LETTERS

The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA

Qiaozhen Ye[1], Robert M. Krug[2] & Yizhi Jane Tao[1]

Influenza A viruses pose a serious threat to world public health, particularly the currently circulating avian H5N1 viruses. The influenza viral nucleoprotein forms the protein scaffold of the helical genomic ribonucleoprotein complexes, and has a critical role in viral RNA replication[1]. Here we report a 3.2 Å crystal structure of this nucleoprotein, the overall shape of which resembles a crescent with a head and a body domain, with a protein fold different compared with that of the rhabdovirus nucleoprotein[2,3]. Oligomerization of the influenza virus nucleoprotein is mediated by a flexible tail loop that is inserted inside a neighbouring molecule. This flexibility in the tail loop enables the nucleoprotein to form loose polymers as well as rigid helices, both of which are important for nucleoprotein functions. Single residue mutations in the tail loop result in the complete loss of nucleoprotein oligomerization. An RNA-binding groove, which is found between the head and body domains at the exterior of the nucleoprotein oligomer, is lined with highly conserved basic residues widely distributed in the primary sequence. The nucleoprotein structure shows that only one of two proposed nuclear localization signals are accessible, and suggests that the body domain of nucleoprotein contains the binding site for the viral polymerase. Our results identify the tail loop binding pocket as a potential target for antiviral development.

We expressed the nucleoproteins of several different influenza A virus strains in *Escherichia coli* and insect cells. These recombinant nucleoproteins form non-homogeneous assemblies ranging from monomers to large insoluble aggregates. On the basis of results from gel filtration chromatography, the most abundant molecular species are small oligomers of two to three hundred kilodaltons (Fig. 1). The degree of nucleoprotein oligomerization also varies between influenza A virus strains (Fig. 1). When the nucleoproteins collected from the monomeric peak position were reapplied to the same gel filtration column, most of the nucleoprotein re-appeared as oligomers. Therefore, monomeric and oligomeric forms are interchangeable, and different nucleoprotein assembly states are in dynamic equilibrium. Electron microscopy of the influenza A/WSN/33 virus nucleoprotein taken from the oligomeric peak fraction of the gel filtration (~150 kDa) shows many closed ring structures (Fig. 2a). Although trimers are most common, larger oligomers and monomers are also observed. When viewed along the symmetry axis of an oligomer, each nucleoprotein molecule has the shape of an ellipsoid, with a small lateral surface area buried between neighbouring nucleoprotein subunits.

Although we were unable to isolate nucleoprotein in a single molecular form, nucleoprotein samples collected from the oligomeric peak fractions of a gel filtration column produced thin-plate crystals that diffract to 3.0 Å at synchrotron radiation sources. These thin-plate crystals belong to the space group $C222_1$ with one complete trimer per asymmetric unit. The three-fold non-crystallographic symmetry (NCS) does not fall on any of the major crystallographic symmetry axes. Using multiple isomorphous replacement and anomalous scattering (MIRAS), we determined the crystal structure of influenza A virus nucleoprotein (strain A/WSN/33) to 3.2 Å resolution (Supplementary Table 1 and Supplementary Fig. 1). Of the three subunits in an NCS trimer, subunits A and B each contain 426 amino acid residues in the final model, including residues 21–72, 92–202, 213–396, 402–428 and 438–489. Subunit C contains ten more residues, as residues 203–212 are ordered in this subunit. There are six cysteine residues in each nucleoprotein monomer, but none is close enough for disulphide bond formation. The structure of nucleoprotein is mostly α-helical and each monomer has a curved, crescent-like shape (Fig. 3; see also Supplementary Fig. 2).The overall structure can be divided into two domains: a head and a body, which are joined together by the polypeptide chain at three regions. The head and body domains are formed by non-contiguous polypeptide regions: the head domain is formed by residues 150–272 and 438–452, and the body domain consists of the three polypeptide segments 21–149, 273–396 and 453–489. Consequently, the overall fold and topology of the influenza virus nucleoprotein are different from that of the rhabdovirus nucleoproteins in which the two domains are formed by collinear polypeptide sequences[2,3].

