Biochemical and Structural Evidence in Support of a Coherent Model for the Formation of the Double-Helical Influenza A Virus Ribonucleoprotein

Qiaozhen Ye, Tom S. Y. Guu, Douglas A. Mata, Rei-Lin Kuo, Bartram Smith, Robert M. Krug, and Yizhi J. Tao

Department of Biochemistry and Cell Biology, Rice University, Houston, Texas, USA, and Department of Molecular Genetics and Microbiology, Institute for Cellular and Molecular Biology, University of Texas at Austin, Austin, Texas, USA

Q.Y. and T.S.Y.G. contributed equally to this article.

ABSTRACT Influenza A virions contain eight ribonucleoproteins (RNPs), each comprised of a negative-strand viral RNA, the viral polymerase, and multiple nucleoproteins (NPs) that coat the viral RNA. NP oligomerization along the viral RNA is mediated largely by a 28-amino-acid tail loop. Influenza viral RNPs, which serve as the templates for viral RNA synthesis in the nuclei of infected cells, are not linear but rather are organized in hairpin-like double-helical structures. Here we present results that strongly support a coherent model for the assembly of the double-helical influenza virus RNP structure. First, we show that NP self-associates much more weakly in the absence of RNA than in its presence, indicating that oligomerization is very limited in the cytoplasm. We also show that once NP has oligomerized, it can dissociate in the absence of bound RNA, but only at a very slow rate, indicating that the NP scaffold remains intact when viral RNA dissociates from NPs to interact with the polymerase during viral RNA synthesis. In addition, we identify a previously unknown NP-NP interface that is likely responsible for organizing the double-helical viral RNP structure. This identification stemmed from our observation that NP lacking the oligomerization tail loop forms monomers and dimers. We determined the crystal structure of this NP dimer, which reveals this new NP-NP interface. Mutation of residues that disrupt this dimer interface does not affect oligomerization of NPs containing the tail loop but does inactivate the ability of NPs containing the tail loop to support viral RNA synthesis in minigenome assays.

IMPORTANCE Influenza virus, the causative agent of human pandemics and annual epidemics, contains eight RNA gene segments. Each RNA segment assumes the form of a rod-shaped, double-helical ribonucleoprotein (RNP) that contains multiple copies of a viral protein, the nucleoprotein (NP), which coats the RNA segment along its entire length. Previous studies showed that NP molecules can polymerize via a structural element called the tail loop, but the RNP assembly process is poorly understood. Here we show that influenza virus RNPs are likely assembled from NP monomers, which polymerize through the tail loop only in the presence of viral RNA. Using X-ray crystallography, we identified an additional way that NP molecules interact with each other. We hypothesize that this new interaction is responsible for organizing linear, single-stranded influenza virus RNPs into double-helical structures. Our results thus provide a coherent model for the assembly of the double-helical influenza virus RNP structure.

INTRODUCTION Influenza A virus, the causative agent of human pandemics and annual epidemics, is a negative-sense, single-stranded RNA virus in the Orthomyxoviridae family. Within the pleomorphic lipid envelope of the virion are eight segments of virion RNA (vRNA). Each vRNA segment interacts with multiple nucleoproteins (NPs) and a heterotrimeric polymerase complex (3P, comprised of PA, PB1, and PB2) to form a viral ribonucleoprotein (vRNP) that functions in transcription, replication, and packaging of the viral genome (1). Approximately 24 nucleotides (nt) of vRNA associate with each NP molecule (2, 3). Electron microscopy showed that influenza virus RNPs are not linear but rather are hairpin-like double-helical structures (2, 4–6). The viral polymerase, which binds to the common 3’- and 5’-terminal sequences of the vRNA segments, is located at the hairpin termini (7). The location of the viral polymerase at the juxtaposed 3’ and 5’ termini was also observed in a reconstituted mini-RNP (8). In contrast to influenza virus RNPs, the RNPs of nonsegmented, negative-strand RNA viruses (e.g., rhabdoviruses) are linear and form single coils (9). The molecular basis for the double-helical structure of influenza virus RNPs has not been determined.

The atomic structures of the RNPs of two influenza A virus strains, H1N1 (influenza A/WSN/33 [WSN]) and H5N1 (influenza A/Hong Kong/483/97 [HK]), both in the form of a trimer,
have been determined to 3.2- and 3.5-Å resolutions, respectively (10, 11). These two structures show that NP-NP interaction is mediated largely by a tail loop consisting of amino acid residues from positions 402 to 429. A highly positively charged groove is found at the exterior of the NP trimer, indicating that RNA is bound at the outer periphery of the viral RNPs (4, 12). This mode of RNA binding is consistent with previous findings that the viral RNAs in influenza virus RNPs are readily digested by RNase and are exposed to the solvent (12) and that polyvinylsulfate (PVS), a negatively charged polymer, is able to completely displace RNA from influenza virus RNPs (13). The putative RNA-binding groove of the influenza NP is lined with a large number of basic residues scattered along the NP primary sequence (10). Mutagenesis analysis identified two polypeptide regions in the groove that are essential for RNA binding, one containing residues R74 and R75 and the other containing residues R174, R175, and R221 (11, 14, 15).

