371/journal.ppat.1003624

**Editor:** Félix A. Rey, Institut Pasteur, France

**Received:** April 16, 2013; **Accepted:** July 31, 2013; **Published:** September 12, 2013

**Copyright:** © 2013 Zheng et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by the Welch Foundation (C-1565 to YJT and C-0612 to JO), the National Institutes of Health (AI067638 to YJT and GM35649 to JO), and the Kresge Science Initiative Endowment Fund at Rice University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: [ytao@rice.edu](mailto:ytao@rice.edu)

## Introduction

Influenza viruses represent a serious public health concern due to their potential in causing widespread epidemics and pandemics. The genomes of the influenza A and B viruses are both comprised of eight segments of negative-sense RNA [1]. These eight gene segments are encapsidated in the form of rod-shaped, double-helical ribonucleoprotein (RNP) complexes in virions as well as in infected cells [2]. Each RNP contains a viral RNA, a heterotrimeric viral polymerase (consisting of PA, PB1, and PB2) and multiple copies of the viral-encoded nucleoprotein (NP) that bind viral RNA in a stoichiometric manner.

The segmented genome of influenza viruses offers a number of replication advantages such as gene recombination and transcriptional regulation, but at the same time it creates challenges for genome packaging. Once controversial, it is now widely accepted that the vRNPs of the influenza A viruses are specifically packaged into the virion, so that each particle contains eight unique segments of viral RNA (vRNA). Early evidence from defective-interfering RNAs (DI RNAs) and reverse genetics [3,4] indicated that each gene segment possesses distinct packaging signals, and that unique sequences from both non-coding and coding regions at the 5' and 3' ends of all eight vRNA segments are essential for vRNA packaging [5–9]. Electron microscopy and tomography images showed that vRNPs in budding virions were organized in a distinct “7+1” pattern, with a long segment in the center surrounded by seven segments of various lengths at the periphery

[10–13]. During virus assembly, vRNPs are aligned at the budding tip and interconnect with each other to form a supra-molecular assembly. Fluorescence *in situ* hybridization (FISH) analysis of single virus particles further confirmed that the eight unique RNPs are selectively incorporated into progeny virions and most virus particles contain only one copy of each of the eight RNP complexes [14].

Early electron microscopy studies of isolated influenza A virus RNPs revealed rod-shaped structures with a double-helical arrangement [15,16]. Recently, cryo-EM reconstructions of the influenza A virus RNP have been calculated to ~20 Å resolution [17,18]. The double-helical stem region of the RNP shows a rise between two neighboring NP of ~30 Å with five to six NP molecules per turn. Interaction between the two opposing arms of an RNP hairpin is solely mediated by the NP. Adjacent NP molecules on the same RNA strand interact through the extended NP tail loop. The putative RNA binding groove of the NP scaffold is exposed on the outer surface of the RNP.

Previous studies based on genome-wide stoichiometry calculations showed that there are ~24–28 nts of RNA for each NP molecule [15,19,20]. In the absence of NP, vRNAs are able to mediate a comprehensive interaction network, with each vRNA segment interacting with at least one other vRNA partner, suggesting that RNPs are likely held together by direct base-pairings between packaging signals [12]. This model raises an intriguing question as to how the packaging signals are presented on the surface of the RNPs. Is vRNA maintained in an extended

## Author Summary

Orthomyxoviruses are a family of RNA viruses that include the various types of influenza viruses. The genome of orthomyxoviruses consists of multiple segments of negative-sense, single-stranded RNA molecules, each packaged in the form of rod-shaped, double-helical ribonucleoprotein (RNP) complexes. How different RNPs interact with each other to ensure specific genome packaging is a long-standing question and crucial to our understanding of orthomyxovirus replication and influenza virus gene reassortment. Our study of a fish orthomyxovirus, the infectious salmon anemia virus (ISAV), shows that its nucleoprotein (NP), which forms the protein scaffold backbone of the viral RNP, has a bi-lobular structure like the influenza virus NP. Because ISAV-NP forms stable dimers in solution, we were able to determine ISAV-NP RNA binding stoichiometry by biochemical assays, electron microscopy and dynamic light scattering. Our results indicate that each ISAV-NP binds ~12-nt RNA, shorter than the 24–28 nts originally estimated for the influenza A virus based on population average. We propose that NP-free RNA exists on orthomyxovirus RNPs, and such RNA regions likely mediate specific RNP-RNP interactions during genome packaging. Further elucidation of the RNA-mediated RNP-RNP interactions will help us determine the molecular basis of gene reassortment by orthomyxoviruses including the influenza viruses.

configuration devoid of any secondary structure by tight wrapping around the NP, or perhaps some NP-free RNAs exist in the RNP to allow vRNA-vRNA interaction?

To better understand the mechanism of influenza virus RNP packaging, we performed a series of structural and functional studies of the NP from the infectious salmon anemia virus (ISAV). ISAV is the first orthomyxovirus isolated from fish, and it (genus: *Isavirus*) is capable of causing wide-spread, severe anemia in Atlantic salmon and other salmonid species [21,22]. Although their gene sequences are highly divergent, ISAV and influenza viruses share many similarities in their genomic, structural, morphological, and physicochemical properties. Here we determined the first crystal structure of an ISAV NP to 2.7 Å resolution. The crystal structure of ISAV-NP shows a head and a body domain that resemble those of the influenza A virus NP, despite the lack of sequence similarity. Using fluorescence anisotropy assays with synthetic RNA oligos, we showed that each ISAV-NP molecule binds to 12 nts of RNA, much shorter than the 24 nts originally estimated by genome-wide calculation for the related influenza A virus [15,19,20]. The 12-nt RNA binding stoichiometry is further supported by results obtained by electron microscopy and dynamic light scattering, as the addition of a 48-nt RNA to a solution of ISAV-NP led to the formation of NP tetramers or the dimerization of NP dimers. Our results thus suggest that a considerable amount of genomic RNA may exist in a NP-free state within the orthomyxovirus RNPs, and this free RNA may mediate the formation of high-order structures necessary for the specific recognition of the eight individual RNPs during virus assembly.

## Results and Discussion

### ISAV-NP forms stable dimers in solution

ISAV-NP was overexpressed in *E. coli* and purified as dimers with an approximate molecular weight of 140 kD, as estimated from the gel filtration chromatogram (Fig. 1). The ISAV-NP dimer

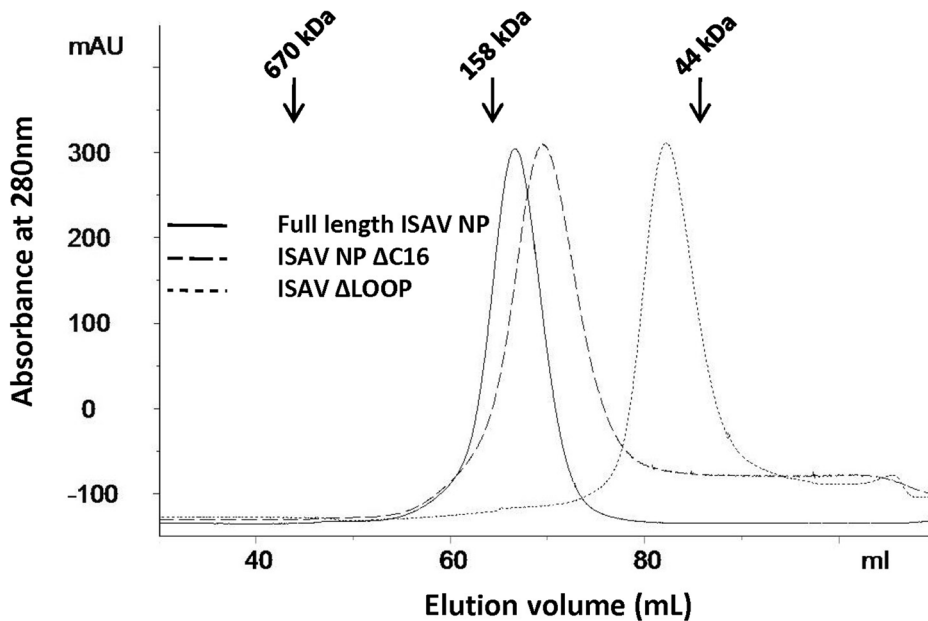
was stable and did not dissociate under a wide range of conditions when tested *in vitro* (e.g. 0.01 to 10 mg/mL protein, NaCl ranging from 0.05 to 1 M, at 4 or 25°C for up to one month, etc). These unusual features of ISAV-NP make it an excellent system for studying NP RNA binding properties for the orthomyxoviruses, as the recombinant NPs of both influenza A and B viruses are known to form multiple oligomeric species [23–25]. Furthermore, influenza A virus NP oligomers and monomers undergo dynamic equilibrium in solution [23], and the addition of short RNA oligos led to the formation of larger oligomers possibly due to quaternary structure re-arrangement [24,26]. RNPs from ISAV and influenza viruses exhibit similar double-helical morphologies in EM [15,17,18,27].