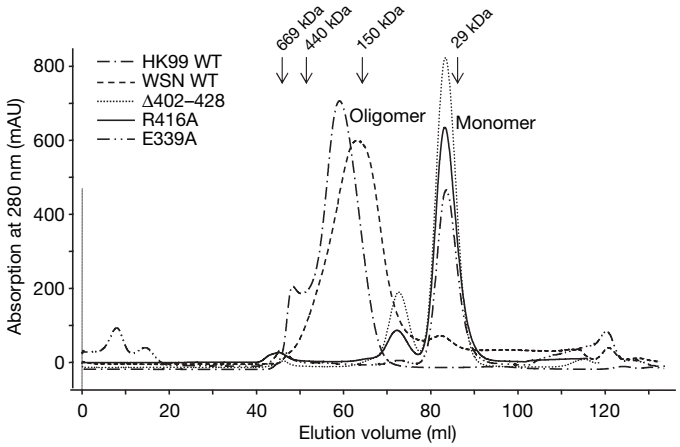

Figure 1 | Nucleoprotein oligomerization behaviour shown by chromatogram. Wild-type nucleoprotein from A/Hong Kong/1074/99 (H9N2) (HK99 WT), wild-type nucleoprotein from A/WSN/33 (WSN WT), and Δ402–428, R416A and E339A mutant nucleoprotein from A/WSN/33 (H1N1) were eluted from a Superdex 200 size-exclusion column. The peak positions of protein standards are marked by arrows. Nucleoprotein monomer and oligomer peaks are also marked.

[1]Department of Biochemistry and Cell Biology, Rice University, 6100 Main Street, MS140, Houston, Texas 77005, USA. [2]Institute for Cellular and Molecular Biology, University of Texas at Austin, 2500 Speedway, Austin, Texas 78712-1095, USA.

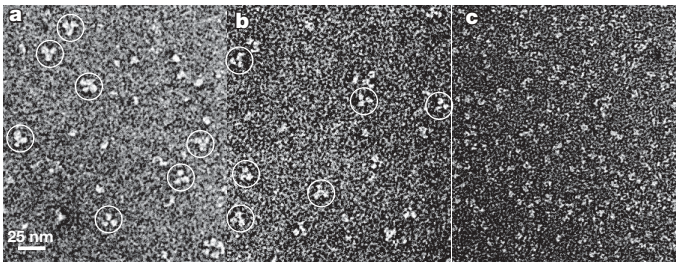

Figure 2 | Electron microscopy images of influenza virus nucleoprotein (strain A/WSN/33). a, Recombinant nucleoprotein. Nucleoprotein trimers are highlighted by white circles. A small number of nucleoprotein monomers and larger oligomers are also observed. **b**, Nucleoprotein in a complex with a 24-nucleotide ssRNA. The nucleoprotein:RNA complex was prepared by mixing nucleoprotein and RNA at the ~1:2 ratio. **c**, The R416A mutant nucleoprotein. All samples were stained using uranyl formate as described in Methods. Scale bar, 25 nm.

Each nucleoprotein contains a tail loop, formed by residues 402–428, at the back of the crescent-shaped molecule (Fig. 3d). The polypeptide connecting the tail loop to the main body of the nucleoprotein structure is disordered. The tail loop has a rather extended conformation and interacts with a neighbouring molecule in an NCS trimer. This tail loop is inserted into the body domain of a neighbouring molecule in an anticlockwise direction when viewed along the three-fold axis from the side of the head domain. Intra-molecular interactions for the tail loop are not possible because the distance is too far for the tail loop to reach the binding pocket located on the opposite side of the same molecule. The tail loop contains three β-strands, with the last two forming an intra-molecular β-hairpin. The tail loop from one molecule makes extensive interactions with a neighbouring subunit through several types of molecular forces, including intermolecular β-sheets (for example, β7 with β8; β6 with β4), hydrophobic interactions (for example, Pro 453(A), Phe 420(B) and Ile 453(A)), and salt bridges (for example, Glu 339(A) with Arg 416(B); Arg 422(B) with Glu 449(A)) (Fig. 4). (The last capitalized letter for amino acids indicates the molecular origin of that residue; for example, (A) indicates subunit A.) The interactions mediated by the tail loop are almost equivalent among the three NCS-related subunits (Supplementary Table 2). In addition, because the oligomerization tail loop is loosely attached to the rest of the nucleoprotein through highly flexible linkers that are disordered in electron density maps, the tail loop can make similar inter-subunit interactions in both closed rings (as observed by electron microscopy in Fig. 2a) and helical ribonucleoprotein (RNP)-like structures, as has been proposed for rhabdovirus nucleocapsid proteins[2,3]. The carboxy-terminal region of each monomer (residues 440–end) also makes some inter-subunit contacts, but these are far fewer in number compared to the tail loop (Supplementary Table 2). In addition, many such interactions among the three subunits are not conserved and therefore also suggest that the molecular interface mediated by the C-terminal region is labile and not critical for nucleoprotein oligomerization.