NP that is not associated with RNA is required for viral RNA replication, which occurs in two steps (16). First, a full-length copy of vRNA, termed complementary RNA or cRNA, is made and is then copied to produce vRNA. Whereas viral mRNA synthesis is initiated with capped RNA primers derived from cellular pre-mRNAs, the synthesis of cRNA and vRNA is initiated without a primer. Three amino acids (R204, W207, R208) in the loop at the top of the head domain of NP are required for both its binding to the viral polymerase and its ability to support viral RNA synthesis catalyzed by the viral polymerase (17). Several roles for the NP-polymerase interaction have been proposed, but no role has been definitively established (18–22). NP is deposited along the newly synthesized cRNA and vRNA chains during viral RNA synthesis (23, 24), so that cRNA and vRNA synthesis and RNP assembly are coupled. This coupling leads to the selective binding of NP to cRNA and vRNA but not to viral mRNA or cellular RNAs in infected cells. By itself, NP binds RNA nonspecifically in vitro (12).

It has not been determined how the transition of NP from monomer to oligomer is regulated in infected cells. Several observations suggested that the oligomeric state of NP might be the predominant form of NP. For example, oligomeric ring structures of NP were predominantly present in purified, recombinant NP that was overexpressed in both Escherichia coli and insect cells (10, 11). In addition, when vRNA was removed from vRNPs, a large proportion of the NP molecules still remained associated with each other in an RNP-like structure (25, 26). Because of these findings, Ruigrok and Baudin (25) raised an intriguing issue about RNP assembly: “it will be interesting to know whether NP has only a weak self-association activity in the absence of RNA. Consequently, NP likely exists primarily as monomers in the cytoplasm before it is imported to the nucleus and participates in the assembly of viral RNPs. We conclude that it is probably unnecessary to invoke the action of another protein to prevent NP oligomerization in the cytoplasm.

Assessing the role of NP oligomerization during viral RNA replication. We previously showed that NP-NP interactions are largely mediated by a tail loop consisting of amino acid residues from 402 to 429 (10). Because cRNA and vRNA synthesis is coupled with RNP assembly (23, 24), we considered it likely that NP proteins lacking a functional tail loop would be inactive in viral RNA replication. To determine whether this is the case, we assayed such NP mutant proteins for their ability to support viral RNA synthesis using minigenome assays and a vRNA-sense luciferase reporter, as described in Materials and Methods. As shown in Fig. 2a, a WSN NP mutant that lacks the tail loop (Δ402–429NP), as well as NP proteins that contain the R416A or E339A mutation, which disrupts a critical intersubunit salt bridge, have completely lost the ability to support viral RNA synthesis. These results are consistent with results from earlier mutagenesis analysis (30, 31). In contrast, the ability of NP to interact with polymerase proteins, as assayed in cotransfection assays, was not affected by such mutations (Fig. 2b).

During viral RNA synthesis, a viral template RNA (vRNA or cRNA) has to dissociate from several, if not many, NP molecules in order to interact with the large trimeric polymerase. The NP scaffold in the RNA-free region remained intact only if NP oligomers are kinetically stable. To measure the stability of RNA-free NP oligomers, purified recombinant NP lacking RNA, which are in the form of oligomers (i.e., mostly trimers), were incubated at 37°C. These NP oligomers broke down to monomers in solution very slowly. For the NP of H1N1 WSN virus, gel filtration analysis of purified NP trimers showed that they dissociated to monomers via an approximate first-order reaction with a half-life ($t_{1/2}$) of $\approx 10$ h (Fig. 1b). A large $t_{1/2}$ was also found for the dissociation of the NP trimers of other influenza A virus strains, namely, an H3N2 virus (influenza A/Utah/72 virus), Hong Kong (HK)/H5N1 virus, and an H9N2 virus (influenza A/HK/1074/99). Con-
sequentially, these NP-NP interactions have extremely low dissociation rates, indicating that the NP scaffold in the RNP complex likely remains intact during viral RNA synthesis, when template RNA has to dissociate from some NPs in order to interact with the polymerase. Consistent with this conclusion, RNPs have been found to retain their rod-like structure even after RNA was stripped away (25, 26).