### Crystal structure of ISAV-NP

Crystals of the full-length ISAV-NP were obtained, which only diffracted to ~10 Å resolution. A number of N-terminal truncation mutants, including ΔN39, ΔN50, ΔN64, and ΔN84, and ΔN111, were subsequently cloned and expressed. Although ΔN50, ΔN64, and ΔN84 produced crystals, their diffraction was also poor. Visual inspection of the ISAV-NP sequence revealed a highly acidic C-terminal tail (i.e. <sup>600</sup>EIEFDEDDDEEEEDIDI<sup>616</sup>, with 12 out of the last 16 residues being either Glu or Asp); a feature that has been previously noted in many orthomyxovirus NPs. Without the last 16 residues, the ISAV-NP ΔC16 mutant also formed dimers, similar to the full-length protein (Fig. 1), and produced crystals that diffracted far better up to 2.7 Å resolution. The slight increase in the elution volume of the ISAV-NP ΔC16 mutant in gel filtration chromatography is likely due to changes in NP's radius of gyration or differences in protein-resin interaction.

The crystal structure of the ISAV-NP ΔC16 was determined by single-wavelength anomalous dispersion (SAD). The final model contains 495 amino acid residues in total, including residues 5–25, 27–38, 41–45, 61–75, 86–112, 133–182, 201–320, 325–497, 504–530, and 542–586. Each NP molecule can be divided into three domains: an N-terminal domain, a head domain, and a body domain (Fig. 2A). While the N-terminal domain is made up of a single polypeptide stretch containing the first 111 residues, the head and body domains are not colinear, with the head domain made from two excursions of the polypeptide chain (residues 272–380 and 494–558) out of the body domain. Structural comparison shows that the head and body domains of ISAV-NP closely resemble those of the influenza A virus NP, despite the lack of significant sequence similarity (Figure 2B, E).

Although there is only one NP in each crystallographic asymmetric unit, dimers are found around the crystallographic two-fold symmetry axes (Fig. 2A), which are presumably the same as those found in the solution (Fig. 1). The tail loop structure, which plays a critical role in influenza A virus NP oligomerization, is also found in ISAV-NP, although these two NP proteins form oligomers of different sizes (dimer vs. trimer). The tail loop of ISAV-NP (residues 504–530) folds into an extended β-hairpin that protrudes into a deep pocket in a neighboring subunit to mediate dimer formation. The interaction made by the tail loop is primarily hydrophobic in nature. A single charged residue Asp513 is found near the tip of the tail loop. The side chain of Asp513 hydrogen bonds to four main-chain nitrogens in the opposing peptide strand from the same tail loop, thus stabilizing the formation of the β-hairpin within the tail loop. The ISAV-NP dimer interface is ~3520 Å<sup>2</sup> in size, of which a 1518 Å<sup>2</sup> area is mediated by the tail loop alone. Deletion of the tail loop (ΔLOOP or Δ504–530) resulted in the formation of NP monomers only (Fig. 1), which suggests an important oligomerization function for the tail loop structure in orthomyxoviruses in general.



**Figure 1. Gel filtration chromatograms for ISAV-NP proteins.** The eluted positions of three protein standards are marked by arrows.  
doi:10.1371/journal.ppat.1003624.g001

### The ISAV-NP RNA binding groove

Analogous to the influenza A virus NP, there is a potential RNA-binding groove between the head and body domains of the ISAV-NP. The groove is positively charged according to its calculated electrostatic potentials, and is covered with a large number of basic amino acid residues, which presumably interact with the negatively charged phosphodiester backbone of bound RNA (Fig. 2D, Fig. S1 – Supporting Information). All of these basic residues, except those from the N-terminal domain, are also found in the influenza A virus NP, although their positions may slightly shift on the primary sequence (Fig. 2E). The aromatic residue F274, analogous to Y148 from the influenza A virus NP [23], is found at one end of the RNA-binding groove, and may mediate nucleotide base stacking.

To test whether ISAV-NP binds RNA in the same manner as the influenza A virus NP, we made two double mutants: K185A/R186A and K296A/R299A. These two pairs of basic residues are from two flexible loops in the potential RNA binding groove that are partially disordered in the ISAV-NP structure. At equivalent positions in the influenza A virus NP, R74/R75 and R174/R175 were found to play critical roles in RNA binding [24]. As expected, mutations of these residues resulted in a dramatic decrease in the RNA binding affinity, as the  $K_d$  changed from 24.5 nM for the wild-type (*wt*) protein to 355 nM and 472 nM for the K185A/R186A and K296A/R299A mutants, respectively (Supporting Information, Fig. S2). This result indicates that NPs of both ISAV and the influenza A virus likely bind RNA with their positively charged groove.

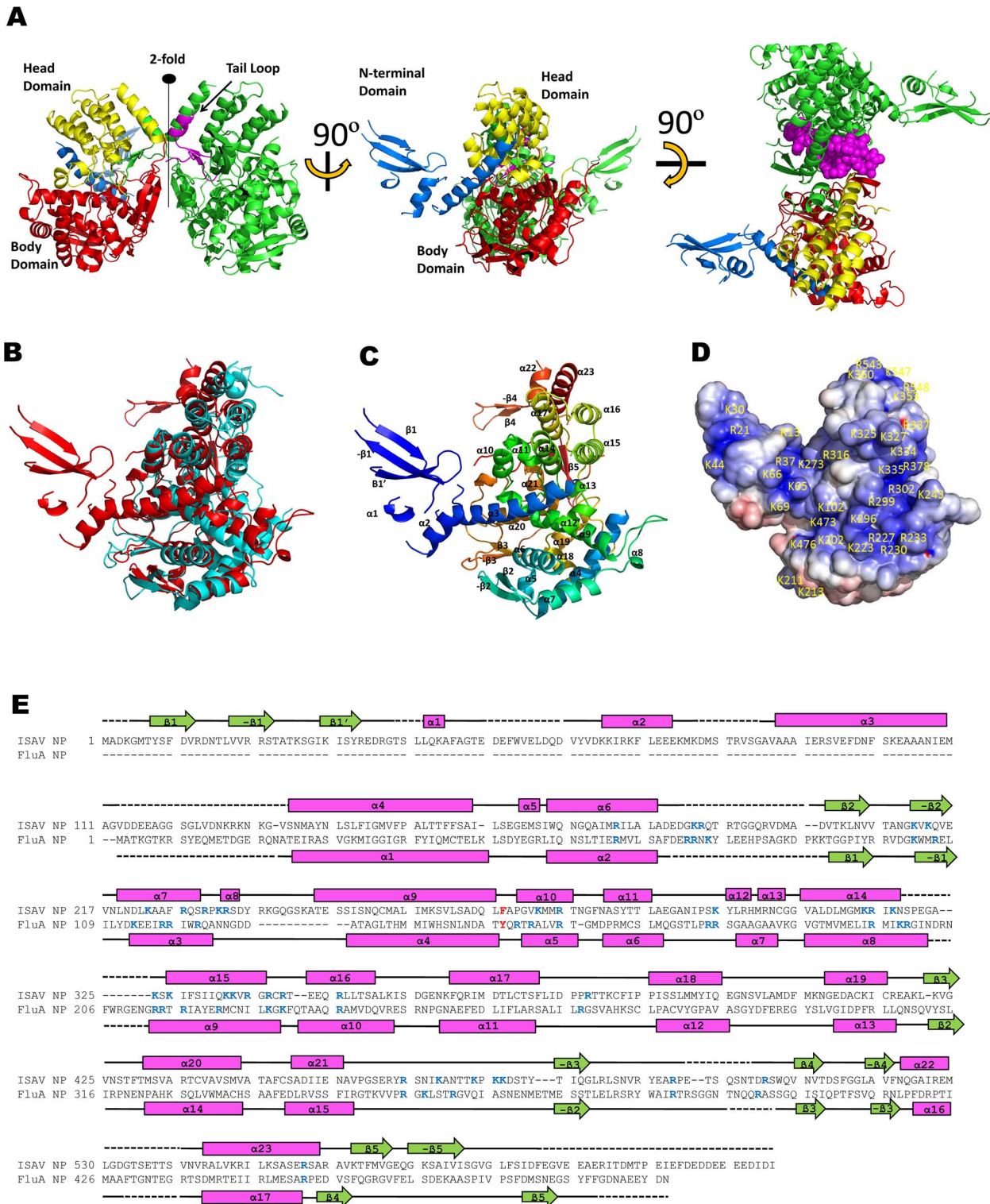
### ISAV-NP N-terminal domain and C-terminal tail