To determine whether the tail loop is essential for nucleoprotein oligomerization, we made two single residue mutants, R416A and E339A, and a deletion mutant, Δ402–428, that removes the tail loop.

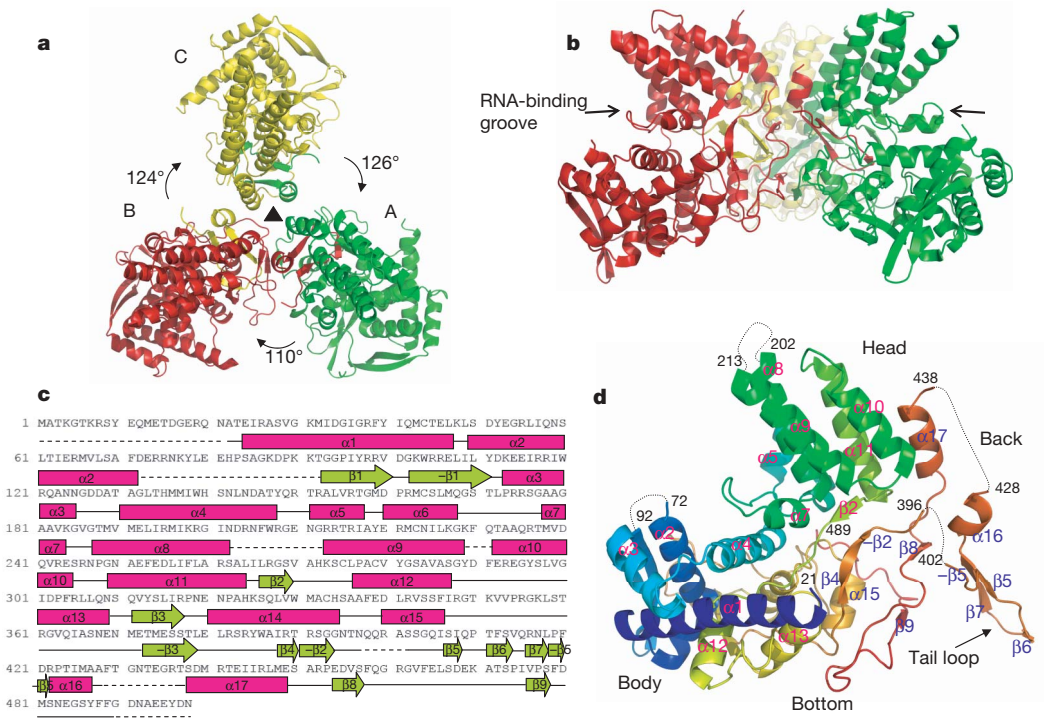

Figure 3 | Nucleoprotein crystal structure. a, Nucleoprotein trimer viewed along the NCS three-fold axis, with three subunits shown in different colours. The rotation angles that relate the three subunits are marked. **b**, Side view of a nucleoprotein trimer, with the three subunits coloured as in **a**. **c**, Nucleoprotein secondary structure assignment (subunit B). α-Helices and β-strands are shown by rods and arrows, respectively. β-Strands with opposite signs form antiparallel β-strands. **d**, Nucleoprotein subunit B. The polypeptide is coloured continuously from blue to red. Residues preceding and succeeding disordered regions are marked and connected by dotted lines.

Residues R416 and E339 form an inter-subunit salt bridge. When the oligomerization behaviour of R416A, E339A and Δ402–428 was evaluated by gel filtration chromatography, all three mutants did not self-associate and formed only monomers (Fig. 1). Electron microscopy of the R416A mutant protein confirmed that only monomers were formed (Fig. 2c), in contrast to wild-type nucleoprotein where trimers and larger oligomers were frequently observed (Fig. 2a). Proper self-association is important for the biological activities of nucleoprotein, because the R416A mutant was shown to be defective in RNA binding and also in viral RNA synthesis[4]. This observation may indicate that the displacement of the tail loop from its binding pocket causes significant structural rearrangements in nucleoprotein. We anticipate that chemical compounds which competitively displace the tail loop from its binding pocket would interfere with viral genome replication, and therefore serve as promising leads for anti-influenza drug development.