NP without the tail loop forms dimers via a different NP-NP interface. When analyzed by gel filtration, the Δ402–429NP mutant of WSN virus, as expected, no longer forms homo-trimers or larger homo-oligomers (Fig. 2c), as is also the case for the R416A or E339A mutant NP protein (data not shown). Unexpectedly, the NP tail loop mutant not only forms monomers but also forms a species at the ~110-kDa position (Fig. 2c), which corresponds to the molecular mass of an NP dimer containing two ~55-kDa NP monomers. This dimer species was also observed with the tail loop mutant of HK H5N1 NP (Fig. 2c). This formation of homodimers indicates that NP is able to dimerize presumably through a new interface that does not involve the tail loop. In fact, it appears that native NP is also capable of forming dimers. As shown in Fig. 1b, a small population of dimer species was clearly visible in the native NP sample after most of the NP trimers had dissociated into monomers (Fig. 1b).

Crystal structures of the NP dimers. To identify the new dimer interface, we determined the crystal structure of the WSN Δ402–429NP dimer to 2.8-Å resolution (Table 1; Fig. 3). We also attempted to crystallize the Δ402–429NP monomer but without success. In the NP dimer crystal, a dimer is found in each asymmetric unit, and it appears to be the only stable dimer based on the newly defined regions are residues 73 to 91 and 490 to 498, both of which are disordered in the original trimeric H1N1 NP structure.

The C-terminal tail, which contains amino acids 490 to 498, extends from the back of the body domain and winds into the putative RNA-binding groove (Fig. 3c). This C-terminal tail contains four acidic amino acid residues (D491, E494, E495, and D497) and two asparagines (N492 and N498), and therefore

![Figure 1](http://example.com/figure1.png)

**Figure 1** Oligomerization behavior of NP. (a) NP oligomerization at equilibrium. Freshly purified NP was diluted to different concentrations and allowed to equilibrate for 5 days at 4°C before being injected into a Superdex 200 size exclusion column. For NP samples at 16.5 mg/ml (300 μM) and 1 mg/ml (18 μM), the proportions of trimers/monomers were estimated to be 70%/30% and 10%/90%, respectively, by integrating the areas under the two well-separated peaks in the chromatogram. These measurements gave rise to two independently determined $K_d$ values: $10 \times 10^{-3}$ M$^{-1}$ and $7.1 \times 10^{-3}$ M$^{-1}$, assuming monomer $\leftrightarrow$ trimer transition. The average of the two yielded the final $K_d$ of $8.5 \times 10^{-3}$ M$^{-1}$. (b) Time-dependent dissociation behavior of the NP trimer (A/WSN/1933). Purified NP trimers, at a 1-mg/ml or 18 μM concentration, prepared in the storage buffer (50 mM Tris-HCl, pH 7.5, and 200 mM NaCl) were applied to a Superdex 200 size exclusion column to assess their size distribution at different time points (marked with different colors). A measurement taken at 72 h produced a curve that was nearly identical to the one at 40 h and therefore was omitted for clarity.
its electrostatic property is opposite to that of the positively charged surface of the putative RNA-binding groove. The acidic nature of the C-terminal tail is conserved across all members of the Orthomyxoviridae family. In one extreme case, the infectious salmon anemia virus (ISAV), an isavirus that is distantly related to the influenza A virus, has 12 acidic residues in the last 16 residues of its NP sequence (32).

Comparison of the trimeric and dimeric NP interaction interfaces shows that two different sets of residues are used for these two types of interactions (Fig. 4). The tail loop interaction is responsible for NP oligomerization and involves residues in and around the tail pocket at the back of the NP molecule (Fig. 4b). In contrast, dimer interactions are mediated mostly by two polypeptide regions: a helix-turn-helix motif comprised of residues 149 to 167 and the C terminus from residues 482 to the end (Fig. 4a).

Near the 2-fold symmetry axis, the side chains of residues 487 to 489 close to the C-terminal tail are rearranged compared to those of the NP trimer. These movements allow the side chains of F489 from the two subunits to stack against each other. Overall, the intersubunit interface buried in the trimer is more extensive (≈8,331 Å²) than that in the Δ402–429NP dimer (≈2,122 Å²).