The N-terminal domain of ISAV-NP is a unique structural feature that was not observed in the crystal structures of either the influenza A or influenza B virus NP. Consisting of three anti-parallel  $\beta$ -strands and two  $\alpha$ -helices, the ISAV N-terminal domain adopts an extended shape and interacts with the rest of the structure mainly through the long  $\alpha$ -helix made of residues 86–112. This large  $\alpha$ -helix partially covers the potential RNA binding

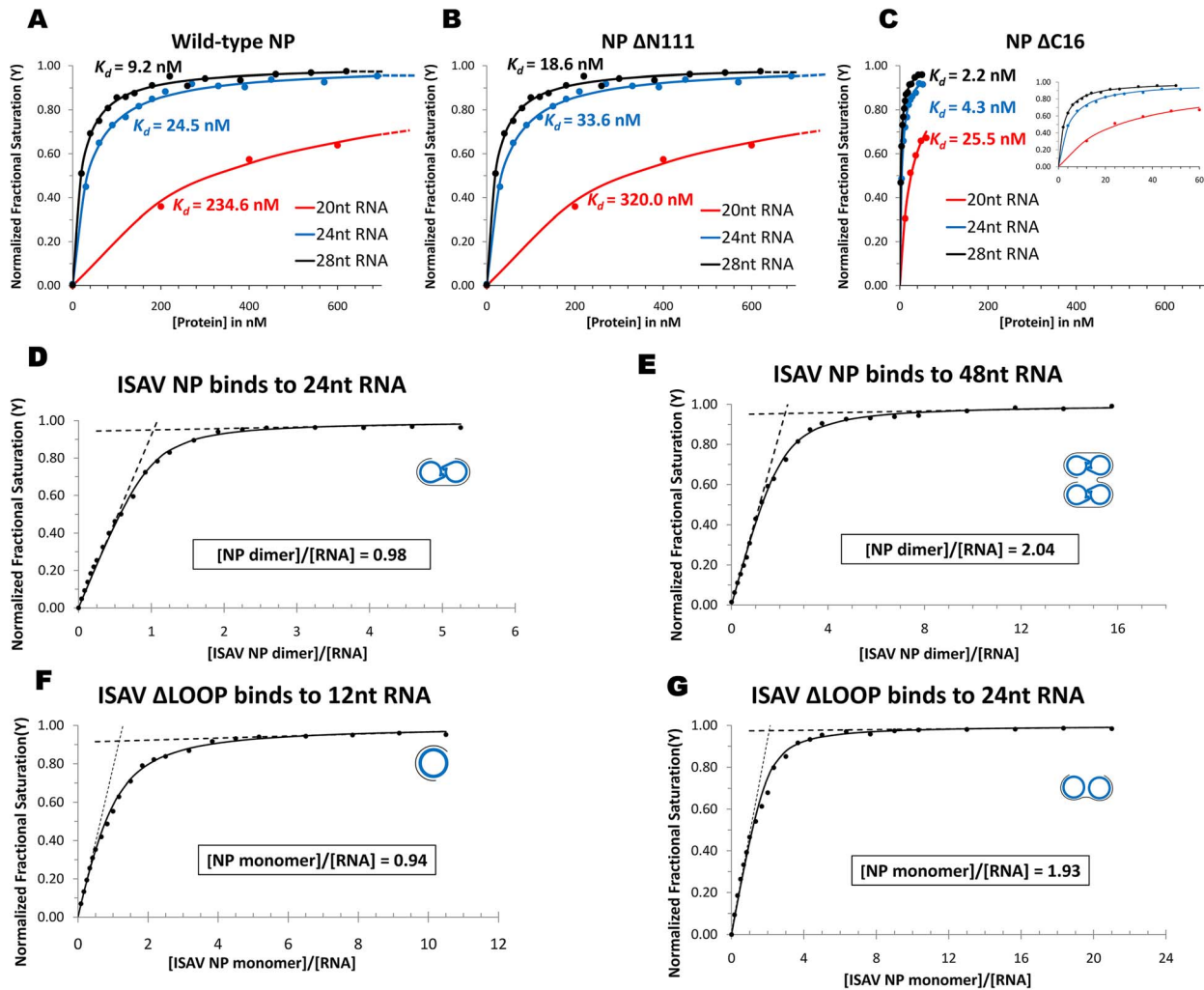
groove in between the head and body domain. Deleting the entire N-terminal domain had little effect on the equilibrium-state RNA binding, as *wt* and  $\Delta$ N111 NP exhibited very similar RNA binding affinities ( $K_d = 24.5$  nM for *wt* vs. 33.6 nM for  $\Delta$ N111, using a 24-nt RNA) (Fig. 3, see the section below for more details). Nevertheless, it is plausible that the N-terminal domain may regulate the dynamics of the NP RNA binding. For instance, it may help to attract RNA substrates and/or serve as a cap to prevent the dissociation of RNA prior to being locked in a bound position. We speculate that the long  $\alpha$ -helix from the N-terminal domain may need to be displaced from its current location upon RNA binding. The inherent structural flexibility of the N-terminal domain is evident from its high temperature factors, averaging  $85.0 \text{ \AA}^2$  for the N-terminal domain, as compared to  $54.8 \text{ \AA}^2$  for the rest of the structure.

A multi-sequence alignment based on secondary structure shows that a N-terminal domain is consistently found in every orthomyxovirus NP, but it is highly variable in both size and sequence (*e.g.* 21aa for influenza A virus, 72aa for influenza B virus, and 111aa for ISAV). Furthermore, the N-terminal domain is structurally disordered in both the influenza A and influenza B virus NPs [23–25]. In influenza A virus, the first 20 amino acid region contains a phosphorylation site and a nuclear localization signal (NLS) that are important for the intracellular trafficking of the NP [28,29]. Such functions have not been attributed to the ISAV N-terminal domain. Although two monopartite nuclear localization signals  $^{230}\text{RPKR}^{233}$  and  $^{473}\text{KPKK}^{476}$  have been identified in ISAV-NP [30], both were mapped to the surface of the body domain. The location of the N-terminal domain suggests that it would be situated on the exterior of assembled RNPs, and therefore may help to mediate interactions with the viral matrix protein and/or NS2 to facilitate the intracellular trafficking of viral RNPs.