The nucleoprotein crystal structure shows that a deep groove between the head and body domain, which is at the exterior of nucleoprotein oligomers, is probably the RNA-binding site (Figs 3b, d and 5). The surface of the groove is made up of a large number of basic residues, including R65, R150, R152, R156, R174, R175, R195, R199, R213, R214, R221, R236, R355, K357, R361 and R391, that are likely to function in RNA binding by interacting with the phosphodiester backbone. An aromatic residue, Y148, is found at one end of the groove and may stack with a nucleotide base (Fig. 5b). Eleven out of the sixteen basic residues observed at the putative RNA-binding surface are conserved among influenza A, B and C viruses (of strains A/WSN/33, B/Ann Arbor/1/66, and C/California/78, respectively), further supporting a role in RNA binding. The aromatic residue Y148 is also conserved. Because these positively charged residues are widely distributed along the nucleoprotein polypeptide, previous mutagenesis studies have identified only a small subset of the basic amino acids in the RNA-binding groove[4–6]. With RNA bound to the groove at the outer periphery of oligomeric nucleoprotein, the distance between two neighbouring binding sites in a nucleoprotein trimer is approximately 70 Å, consistent with the stoichiometry of one nucleoprotein per 24 nucleotides of RNA, as previously determined[7].

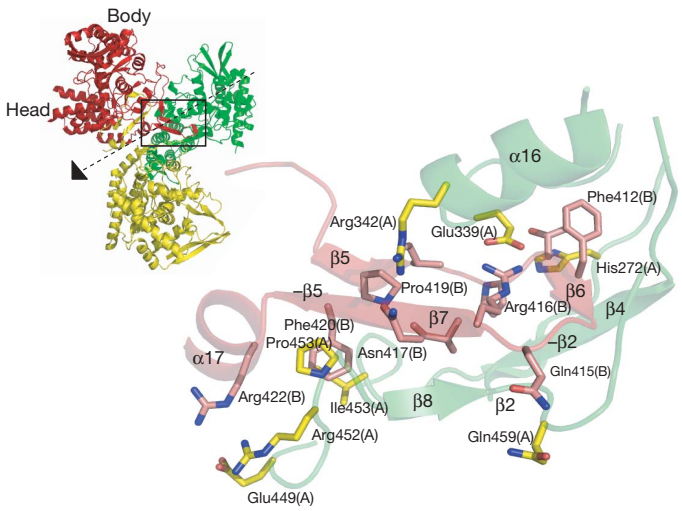

Figure 4 | Inter-subunit interactions mediated by the tail loop. Subunits A and B are shown in green and red, respectively. Amino acid side chains are drawn only for those residues making important interactions (for a thorough list, see Supplementary Table 2). The last capitalized letter in amino acid labels is used to show the molecular origin of that residue; for example, Glu 339(A) means residue Glu 339 from subunit A. The tail loop region is marked on a nucleoprotein trimer (boxed region), the three-fold axis of which is shown by a dotted arrow. At this orientation, the nucleoprotein trimer is tilted with the body domain situated on top.

To determine whether our nucleoprotein preparation possesses RNA-binding activity, we carried out gel-shift experiments using a 24-nucleotide single-stranded (ss)RNA, which is predicted to bind only a single nucleoprotein molecule[7,8]. At a nucleoprotein:RNA ratio of 1:1 and 2:1, 95% and 100% of the RNA, respectively, is bound to nucleoprotein (Supplementary Fig. 3). These nucleoprotein:RNA complexes form trimeric oligomers like those observed for RNA-free nucleoprotein (Fig. 2b), indicating that RNA binding itself does not affect nucleoprotein–nucleoprotein association, and that RNA-bound nucleoprotein is likely to use a similar set of oligomerization forces as those in free nucleoprotein. The polymerization of nucleoprotein into large RNP-like structures is likely to be the direct consequence of cooperative binding of nucleoprotein to long RNA molecules. Because the oligomerization tail loop is flexibly attached to the nucleoprotein body domain, we anticipate that many of the inter-subunit interactions now observed in the nucleoprotein crystal structure would be preserved in the continuous nucleoprotein:RNA helical fibre in RNPs, and that similar nucleoprotein–nucleoprotein interactions would also be maintained when the RNP is partially unwound during RNA synthesis. On the basis of the observation that RNA-bound nucleoprotein appears to curve slightly towards the interior in a mini-RNP[9], it is possible that RNA binding to the exterior groove in the nucleoprotein induces a small change in the relative orientation of the body and head domains. Our results indicate that RNA would be largely exposed on the external surface of nucleoprotein molecules in the RNP, consistent with the fact that influenza virus RNA coated with nucleoprotein molecules is RNase-sensitive[10]. In contrast, the rhabdovirus and bornavirus nucleoproteins have their RNA-binding groove facing the interior[2,3,11], and would be expected to protect the bound RNA from RNase digestion, as has been observed[12].