The NP dimer interface is required for viral RNA replication. To verify that the dimer interface is biologically relevant, we tried to identify amino acid residues that are critical for NP dimerization but do not affect oligomerization or RNA binding. We first determined which mutation in the NP mutant with the loop deleted eliminated dimer formation. After testing various mutations, we showed that introducing a D491A mutation eliminated dimer formation (Fig. 5a and b). Although D491 is not on the dimer interface, its mutation apparently disturbs the conforma-
tion of the C-terminal tail, eliminating dimer formation. It was also confirmed that the D491A mutation did not have any effect on regular NP oligomerization/trimerization through the tail loop (Fig. 5b). In addition to identifying D491A, we identified two other mutations, F487D and F487R, that were able to affect dimer interaction but not trimer formation (Fig. 5a and data not shown). Of the other dimer mutations that we have made, R152A, D160A, R261A, and F489A result in partial but not complete disruption of NP dimer formation. The Δ491–498 mutation completely disrupted NP dimer formation, but considering the extensive nature of the deletion, functional characterization of this mutant was not further pursued.

The ability of these three single-residue NP dimerization mutants to support viral RNA synthesis was then tested in minigenome assays using both a vRNA-sense luciferase template and a cRNA-sense luciferase template (Fig. 5c and d). The Y487R NP mutant did not show any activity in either assay. Although the Y487D and D491A mutants showed approximately 10 to 20% activity in the vRNA luciferase assay, we did not detect any mRNA synthesis real-time PCR (RT-PCR). These two mutants showed similarly low activities in the cRNA assay, but, as shown, no vRNA-sense RNA was seen with 25 cycles of RT-PCR. Likewise, no band was observed with 30 cycles either. Additionally, we did

| Parameter                  | Value(s) |
|---------------------------|----------|
| Wavelength (Å)            | 0.9179   |
| Range of resolution (Å)   | 30–2.8   |
| Space group               | C2       |
| a = 60.2, b = 155.6, c = 97.8, β = 90.9° |
| Mosaicity (°)             | 0.83     |
| No. of unique reflections | 21,284   |
| Redundancy                | 4        |
| Mean I/σ(I)               | 9.1 (2.5)|
| Rmerge (%)                | 11.5 (43.9)|
| Completeness (%)          | 97.6 (82.6)|

| Refinement statistics     | Value(s) |
|---------------------------|----------|
| Rwork/Rfree (%)           | 24.2/29.8|
| RMS bond length (Å)/angle (°) | 0.010/1.568 |
| Avg resolution of B factors (Å²) | 56.7   |
| Ramachandran plot data    |          |
| Most favored regions (%)  | 83.4     |
| Additional allowed regions (%) | 16.6   |
| Generously allowed regions (%) | 0       |
| Disallowed regions (%)    | 0        |

* Data in parentheses are for the highest-resolution shell.

**FIG 3** Structure of the Δ402–429NP dimer. (a) Δ402–429NP dimer. The molecule is viewed along the noncrystallographic 2-fold symmetry axis. The two subunits are shown in red and green. Arrows point to the potential RNA-binding grooves. Structurally disordered regions in each molecule are noted in the main text. (b) Subunit B of the Δ402–429NP dimer. The polypeptide chain is rainbow colored from blue to red continuously from the N to the C terminus. The four loops (L1 to L4) in/near the potential RNA-binding groove are highlighted. Important residues from loops L1 to L3 are shown as sticks and highlighted in magenta. (c) C-terminal tail of Δ402–429NP. Residues 490 to 498 are shown as sticks, and the rest of the molecule is a surface representation colored according to electrostatic potential, with positive in blue and negative in red. The viewing direction is similar to that in panel b.
We found that the RNA-binding functional states (e.g., monomeric versus oligomeric NP, RNA-binding activity and) adopt different conformations in different interaction, the C-terminal tail of NP may help to regulate NP RNA-oligomers (data not shown). In addition to mediating dimer in-

due 149 to 167 and the C-terminal tail from residue 482 to the
structure shows that two polypeptide regions constitute most of
the dimer interface: the helix-turn-helix motif comprised of resi-
dues 491 to 498 had no effect on NP's ability to form trimers and larger
trimers oligomerize by interacting with each other largely via the
tail loop and contain two structural domains, as previously de-
dscribed (10, 11). The NP dimer lacking the tail loop consists of two
NP subunits that pack against each other side-by-side in an an-
tiparallel fashion through mostly hydrophobic interactions and
cannot further polymerize. The NP monomer, which likely has a
structure similar to that of trimeric NP, is in dynamic equilibrium
with NP oligomers, but these NP-NP interactions are weak in the
absence of any associated RNA.