At the C-terminal end of the ISAV-NP polypeptide is a highly acidic C-terminal tail that was excluded from our NP construct used for crystal structure determination (see discussions above). The last ordered residue in the crystal structure of ISAV-NP  $\Delta$ C16 (consisting of residues 1 to 600) is residue 586, suggesting that the



**Figure 2. ISAV-NP crystal structure.** (A) Three orthogonal views of an NP dimer. One subunit is colored in green, whereas the other subunit is colored by domains with blue for the N-terminal domain, yellow for the head, red for the body, and magenta for the tail loop. In the last panel, the magenta tail loop is shown by space-filling model to highlight the extensive interaction mediated by the tail loop. (B) ISAV-NP (red) superimposed onto the influenza A virus NP (cyan, PDB ID: 2IQH) [23]. (C) ISAV-NP monomer. The molecule is colored continuously from blue to red for the N-terminus to C-terminus, respectively. Secondary structural elements are numbered. (D) Calculated electronic potential for an NP monomer. Positively charged residues are highlighted. (B–D) are shown in the same orientation by viewing into the putative RNA binding groove, same as the middle panel in (A). (E) ISAV-NP secondary structure assignment.  $\alpha$ -helices are shown by cylinders and the  $\beta$ -strands are represented by arrows. Conserved charged residues from the RNA binding groove are highlighted in blue. The conserved aromatic residue F274 in the groove is shown in red. The NP sequences from ISAV and the influenza A virus were manually aligned, based on tertiary structures. doi:10.1371/journal.ppat.1003624.g002



**Figure 3. ISAV-NP RNA binding.** (A–C) RNA binding affinity measurements for the *wt* NP (A),  $\Delta$ N111 (B) and  $\Delta$ C16 (C). FA was performed using three RNA oligos each containing 20, 24, and 28 nts. The binding curves are plotted using the same protein concentration range. For the  $\Delta$ C16 mutant, an inset is added in (C) to provide a more spread-out view at lower protein concentrations. Fully saturated curves for the 20-nt RNA are shown in Fig. S3 in Supporting Information. (D–G) RNA binding stoichiometry measurements for the *wt* NP and the  $\Delta$ LOOP monomeric mutant. NP:RNA complexes assembled in each experiments are schematically shown with blue handcuffs for NP dimers and grey curves for RNA molecules. doi:10.1371/journal.ppat.1003624.g003

entire polypeptide region after residue 586 is likely to be disordered. We have shown that removing the last 16 residues of the ISAV-NP resulted in an enhanced RNA binding affinity (i.e.  $\sim 6$ -fold increase for a 24-nt RNA) (Fig. 3A, C, see the section below for more details). The exact functional role of the NP C-terminal tail during virus replication remains unclear, and it is possible that the C-terminal tail may assume a different conformation at other functional states. For the influenza A virus NP, the last eight residues are also negatively charged. These last residues of the influenza A virus NP have been found to assume three different structural conformations: (1) structurally disordered [23]; (2) structurally ordered and bound to the putative RNA binding groove [26,31]; and (3) structurally ordered but pointing away from the RNA binding groove [24]. It was proposed that the acidic C-terminal tail of the influenza A virus NP may function to negatively regulate NP RNA binding activity [26,31]. Here, our finding that the ISAV-NP $\Delta$ C16 mutant has a stronger RNA binding affinity appears to support this hypothesis.

### ISAV-NP RNA binding

Using synthetic RNA oligos and fluorescence polarization assays, we first measured the RNA-binding affinity of three NP variants, *wt* NP, NP  $\Delta$ N111, and NP  $\Delta$ C16.  $\Delta$ N111 lacks the N-terminal domain, so its structure is comparable to that of the influenza A virus NP. Because genome-wide stoichiometry estimation revealed that each influenza A virus NP binds to 24 nts of RNA [19], the oligonucleotide (AC)<sub>12</sub> (consisting of 12 AC repeats) was chosen for our initial assays. Interestingly, *wt* NP and  $\Delta$ N111 exhibited similar RNA-binding affinities (Fig. 3A and B), suggesting that the N-terminal domain is not actively involved in equilibrium-state RNA binding. On the other hand, NP  $\Delta$ C16 exhibited a much stronger RNA binding affinity than *wt* NP, indicating the acidic C-terminal tail negatively regulates NP-RNA binding. When the RNA oligo was shortened from 24 nts to 20 nts, the binding affinity of NP dropped sharply with a  $\sim 10$ -fold increase in  $K_d$  (Fig. 3A). On the other hand, when the size of the RNA oligo was increased from 24 nt to 28 nt; there was little change in NP-RNA binding affinity (Fig. 3A). Similar to the full-

length NP, RNA size has little effect on the  $K_d$ 's for binding to the  $\Delta$ N111 and  $\Delta$ C16 mutants, even though the latter shows a  $\sim$ 10-fold higher intrinsic affinity (Fig. 3B and C). Therefore, the 24 nt-long RNA has saturated the NP-RNA binding site, and the small increase in binding affinity from 24-nt to 28-nt RNA was probably due to non-specific binding. The same trend was also observed when DNA oligos ranging from 15 to 30 nts long were used for binding (Supporting Information, Fig. S4).

To determine whether each 24-nt RNA binds to a single NP subunit or to both subunits in a NP dimer, we used the same fluorescence polarization assay to quantify binding stoichiometry. To ensure stoichiometric binding, we adjusted the experimental conditions such that the total concentration of RNA was significantly above the binding dissociation constant  $K_d$  (see Online Methods). When the *wt* NP was titrated in the solution, the fraction of bound RNA, as noted from millipolarization (mP) signal, first increased linearly and then gradually approached saturation (Fig. 3D). Using either the inflection point extrapolation or direct fitting of the general titration equation, we arrived at the same conclusion that each 24-nt RNA binds to an NP dimer. Subsequently, a similar experiment using a 48-nt RNA gave a stoichiometry of two NP dimers per 48-nt long RNA (Fig. 3E), consistent with our earlier result of each NP monomer binding to  $\sim$ 12 nts of RNA.

To rule out the possibility that one NP subunit binds 24 nts of RNA and that RNA binding by one NP subunit allosterically inhibits RNA binding by the other subunit in the NP dimer, a monomeric  $\Delta$ LOOP mutant (Fig. 1) was used for RNA binding assays as described above. Circular dichroism spectra of the  $\Delta$ LOOP mutant were very similar to that of the *wt* NP trimer, indicating similar overall secondary structure contents (Supporting Information, Fig. S5). For the influenza A virus, the NP  $\Delta$ LOOP mutant has a tertiary structure that is nearly identical to the *wt* NP except for the lacking of the tail loop [31]. A monomeric mutant with a single amino acid substitution (*i.e.* R416A) in the tail loop of the influenza A virus NP was found to undergo some structural rearrangements in the C-terminal part of the polypeptide, but the potential RNA binding groove remains largely intact [26].

As expected, the RNA binding affinity of the ISAV-NP  $\Delta$ LOOP monomer was very similar to that of the *wt* NP for a 12-nt long RNA, which supposedly contains only one NP binding motif ( $K_d = 488.4$  nM for *wt* vs. 1076.1 nM for  $\Delta$ LOOP NP). However, the protein binding stoichiometries of the 12-nt and 24-nt RNAs were unexpectedly different. While each 12-nt RNA bound to one  $\Delta$ LOOP NP monomer, each 24-nt RNA recruited two NP  $\Delta$ LOOP monomers when fully saturated (Fig. 3F, G). Therefore, results obtained from the  $\Delta$ LOOP mutant are consistent with our earlier conclusion based on the *wt* NP dimer.

We then used negative-staining electron microscopy (EM) to confirm the RNA-binding stoichiometry by measuring the number of NP molecules associated with each RNA molecule. As shown in Fig. 4, ISAV-NP formed dimers in the absence of RNA, with each NP molecule assuming the shape of a  $4 \times 5$  nm<sup>2</sup> oval (Fig. 4A, B). When the 24-nt RNA was added to NP at the 1:2 molar ratio (*i.e.* one RNA per NP dimer), the same dimeric assemblies were observed (Fig. 4C, D). However, when the 48 nt-RNA were added to ISAV-NP at the 1:4 ratio (*i.e.* one RNA for four NP subunits), we observed mostly tetramers, which were presumably formed by simultaneous binding of two NP dimers to one 48-nt RNA (Fig. 4E, F).