NP residues that mediate dimer formation are mostly different
from those used for tail loop interactions, suggesting that these
two types of interactions are nonexclusive and can occur simulta-
aneously within a higher-order complex. The NP dimer crystal
structure shows that two polypeptide regions constitute most of
the dimer interface: the helix-turn-helix motif comprised of resi-
dues 149 to 167 and the C-terminal tail from residue 482 to the
tail. The critical role of the C-terminal tail of NP in NP dimer
formation is evident from our mutagenesis data, as the disruption
of the structural conformation of the C-terminal tail alone re-
sulted in NP mutants that were not able to form any dimers. In-
terestingly, residues 490 to 498 are disordered in the original NP
trimmer crystal structure, and as expected, the deletion of residues
491 to 498 had no effect on NP's ability to form trimers and larger
oligomers (data not shown). In addition to mediating dimer in-
teraction, the C-terminal tail of NP may help to regulate NP RNA-
binding activity and adopt different conformations in different
functional states (e.g., monomeric versus oligomeric NP, RNA-
bound versus RNA-free NP). We found that the RNA-binding
activity of the Δ491–498 mutant was enhanced by ~2-fold com-
pared to that of the wild-type NP protein ($K_d$ was equal to 1.7 ±
0.1 nM for the Δ491–498 mutant versus 3.6 ± 0.5 nM for the wild
type).

While the tail loop functions in maintaining interactions be-
tween adjacent NP molecules in influenza virus RNPs, the NP
dimer interaction has the potential to organize the RNP into its
double-helical structure (Fig. 6). In any double-helical structures,
two sets of stabilizing forces are required. The first set is between
adjacent molecules on the same polymer strand, and the second is
between molecules from the two opposing polymer strands of the
double-helical structure. As shown in Fig. 6, our double-helical
RNP model entails two types of NP interactions. One is the tail
loop interaction, which is used to maintain connection between
adjacent NP molecules that point in the same direction and bind
to the same RNA strand. The other is the dimer interaction made
between antiparallel NP molecules that are associated with two
antiparallel RNA strands. Disruption of the NP dimer interface
leads to the unwinding of the RNP, resulting in the dissociation
of the two antiparallel RNA strands, as previously observed under
high-salt and low-salt buffer conditions (7). Because the RNA-
binding groove is located at the exterior of the NP dimer structure,
bound RNAs are likely to be solvent exposed in the influenza virus
RNPs. Since NP forms the internal protein scaffold of RNP based
on our model, the overall structure of RNP is expected to remain
intact after bound RNA is dissociated, consistent with an earlier
observation (25). We speculate that the RNPs of nonsegmented,
negative-strand RNA viruses lack such a dimer interface, explain-
ning why the RNPs of these viruses are linear rather than double-
helical. Interestingly, the RNPs of another group of segmented,
negative-strand RNA viruses, bunyaviruses, also appear to be
double-helical (33), suggesting that a dimer interface similar to
that described here for the influenza virus NP may be shared by
other members of this group of RNA viruses.

We have shown that NP mutants that cannot dimerize also
cannot support viral RNA synthesis, as determined in minig-
encode assays. In these assays, the vRNAs and cRNAs that are
produced are presumably in the form of RNPs, as is the case in
infected cells. These assays measure the steady-state levels of
vRNPs or complementary RNPs (cRNPs) that are produced. Con-
sequently, the absence of newly synthesized vRNPs and cRNPs are
consistent with two possibilities, which are not mutually exclu-
sive: (i) elongation of vRNA and cRNA chains stops at the point at

FIG 4 Δ402–429NP dimer interface. (a) Subunit B of the NP dimer in ribbons (left) and as a surface representation (right). Residues within a 3.6-Å distance
of the other subunit of the dimer are highlighted in red. The viewing direction is from the back of the NP molecule, such that the RNA-binding groove is located
on the other side of the molecule. (b) Subunit C of the NP trimer (A/WSN/33). Surface regions within a 3.6-Å distance of the tail loop of subunit A are shown in
red. Molecules in both panels are shown in the same orientation.
which the normal double-helical region of the RNPs is reached, and the resulting small single-stranded RNPs are unstable, and/or (ii) partial or complete single-stranded RNPs are produced but are inherently unstable or cannot be used as the template for further RNA synthesis. In any case, these results indicate that the ability to form the NP-NP interface needed for the formation of a double-helical structure is biologically important.

**NP trafficking and functions in infected cells.** Our results support the model for NP trafficking and functions in infected cells outlined in Fig. 6b. Newly synthesized NP exists as monomers in the cytoplasm due to the weak NP-NP association in the absence of bound RNA. After importation into the nucleus, NP molecules are added to the growing end of the assembling RNP complex, where monomeric NP polymerizes with RNA-bound NP.
through the tail loop interaction. Fully assembled RNP adopts a double-helical structure, mediated by the dimer interface revealed in the Δ402–429NP dimer structure described here.