Similar results were also obtained from size measurement using dynamic light scattering (DLS) (Fig. 4G). Assuming that ISAV-NP adopts a roughly globular shape, the molecular weight of the *wt* NP dimer was  $150.9 \pm 13.0$  kD by DLS measurement. The

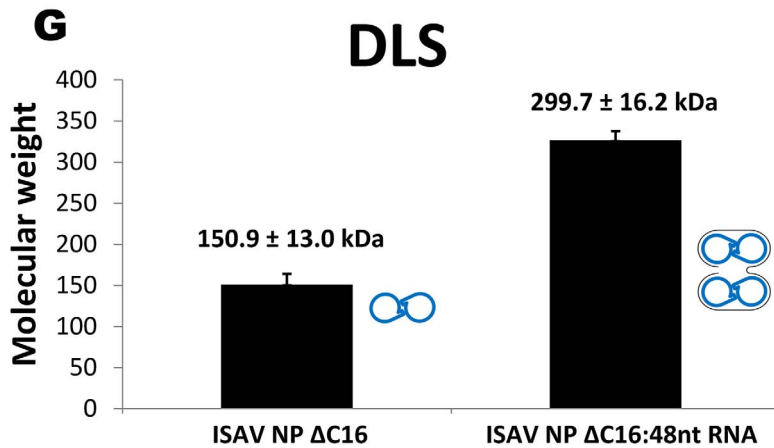
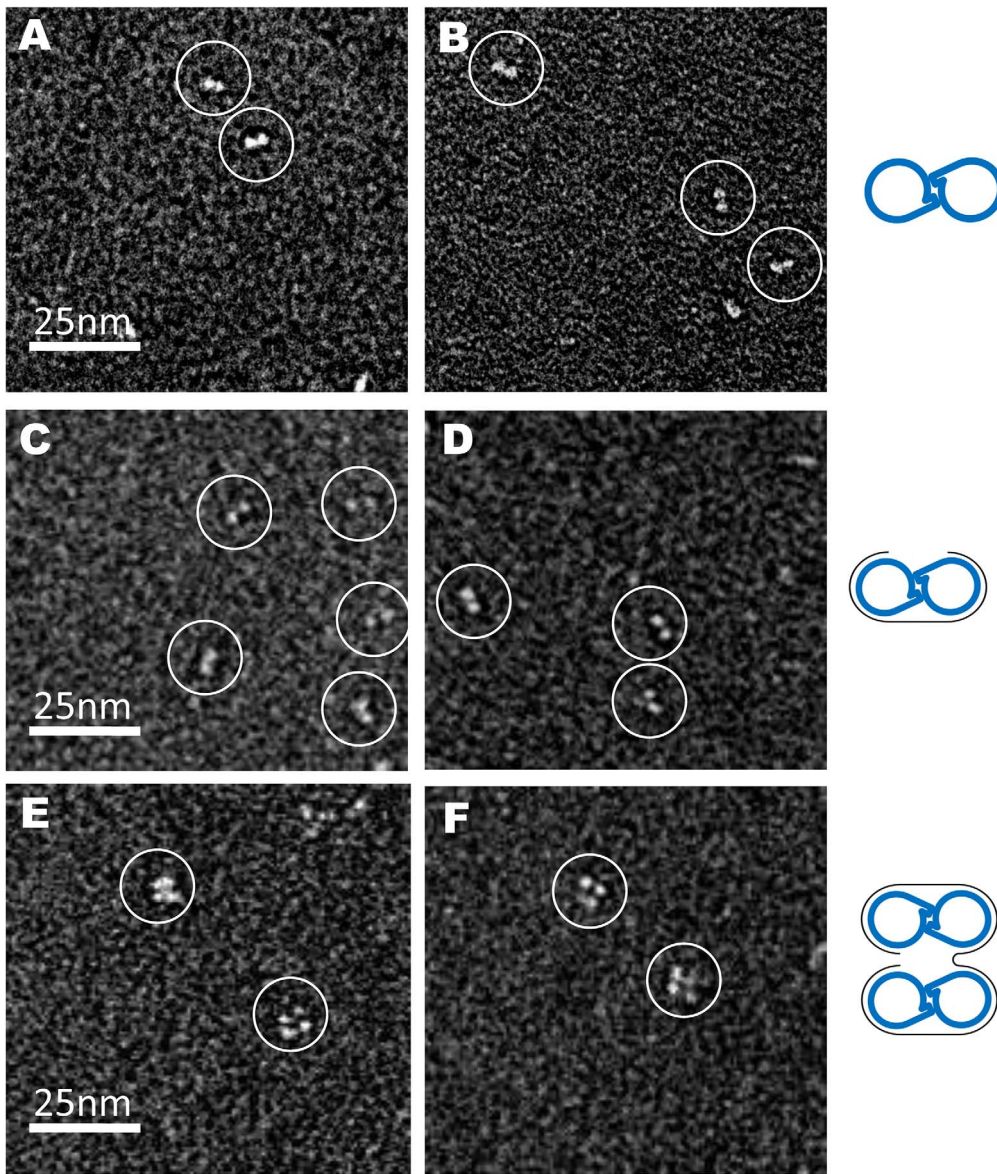
calculated molecular weight of a dimer is 136.5 kD, and the slight difference could be explained by the non-globular shape of the dimer. When the 48-nt RNA was added to the NP sample, the molecular weight of the complex was increased to  $299.7 \pm 16.2$  kD, which matches closely to the calculated molecular weight of two ISAV-NP dimers binding to a 48-nt poly(C) RNA ( $2 \times 136.5 + 14.8 = 287.8$  kD). Our results from EM and DLS once again confirm that one ISAV-NP molecule binds to a 12-nt RNA.

### Functional implication of the orthomyxovirus NP-RNA binding

EM images of ISAV RNPs released from virions showed double-helical structure with similar dimensions compared to those from the influenza A viruses. For instance, it was reported that the longest RNPs of ISAV were  $\sim$ 120 nm long [27], very close to the maximum length of  $\sim$ 110 nm measured for the influenza A virus RNPs [15]. ISAV contains eight gene segments, with the longest one, which encodes PB2, containing  $\sim$ 2.4 kb of nucleotides. This is only slightly larger than the longest RNA segment of the influenza A virus, which is  $\sim$ 2.3 kb in length. Therefore, considering that ISAV and the influenza A virus have similarly sized RNPs as well as RNAs, we expect that the NP RNA binding stoichiometry should be similar for these two viruses. Hence, each ISAV-NP is likely associated with  $\sim$ 24–28 nts of RNA in the infectious virion, as previously determined for the influenza A viruses based on population average.

Since only the head and body domain of ISAV-NP affect equilibrium-state RNA binding, ISAV-NP should serve as a valid model for studying the NP-RNA binding for other orthomyxoviruses including the influenza viruses. We conclude that each ISAV-NP monomer, and possibly each influenza A virus NP, bind  $\sim$ 12 nts of RNA. This stoichiometry of binding is also supported by the observation that the potential RNA-binding groove of the ISAV and the influenza A virus NP is  $\sim$ 60 Å long. Assuming that bound nucleotides are  $\sim$ 5 Å apart (*i.e.* based on single-stranded RNAs binds to nucleocapsids from the tobacco mosaic virus and also rhabdoviruses), each NP molecule should bind no more than  $\sim$ 12 nts of RNA, consistent with our stoichiometric measurements. It has also been noted by Jennings et al. that the length of 22–26 nts exceeds the dimension of NP, so RNA coiling or looping may be necessary for RNP assembly [32]. Indeed, RNase digestion of an influenza A virus RNPs and mini-RNPs produced RNA fragments of  $\sim$ 15–18-nt in size [33,34], indicating that the size of RNA bound by each NP should be shorter than 15–18 nts long.

The cryo-EM reconstruction of influenza A virus RNPs shows that NP molecules follow a regular double-helical arrangement, with the potential RNA binding groove facing outward [17,18]. At the current resolution, RNA densities cannot yet be identified in the EM maps. Compared to NP dimer or longer unconstrained NP polymers, the double-helical stem region of the RNP [17,18] shows a wider separation between adjacent NP molecules along the same NP strand and thus may require short RNA linkers made of up to four extra bases. However, even with such linkers considered, RNA would still follow a highly convoluted path if 24 nts of RNA are modeled for each NP molecule in a RNP [17]. Moreover, if we assume that the double-helical RNP is converted from a linear NP-RNA polymer precursor during assembly, then double-helical RNPs and linear NP-RNA polymers should have similar NP RNA binding stoichiometry. Therefore, we believe that the geometric difference between the helical RNP and the unconstrained NP-RNA polymer cannot completely compensate for the substantial difference in NP RNA stoichiometry (*i.e.*  $\sim$ 24 nts vs.  $\sim$ 12 nts for each NP).



**Figure 4. ISAV-NP:RNA complexes.** (A–F) EM images of the free NP (A, B), NP bound to a 24-nt RNA (C, D), and NP bound to a 48-nt RNA (E, F). (G) Molecular weights of NP and NP:RNA complexes as determined by DLS. Molecular interpretations were made for each figure with blue handcuffs for the NP dimers and grey curves for RNA molecules. doi:10.1371/journal.ppat.1003624.g004

Based on our stoichiometry measurements and the rationale as stated above, we speculate that a fraction of the orthomyxovirus (*e.g.* ISAV, influenza A viruses, etc) genomic RNA likely exists in a NP-free state in the context of RNP. Such NP-free RNA regions are probably rich in secondary or tertiary structures, as it has been shown that double-stranded RNAs bind less favorably to NP compared to single-stranded RNAs. We envision that there is an uneven distribution of NP along the viral RNA that varies depending on the secondary/tertiary structure of the RNA. Because nascent v/cRNA molecules are encapsidated by NP co-transcriptionally, we assume that only local RNA structures occur in RNPs. Although little is known about the details of such RNA structures, presumably they need to be stable enough to resist denaturation by NP binding, as it has been reported that the influenza A virus NP was able to melt the viral RNA panhandle structure [35]. These RNA structures likely play a critical role in the packaging of viral genome during the assembly process, possibly by forming specific higher-order structures. These structures would allow the eight different RNPs to recognize each other and form the “7+1” array previously seen in EM images of budding influenza A virus particles [10–13,36]. Thus, the 2-nm thick filamentous masses connecting neighboring RNPs in budding influenza A viruses are likely RNA structures formed by packaging signals [13]. Because the RNA packaging signals are usually found near both ends of the vRNA molecules [37], each RNP may interact with at least two neighboring RNPs to build up a network of interactions that ensure specific genome packaging.

Crystal structures of NP proteins or NP-RNA complexes have been reported for a number of other negative-sense, single-stranded RNA viruses including members from the rhabdovirus, paramyxovirus, bornavirus, arenavirus and bunyavirus families [38–53]. Among negative strand RNA viruses, rhabdovirus, paramyxovirus, filovirus and bornavirus have a non-segmented genome, whereas arenavirus, bunyavirus and orthomyxovirus have a genome comprised of multiple RNA segments. For non-segmented negative strand RNA viruses, it has been shown by crystal structure and/or genetic evidence that their RNPs are strictly stoichiometric complexes, with NP proteins distributed evenly along the NP-RNA complexes; whereby vRNA is completely unwound and protected by the NP [38–41,54,55]. For rhabdovirus, paramyxovirus, and filovirus, each of their NP proteins binds 9, 6/7, and 6 nucleotides of RNA, respectively [39–41,54,55]. Such properties are consistent with the function of these RNPs, as the genomic sequences of non-segmented negative-sense RNA viruses are not expected to mediate RNP packaging into viral particles. Compared to NPs from non-segmented, negative-sense RNA viruses, orthomyxovirus NPs have a different topology and are relatively large in size. With NP-free RNA exposed on the outer surface, the structural organization of the orthomyxovirus RNP is optimally suited for multi-segment RNA packaging. Future efforts in identifying RNA secondary structure will likely reveal molecular determinants for RNP recognition and genomic RNA reassortment for the influenza viruses and other orthomyxoviruses. It will also be of great interest to identify molecular determinants used by arenaviruses and bunyaviruses to facilitate specific RNP packaging, should such a specific genome package mechanism exist in these segmented RNA viruses as well.

## Materials and Methods

### 1. Cloning

The full-length ISAV-NP (aa 1–616), the  $\Delta$ C16 NP (aa 1–600), and the  $\Delta$ N111 NP (aa 112–616) were cloned with an N-terminal 6×His tag. The polymerase chain reactions were performed using

the *PfuUltra*<sup>TM</sup> high-fidelity DNA polymerase (Stratagene). PCR products were digested with NcoI and XhoI (New England Biolabs) for 2 hours, and ligated into the pET28b+ vector for expression in Rosetta 2 *E. coli* cells (DE3) (Novagen). For the ISAV  $\Delta$ LOOP mutant, a short linker sequence “Gly-Gly-Gly-Gly-Ser” was used to replace the “tail-loop” (aa 504–530).

### 2. Expression, purification, and crystallization

The expression of NP proteins was induced with 0.1 mM IPTG at 15°C for 24 hours after the cell density had reached an OD<sub>600 nm</sub> of 0.8. SeMet-substituted NP proteins were expressed in M9 minimal medium supplemented with seleno-methionine (ACROS organics). Cells were collected and sonicated in lysis buffer containing 300 mM NaCl, 5 mM Imidazole, 10% glycerol (v/v), and 50 mM Tris·HCl, pH 7.5. The lysate was centrifuged at 40,000× *g* for 30 min and the supernatant was used for further purification. Recombinant NP proteins were purified using a HisTrap<sup>TM</sup> HP Column, a HiTrap<sup>TM</sup> Heparin HP Column, and finally a Superdex<sup>TM</sup> 200 Column (GE Healthcare). Purified proteins were at least 95% pure according to Coomassie-stained SDS-polyacrylamide gels. All proteins were concentrated to 10 mg/ml and stored at 4°C.

Crystals of ISAV-NP  $\Delta$ C16 were obtained from a protein-DNA complex containing the oligo dC<sub>18</sub>. Although the DNA oligo could not be seen in the electron density, its presence was required for crystal formation. The best  $\Delta$ C16 crystals were grown at 10°C for two weeks using the sitting-drop vapor diffusion method with the drop made of equal volume of protein solution (10 mg/ml in 200 mM NaCl, 10% glycerol, and 50 mM Tris·HCl, pH 7.5) and well solution (containing 0.8 M lithium chloride, 0.1 M sodium citrate, pH 4.0, and 20% PEG 6000). To improve crystal diffraction, an additional dehydration step was performed by transferring crystals to a mother liquor containing 0.8 M lithium chloride, 0.1 M sodium citrate, pH 4.0, 9% PEG 6000 and 20% PEG 400. After 24 hours, the crystals were directly frozen in liquid nitrogen.

### 3. Data collection, phase determination, model building, and refinement

Diffraction data were collected at the Beamline 4.2.2 at the Advanced Light Source, Berkeley. All diffraction data were processed using the program HKL2000 [56]. The structure was determined by single-wavelength anomalous diffraction (SAD). The heavy atom sites and experimental phases were determined by the AutoSol Wizard in the PHENIX software suite [57]. The protein model was built with Autobuild (PHENIX) and COOT [58], and refined with phenix.refine. All the structure figures were prepared using the program PyMOL unless otherwise specified (The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC). The coordinates have been deposited at the Protein Data Bank (PDB ID 4EWC)

### 4. Fluorescence polarization assays

5'-fluorescein-labeled RNA oligos were purchased from Thermo Fisher Scientific and Sigma-Aldrich. The sequences of the RNA oligos used are simple “AC” repeats (*e.g.*, a 12mer RNA has the sequence of 5'-(AC)<sub>6</sub>-3') to avoid potential secondary structures that may interfere with protein binding.

For the affinity measurements shown in Fig. 3A–C, our binding solutions contained 100 mM NaCl, 25 mM Tris, pH 7.5, 10% glycerol, and fluoresceinated RNA at 0.2 nM or higher (at least 10-fold lower than  $K_d$ ). NPs were titrated into the binding solution until the millipolarization (mP) stabilized. All of the experiments



were carried out at 25°C. The data were plotted and analyzed by using the equation  $P = \{(P_{\text{bound}} - P_{\text{free}})[\text{protein}] / (K_d + [\text{protein}])\} + P_{\text{free}}$ , where  $P$  is the polarization measured at a given total protein concentration,  $P_{\text{free}}$  is the initial polarization of fluorescein-labeled RNA without protein bound,  $P_{\text{bound}}$  is the maximum polarization of RNA when all are bound by protein, and  $[\text{protein}]$  is the protein concentration. The free and total protein concentrations were assumed to be equal since the  $K_d$  was at least 10-fold higher than the concentration of fluorescein-labeled RNA. The hyperbolic curves were fitted by nonlinear least-squares regression analysis assuming a bimolecular model of pseudo first order such that the  $K_d$  values represent the protein concentrations at which half of the ligands are in the bound-state.