During viral RNA transcription and replication, it is necessary for vRNA (and cRNA) to dissociate from at least part of the NP scaffold to access the polymerase active site in order to function as a template. This dissociation does not require the disruption of the NP scaffold because the vRNA is exposed on the outer surface of the RNP. Our results suggest that the structure of RNP is maintained primarily by NP protein-protein contacts (i.e., the tail loop between adjacent NP molecules and the dimer interaction between strands). Because we have shown that NP-NP interaction has a low dissociation rate constant, the NP scaffold is likely to retain intact when the vRNA template locally dissociates from the RNP during RNA synthesis, consistent with a previous observation that the overall structure of RNP is maintained even after vRNA is removed (25). After a vRNA region is copied by the polymerase protein subunits were constructed, coexpressed, and purified as previously described (19). siF21 insect cells were grown in Hink's TNM-FH insect medium (JRH Biosciences) supplemented with 10% (vol/vol) fetal bovine serum (Valley Biomedical) and harvested 48 h postinfection.

Purification of NP, overexpressed in either E. coli or insect cells, was performed as previously described (10). Briefly, clarified cell lysate was initially subjected to a Ni-nitrilotriacetic acid (NTA) chromatography step or to an ammonium sulfate precipitation step (for nontagged proteins). Fractions containing NP were subsequently purified by heparin affinity and gel filtration chromatography. The final yield was about 1.5 mg of NP dimer and 5 mg of NP monomer per liter of cells. The purified protein (−95% pure) was concentrated to 5 mg/ml using a Centricron tube (Millipore) and stored at −80°C.

**Crystallization and structure determination.** Crystallization conditions were initially screened using a Hydra II Plus One microdispenser robot (Apoget Discoveries) with ready-to-use crystal screen kits (Hampton Research, Qiagen). Rectangular crystals (−10 by 20 by 80 μm) of the Δ402–429NP dimer were grown at room temperature by vapor diffusion with hanging drops with a ratio of protein to well solution of 1:1. The well solution contained 100 mM Tris-Cl, pH 8.5, 10% polyethylene glycol 3350 (PEG 3350), 10% glycerol, and 10 mM dithiothreitol (DTT). The Δ402–429NP dimer crystallized as small rectangles. Crystals were transferred to stabilizing solution (100 mM Tris-Cl, pH 8.5, 20% PEG 3500, 10% glycerol, 20% PEG 400) and flash frozen in liquid nitrogen. Data were collected at the Cornell High Energy Synchrotron Source (CHESS) with beamline F1. All data were processed with HKL-2000 (Table 1) (34).

The structure of the Δ402–429NP dimer was solved by molecular replacement to 2.8-Å resolution using Phaser (35) and CCP4 (36). The A/WSN/33 (H1N1) NP trimer structure (10) was used as the initial phasing model, and it was found that there were two molecules per crystallographic asymmetric unit. The initial map was built using program O (37). To reduce model bias, composite omit maps and density modification were used, and the model was subjected to multiple cycles of simulated annealing with the Crystallography & NMR System (CNS) (38). The two nontagged proteins) or to an ammonium sulfate precipitation step (for His-tagged proteins) were coexpressed with untagged PB1 and PB2, respectively.

For insect cell expression, constructs of untagged and C-terminally His-tagged NPs and their mutants were cloned into the baculovirus vector pFastBac and expressed in SF21 insect cells using the Bac-to-Bac baculovirus expression system (Invitrogen). Baculoviruses expressing the three polymerase protein subunits were constructed, coexpressed, and purified as previously described (19). SiF21 insect cells were grown in Hink's TNM-FH insect medium (JRH Biosciences) supplemented with 10% (vol/vol) fetal bovine serum (Valley Biomedical) and harvested 48 h postinfection.

MATERIALS AND METHODS

**Plasmid constructs and protein purification.** NP genes from A/WSN/1933 (H1N1) and A/HK/483/97 (H5N1) were cloned into pET28a(+) vectors and expressed in *E. coli* Rosetta 2(DE3) Singles competent cells (Novagen) (10). A C-terminal 6×His tag was engineered to facilitate purification by affinity chromatography. Nontagged NPs from both the H1N1 and H5N1 viruses were also expressed for control purposes. NP mutants were constructed using standard oligonucleotide mutagenesis methods and were expressed using the same system. *E. coli* cells were grown in Luria broth medium at 37°C to an optical density at 590 nm of 0.6 and induced by 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 6 h at 28°C.

For insect cell expression, constructs of untagged and C-terminally His-tagged NPs and their mutants were cloned into the baculovirus vector pFastBac and expressed in SF21 insect cells using the Bac-to-Bac baculovirus expression system (Invitrogen). Baculoviruses expressing the three polymerase protein subunits were constructed, coexpressed, and purified as previously described (19). SiF21 insect cells were grown in Hink's TNM-FH insect medium (JRH Biosciences) supplemented with 10% (vol/vol) fetal bovine serum (Valley Biomedical) and harvested 48 h postinfection.