For the FA experiments used to determine the binding stoichiometry, the assay conditions were identical to those used for the binding affinity determination, except that the RNA concentration was increased to 60 nM and the NaCl concentration was reduced to 50 mM. These adjustments were made to ensure that the RNA concentration was at least 20-fold higher than the  $K_d$ . The titration curves showed a linear increase (first phase) in millipolarization until RNA-binding sites were saturated (second phase). Data points belonging to each phase were separately fitted by linear least-squares regression analysis to produce the inflection point, at which point the RNAs were stoichiometrically bound with protein.

## 5. Electron microscopy

The ISAV-NP  $\Delta$ C16:48 nt-RNA sample was prepared by mixing the protein with the 48 nt-ssRNA at a 4:1 molar ratio in 100 mM NaCl and 25 mM HEPES, pH 7.5. The  $\Delta$ C16:24 nt-RNA complex was prepared by mixing protein with RNA at a molar ratio of 2:1. Both samples were incubated at 25°C for 30 min and purified by gel filtration before use.

Carbon-coated copper grids (Electron Microscopy Sciences) were glow discharged at 5 mA for 1 min. 3  $\mu$ L of protein samples at 0.02 mg/mL was added onto glow-discharged copper grids, washed twice with distilled water, and stained with freshly prepared 1.5% uranyl acetate solution for 1 min. Images were taken with a JEOL 2010 electron microscope (JEOL USA, Inc.) operated at 100 KeV using a magnification of 40,000 $\times$ .

## 6. Dynamic light scattering

The ISAV-NP  $\Delta$ C16:48 nt-RNA sample was prepared in the same manner as for EM. DLS experiments were performed using a DynaPro99 Dynamic Light Scattering system from Wyatt. All of the samples were centrifuged at 22,000 $\times g$  for 30 min to remove large aggregates prior to the experiments. The DLS experiments

were carried out at 25°C in 100 mM NaCl and 25 mM HEPES, pH 7.5. For each sample, 20 molecular weight readouts were calculated with each being the average of five individual measurements. The 20 readouts were further averaged to obtain the mean and standard deviation of the sample molecular weights.

## Supporting Information

**Figure S1** Calculated electronic potential for an ISAV-NP monomer. Positively charged residues are highlighted. The diagram on the left is the same as the one shown Figure 2D. (DOCX)

**Figure S2** RNA binding by *wt* and mutant ISAV-NP. RNA binding affinity measurements for the *wt* NP and its two double mutants, K185A/R186A and K296A/R299A, were performed by FA using a 24-nt RNA. (DOCX)

**Figure S3** ISAV-NP RNA binding. RNA binding affinity measurements for the *wt* NP,  $\Delta$ N111 and  $\Delta$ C16 were performed by FA using a 20-nt RNA oligo. (DOCX)

**Figure S4** ISAV-NP DNA binding. DNA binding affinity measurements for the *wt* NP were performed by FA using four poly(C) oligos ranging from 15 to 30 nucleotides long. (A) and (B) are plotted against a narrower and broader X-axes, respectively. (DOCX)

**Figure S5** Far-UV CD Spectra of ISAV NP. All three NP proteins, including the *wt* NP (blue),  $\Delta$ LOOP (red) and  $\Delta$ C16 (green) were kept at 5  $\mu$ M concentration. Protein buffer was 200 mM potassium phosphate pH 7.5. Measurements were done at room temperature. (DOCX)

**Table S1** Crystallographic data statistics. (DOCX)

## Acknowledgments

We thank Aaron Collier and Justin Harper for helpful discussions and Jay Nix at the Advanced Light Source for assistance with data collection.

## Author Contributions

Conceived and designed the experiments: YJT WZ. Performed the experiments: WZ YJT. Analyzed the data: WZ YJT JO. Contributed reagents/materials/analysis tools: YJT VV WZ. Wrote the paper: YJT WZ.

## References

1. Palese P, Shaw ML (2007) Orthomyxoviridae: The Viruses and Their Replication. Fields virology. 5th ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins. pp. 1648–1692.
2. Portela A, Digard P (2002) The influenza virus nucleoprotein: a multifunctional RNA-binding protein pivotal to virus replication. J Gen Virol 83: 723–734.
3. Duhaut SD, McCauley JW (1996) Defective RNAs inhibit the assembly of influenza virus genome segments in a segment-specific manner. Virology 216: 326–337.
4. Odagiri T, Tashiro M (1997) Segment-specific noncoding sequences of the influenza virus genome RNA are involved in the specific competition between defective interfering RNA and its progenitor RNA segment at the virion assembly step. J Virol 71: 2138–2145.
5. Fujii K, Fujii Y, Noda T, Muramoto Y, Watanabe T, et al. (2005) Importance of both the coding and the segment-specific noncoding regions of the influenza A virus NS segment for its efficient incorporation into virions. J Virol 79: 3766–3774.
6. Fujii Y, Goto H, Watanabe T, Yoshida T, Kawaoka Y (2003) Selective incorporation of influenza virus RNA segments into virions. Proc Natl Acad Sci U S A 100: 2002–2007.
7. Muramoto Y, Takada A, Fujii K, Noda T, Iwatsuki-Horimoto K, et al. (2006) Hierarchy among viral RNA (vRNA) segments in their role in vRNA incorporation into influenza A virions. J Virol 80: 2318–2325.
8. Watanabe K, Takizawa N, Katoh M, Hoshida K, Kobayashi N, et al. (2001) Inhibition of nuclear export of ribonucleoprotein complexes of influenza virus by leptomycin B. Virus Res 77: 31–42.
9. Liang Y, Hong Y, Parslow TG (2005) cis-Acting packaging signals in the influenza virus PB1, PB2, and PA genomic RNA segments. J Virol 79: 10348–10355.
10. Noda T, Sagara H, Yen A, Takada A, Kida H, et al. (2006) Architecture of ribonucleoprotein complexes in influenza A virus particles. Nature 439: 490–492.
11. Yamaguchi M, Danev R, Nishiyama K, Sugawara K, Nagayama K (2008) Zernike phase contrast electron microscopy of ice-embedded influenza A virus. J Struct Biol 162: 271–276.
12. Fournier E, Moules V, Essere B, Paillart JC, Sirbat JD, et al. (2012) A supramolecular assembly formed by influenza A virus genomic RNA segments. Nucleic Acids Res 40: 2197–2209.