Purification of NP, overexpressed in either *E. coli* or insect cells, was performed as previously described (10). Briefly, clarified cell lysate was initially subjected to a Ni-nitrilotriacetic acid (NTA) chromatography step or to an ammonium sulfate precipitation step (for nontagged proteins). Fractions containing NP were subsequently purified by heparin affinity and gel filtration chromatography. The final yield was about 1.5 mg of NP dimer and 5 mg of NP monomer per liter of cells. The purified protein (−95% pure) was concentrated to 5 mg/ml using a Centricron tube (Millipore) and stored at −80°C.
teins were analyzed by 9% SDS-PAGE and Western blotting. Anti-PB1, anti-PB2, and anti-tetra-His (Qiagen) were used as primary antibodies, depending on which proteins were coexpressed. Goat anti-mouse IgG and anti-rabbit IgG conjugated with alkaline phosphate (Pierce) were used as secondary antibodies. The blot was developed with SigmaFAST 5-bromo-4-chloro-3-indolylphosphate (BCIP)-nitroblue tetrazolium (NBT).

Minigenome assays. Dual-luciferase reporter assays for viral RNA synthesis were carried out as previously described (17). The vRNA template was provided by a pH212 plasmid expressing an RNA containing the 26 5’- and 3’-terminal bases of the U6 NS vRNA sandwiching a negative-sense firefly luciferase reading frame. Alternatively, a cRNA template was provided by a pH212 plasmid expressing the 5’- and 3’-terminal bases of U6 NS cRNA sandwiching a positive-sense firefly luciferase reading frame. To monitor transfection efficiencies, a polymerase II-driven plasmid expressing Renilla luciferase was also transfected. The firefly/Renilla luciferase ratio measures the level of viral RNA synthesis. The amount of viral RNA produced (cRNA or vRNA) in these assays was also directly measured by semiquantitative RT-PCR.

Protein structure accession number. Crystal structure coordinates have been deposited in the Protein Data Bank (PDB; Identifier, 3MIR). All ribbon diagrams were made using the program PyMOL (W. L. Delano; http://www.pymol.org).

ACKNOWLEDGMENTS

We thank Ying Francis Liu, Kevin Mackenzie, John Olson, and Hongli Zhan for helpful discussions, and we thank staff at CHESS (Cornell High Energy Synchrotron Source) for X-ray diffraction data collection.

This work was supported by the NIH (grant AI077785 to Y.J.T. and R.M.K.), the Welch Foundation (grant C-1565 to Y.J.T.), the Hamill Foundation (grant to Y.J.T.), and the Kresge Science Initiative Endowment Fund.

REFERENCES

1. Lamb RA, Krug RM. 2001. Orthomyxoviridae: the viruses and their replication, p 1487–1531. In Knipe DM, Howley PM (ed), Fields virology, vol 1, 4th ed. Lippincott Raven, Philadelphia, PA.

2. Compan RW, Content J, Duesberg PH. 1972. Structure of the ribonucleoprotein of influenza virus. J. Virol. 10:795–800.

3. Ortega J, et al. 2000. Ultrastructural and functional analyses of recombinant influenza virus ribonucleoproteins suggest dimerization of nucleoprotein during virus amplification. J. Virol. 74:156–163.

4. Pons MW, Schulze IT, Hirst GK, Hauser R. 1969. Isolation and characterization of the ribonucleoprotein of influenza virus. Virology 39:250–259.

5. Heggens MH, Smith PR, Ulmanen I, Krug RM, Choppin PW. 1982. Studies on the helical nucleocapsid of influenza virus. Virology 118:466–470.

6. 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.

7. Klump K, Ruigrok RW, Baudin F. 1997. Roles of the influenza virus polymerase and nucleoprotein in forming a functional RNP structure. EMBO J. 16:1248–1257.

8. Martin-Benito J, et al. 2001. Three-dimensional reconstruction of a recombinant influenza virus ribonucleoprotein particle. EMBO Rep. 2:313–317.

9. Schoen G, et al. 2004. The 12 A structure of trypsin-treated measles virus N-RNA. J. Mol. Biol. 339:301–312.

10. Ye Q, Krug RM, Tao YJ. 2006. The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. Nature 444:1078–1082.

11. Ng AK, et al. 2008. Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. FASEB J. 22:3638–3647.

12. 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.

13. Goldstein EA, Pons MW. 1970. The effect of polyvinylsulfate on the ribonucleoprotein of influenza virus. Virology 41:382–384.

14. Albo C, Valencia A, Portela A. 1995. Identification of an RNA binding region within the N-terminal third of the influenza A virus nucleoprotein. J. Virol. 69:3799–3806.

15. Kobayashi M, Toyoda T, Adyshiev DM, Azuma Y, Ishihama A. 1994. Molecular dissection of influenza virus nucleoprotein: deletion mapping of the RNA binding domain. J. Virol. 68:8433–8436.

16. Palese P, Shaw ML. 2007. Orthomyxoviridae: the viruses and their replication, p 1647–1689. In Knipe DM, Howley PM (ed), Fields virology, vol 1, 5th ed. Lippincott Raven, Philadelphia, PA.

17. Marklund JK, Ye Q, Dong J, Tao YJ, Krug RM. 2012. Sequence in the influenza A virus nucleocapsid protein required for viral polymerase binding and RNA synthesis. J. Virol. 86:7292–7297.

18. Kagawuchi A, Momose F, Nagata K. 2011. Replication-coupled and host factor-mediated encapsidation of the influenza virus genome by viral nucleoprotein. J. Virol. 85:6197–6204.

19. Newcomb LL, et al. 2009. Interaction of the influenza A virus nucleocapsid protein with the viral RNA polymerase potentiates unprimed viral RNA replication. J. Virol. 83:29–36.

20. Portela A, Digard P. 2002. The influenza virus nucleoprotein: a multi-functional RNA-binding protein pivotal to virus replication. J. Gen. Virol. 83(Part 4):723–734.

21. Resa-Infante P, Jorba N, Coloma R, Ortin J. 2011. The influenza virus RNA synthesis machine: advances in its structure and function. RNA Biol. 8:202–215.

22. Vreede FT, Jung TE, Brownlee GG. 2004. Model suggesting that replication of influenza virus is regulated by stabilization of replicative intermediates. J. Virol. 78:9568–9572.

23. Beaton AR, Krug RM. 1986. Transcription antitermination during influenza virus template RNA synthesis requires the nucleocapsid protein and the absence of a 5′ capped end. Proc. Natl. Acad. Sci. U. S. A. 83:6283–6286.

24. Shapiro GI, Krug RM. 1988. Influenza virus RNA replication in vitro: synthesis of viral template RNAs and virion RNAs in the absence of an added primer. J. Virol. 62:2285–2290.

25. Ruigrok RW, Baudin F. 1995. Structure of influenza virus ribonucleocapsid particles. II. Purified RNA-free influenza virus ribonucleocapsid forms structures that are indistinguishable from the intact influenza virus ribonucleocapsid particles. J. Gen. Virol. 76:1009–1014.

26. Kingsbury DW, Webster RG. 1969. Some properties of influenza virus nucleocapsids. J. Virol. 4:219–225.

27. Green TJ, et al. 2000. Study of the assembly of vesicular stomatitis virus N protein: role of the P protein. J. Virol. 74:9515–9524.

28. Kingsbury DW, Jones IM, Murti KG. 1987. Assembly of influenza ribonucleocapsid in vitro using recombinant nucleoprotein. Virology 156:496–503.

29. Balz AG, et al. 2012. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol. Cell 46:674–690.

30. Elton D, Medcalf L, Bishop K, Harrison D, Digard P. 1999. Identification of amino acid residues of influenza virus nucleoprotein essential for RNA binding. J. Virol. 73:7357–7367.

31. Menai I, et al. 1999. Mutational analysis of influenza A virus nucleoprotein: identification of mutations that affect RNA replication. J. Virol. 73:1186–1194.

32. Clouthier SC, Rector T, Brown NE, Anderson ED. 2002. Genomic organization of infectious salmon anemia virus. J. Gen. Virol. 83(Part 2):421–428.

33. Pettersson RF, von Bonsdorff CH. 1975. Ribonucleoproteins of Uukuniemi virus are circular. J. Virol. 13:386–392.

34. Otwowski Z, Minor W (ed). 1997. Processing of X-ray diffraction data. CCP4 suite: programs for protein crystallography. Acta Crystallogr. D 53(Part 2):905–924.

35. Storoni LC, McCoy AJ, Read RJ. 2004. Likelihood-enhanced fast rotation functions. Acta Crystallogr. D 60(Part 3):432–438.

36. CCP4 (Collaborative Computational Project Number 4). 1994. The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D 50(Part 5):760–763.

37. Jones TA, Zou JY, Cowan SW, Kjeldgaard M. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A. Found. Crystallogr. 47:110–119.

38. Brunger AT, et al. 1998. Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. 54:905–921.