13. Noda T, Sugita Y, Aoyama K, Hirase A, Kawakami E, et al. (2012) Three-dimensional analysis of ribonucleoprotein complexes in influenza A virus. *Nat Commun* 3: 639.
14. Chou YY, Vafabakhsh R, Doganay S, Gao Q, Ha T, et al. (2012) One influenza virus particle packages eight unique viral RNAs as shown by FISH analysis. *Proc Natl Acad Sci U S A* 109: 9101–9106.
15. Compans RW, Content J, Duesberg PH (1972) Structure of the ribonucleoprotein of influenza virus. *J Virol* 10: 795–800.
16. Pons MW, Schulze IT, Hirst GK, Hauser R (1969) Isolation and characterization of the ribonucleoprotein of influenza virus. *Virology* 39: 250–259.
17. Arranz R, Coloma R, Chichon FJ, Conesa JJ, Carrascosa JL, et al. (2012) The structure of native influenza virion ribonucleoproteins. *Science* 338: 1634–1637.
18. Moeller A, Kirchdoerfer RN, Potter CS, Carragher B, Wilson IA (2012) Organization of the influenza virus replication machinery. *Science* 338: 1631–1634.
19. Ortega J, Martin-Benito J, Zurcher T, Valpuesta JM, Carrascosa JL, et al. (2000) Ultrastructural and functional analyses of recombinant influenza virus ribonucleoproteins suggest dimerization of nucleoprotein during virus amplification. *J Virol* 74: 156–163.
20. Martin-Benito J, Area E, Ortega J, Llorca O, Valpuesta JM, et al. (2001) Three-dimensional reconstruction of a recombinant influenza virus ribonucleoprotein particle. *EMBO Rep* 2: 313–317.
21. Cottet L, Rivas-Aravena A, Cortez-San Martin M, Sandino AM, Spencer E (2011) Infectious salmon anemia virus—genetics and pathogenesis. *Virus Res* 155: 10–19.
22. Raynard RS, Murray AG, Gregory A (2001) Infectious salmon anaemia virus in wild fish from Scotland. *Dis Aquat Organ* 46: 93–100.
23. Ye Q, Krug RM, Tao YJ (2006) The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature* 444: 1078–1082.
24. Ng AK, Zhang H, Tan K, Li Z, Liu JH, et al. (2008) Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. *FASEB J* 22: 3638–3647.
25. Ng AK, Lam MK, Zhang H, Liu J, Au SW, et al. (2012) Structural basis for RNA binding and homo-oligomer formation by influenza B virus nucleoprotein. *J Virol* 86: 6758–6767.
26. Chenavas S, Estrozi LF, Slama-Schwok A, Delmas B, Di Primo C, et al. (2013) Monomeric nucleoprotein of influenza A virus. *PLoS Pathog* 9: e1003275.
27. Falk K, Namork E, Rimstad E, Mjaaland S, Dannevig BH (1997) Characterization of infectious salmon anemia virus, an orthomyxo-like virus isolated from Atlantic salmon (*Salmo salar* L.). *J Virol* 71: 9016–9023.
28. Wang P, Palese P, O'Neill RE (1997) The NPI-1/NPI-3 (karyopherin alpha) binding site on the influenza A virus nucleoprotein NP is a nonconventional nuclear localization signal. *J Virol* 71: 1850–1856.
29. Arrese M, Portela A (1996) Serine 3 is critical for phosphorylation at the N-terminal end of the nucleoprotein of influenza virus A/Victoria/3/75. *J Virol* 70: 3385–3391.
30. Aspehaug V, Falk K, Krossoy B, Thevarajan J, Sanders L, et al. (2004) Infectious salmon anemia virus (ISAV) genomic segment 3 encodes the viral nucleoprotein (NP), an RNA-binding protein with two monopartite nuclear localization signals (NLS). *Virus Res* 106: 51–60.
31. Ye Q, Guu TS, Mata DA, Kuo RL, Smith B, et al. (2012) Biochemical and structural evidence in support of a coherent model for the formation of the double-helical influenza A virus ribonucleoprotein. *MBio* 4: e00467–00412.
32. Jennings PA, Finch JT, Winter G, Robertson JS (1983) Does the higher order structure of the influenza virus ribonucleoprotein guide sequence rearrangements in influenza viral RNA? *Cell* 34: 619–627.
33. Coloma R, Valpuesta JM, Arranz R, Carrascosa JL, Ortin J, et al. (2009) The structure of a biologically active influenza virus ribonucleoprotein complex. *PLoS Pathog* 5: e1000491.
34. Yamanaka K, Ishihama A, Nagata K (1990) Reconstitution of influenza virus RNA-nucleoprotein complexes structurally resembling native viral ribonucleoprotein cores. *J Biol Chem* 265: 11151–11155.
35. Baudin F, Bach C, Cusack S, Ruigrok RW (1994) Structure of influenza virus RNP. I. Influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent. *EMBO J* 13: 3158–3165.
36. Harris A, Cardone G, Winkler DC, Heymann JB, Brecher M, et al. (2006) Influenza virus pleiomorphy characterized by cryoelectron tomography. *Proc Natl Acad Sci U S A* 103: 19123–19127.
37. Hutchinson EC, von Kirchbach JC, Gog JR, Digard P (2010) Genome packaging in influenza A virus. *J Gen Virol* 91: 313–328.
38. Rudolph MG, Kraus I, Dickmanns A, Eickmann M, Garten W, et al. (2003) Crystal structure of the borna disease virus nucleoprotein. *Structure* 11: 1219–1226.
39. Albertini AA, Wernimont AK, Muziol T, Ravelli RB, Clapier CR, et al. (2006) Crystal structure of the rabies virus nucleoprotein-RNA complex. *Science* 313: 360–363.
40. Green TJ, Zhang X, Wertz GW, Luo M (2006) Structure of the vesicular stomatitis virus nucleoprotein-RNA complex. *Science* 313: 357–360.
41. Tawar RG, Duquerroy S, Vornrhein C, Varela PF, Damier-Piolle L, et al. (2009) Crystal structure of a nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus. *Science* 326: 1279–1283.
42. Qi X, Lan S, Wang W, Schelde LM, Dong H, et al. (2010) Cap binding and immune evasion revealed by Lassa nucleoprotein structure. *Nature* 468: 779–783.
43. Hastie KM, Kimberlin CR, Zandonatti MA, MacRae JJ, Saphire EO (2011) Structure of the Lassa virus nucleoprotein reveals a dsRNA-specific 3' to 5' exonuclease activity essential for immune suppression. *Proc Natl Acad Sci U S A* 108: 2396–2401.
44. Guo Y, Wang W, Ji W, Deng M, Sun Y, et al. (2012) Crimean-Congo hemorrhagic fever virus nucleoprotein reveals endonuclease activity in bunyaviruses. *Proc Natl Acad Sci U S A* 109: 5046–5051.
45. Carter SD, Surtees R, Walter CT, Ariza A, Bergeron E, et al. (2012) Structure, function, and evolution of the Crimean-Congo hemorrhagic fever virus nucleocapsid protein. *J Virol* 86: 10914–10923.
46. Wang Y, Dutta S, Karlberg H, Devignot S, Weber F, et al. (2012) Structure of Crimean-Congo hemorrhagic fever virus nucleoprotein: superhelical homooligomers and the role of caspase-3 cleavage. *J Virol* 86: 12294–12303.
47. Ferron F, Li Z, Danek EI, Luo D, Wong Y, et al. (2011) The hexamer structure of Rift Valley fever virus nucleoprotein suggests a mechanism for its assembly into ribonucleoprotein complexes. *PLoS Pathog* 7: e1002030.
48. Raymond DD, Piper ME, Gerrard SR, Skiniotis G, Smith JL (2012) Phleboviruses encapsidate their genomes by sequestering RNA bases. *Proc Natl Acad Sci U S A* 109: 19208–19213.
49. Raymond DD, Piper ME, Gerrard SR, Smith JL (2010) Structure of the Rift Valley fever virus nucleocapsid protein reveals another architecture for RNA encapsidation. *Proc Natl Acad Sci U S A* 107: 11769–11774.
50. Li B, Wang Q, Pan X, Fernandez de Castro I, Sun Y, et al. (2013) Bunyamwera virus possesses a distinct nucleocapsid protein to facilitate genome encapsidation. *Proc Natl Acad Sci U S A*.
51. Dong H, Li P, Elliott RM, Dong C (2013) Structure of Schmallenberg orthobunyavirus nucleoprotein suggests a novel mechanism for genome encapsidation. *J Virol* 87:5593–601.
52. Niu F, Shaw N, Wang YE, Jiao L, Ding W, et al. (2013) Structure of the Leanyer orthobunyavirus nucleoprotein-RNA complex reveals unique architecture for RNA encapsidation. *Proc Natl Acad Sci U S A* 110: 9054–9.
53. Reguera J, Malet H, Weber F, Cusack S (2013) Structural basis for encapsidation of genomic RNA by La Crosse Orthobunyavirus nucleoprotein. *Proc Natl Acad Sci U S A* 110: 7246–51.
54. Calain P, Roux L (1993) The rule of six, a basic feature for efficient replication of Sendai virus defective interfering RNA. *J Virol* 67: 4822–4830.
55. Weik M, Enterlein S, Schlenz K, Muhlberger E (2005) The Ebola virus genomic replication promoter is bipartite and follows the rule of six. *J Virol* 79: 10660–10671.
56. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Macromolecular Crystallography, Pt A* 276: 307–326.
57. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66: 213–221.
58. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60: 2126–2132.