The nucleoprotein structure indicates that the non-conventional nuclear localization signal (nNLS: ^{3}**S**x**GTKRSY**xx**M**, with important residues highlighted in bold)[13–15] previously identified at the amino terminus of the influenza nucleoprotein can function as its NLS. Because influenza viral RNA replication and RNP assembly both take place in the nucleus[1], newly synthesized nucleoproteins have to be transported into the nucleus. The nNLS sequence is located at the outer periphery of the nucleoprotein trimer where it is highly accessible to solvent (Fig. 3d). In addition, this region is disordered in electron density maps, and this flexibility should allow the nNLS to fit easily into the substrate recognition pocket on importin-α. Another NLS, a classical bipartite NLS (cNLS: ^{198}**KR**GINDRNFWRG ENG**RKTR**) that is located near the middle of the polypeptide, has been described[16]. The nucleoprotein structure suggests that this sequence does not function as a NLS, because the ^{198}**KR** and the ^{213}**RKTR** motifs are ~15 Å apart, a distance far shorter than the

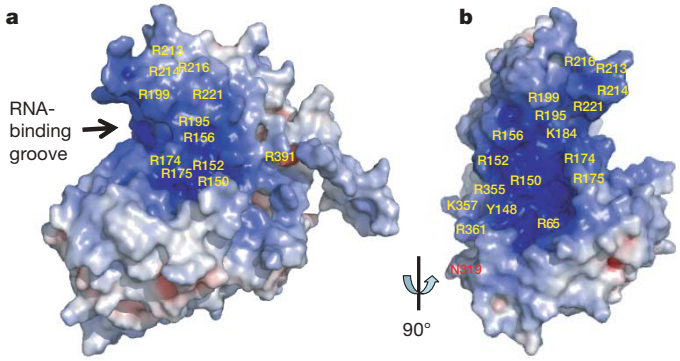

Figure 5 | RNA-binding by nucleoprotein. a, **b**, Electrostatic potential distribution for nucleoprotein (subunit B) identifies a potential RNA-binding site. Panel **a** is in the same orientation as the red subunit in Fig. 3b. Positive and negative electrostatic potentials are shown in blue and red, respectively.

minimal separating distance of 28 Å needed for efficient binding of a bipartite NLS[17]. It has also been proposed that nucleoprotein contains a cytoplasmic accumulation signal (CAS, residues 327–345), which interacts with F-actin and causes cytoplasmic retention of nucleoprotein late in infection[16,18]. However, this region is buried internally in the nucleoprotein structure and in this conformational state would not be able to bind F-actin.

The nucleoprotein structure provides a framework for delineating the mechanism by which nucleoprotein interacts with the viral RNA polymerase to promote viral RNA replication. The interaction of the nucleoprotein with the viral polymerase is required for viral RNA replication, but not for the synthesis of viral messenger RNAs (transcription)[19–21]. Previous experiments implicated three nucleoprotein regions—amino acids 1–160, 256–340 and 340–498—in binding to the PB1 and PB2 subunits of the viral polymerase[22]. Most of these amino acids (1–149, 273–340, 453–498) are localized in the body domain of the nucleoprotein (Fig. 3c, d). We propose that conserved amino acid regions in the body domain that are on the surface of nucleoprotein mediate nucleoprotein–polymerase interactions that are crucial for viral RNA replication. This hypothesis is consistent with the observation that a single amino acid change of N→K in residue 319, which is located on the surface at the right-hand side of the nucleoprotein body domain (Fig. 5b), resulted in an increase of polymerase activity[23]. This nucleoprotein structure will enable investigators to make rational mutations to determine the roles of specific nucleoprotein amino acids in functional interactions with the viral RNA polymerase.

METHODS

The nucleoprotein gene from influenza virus A/WSN/33, which codes for a 498-residue protein, was cloned into vector pET28a and expressed in *E. coli* strain Rosetta 2 (DE3) Singles competent cells (Novagen). The sequence encoding the first seven amino acid residues was deleted. Expressed nucleoprotein contains a C-terminal His-tag provided by the vector. The nucleoprotein gene from influenza virus A/Hong Kong/1074/99 (H9N2), which also encodes a protein of 498 residues, was cloned into a baculovirus vector with an extra sequence 'MPRGSHHHHHHGMASMTGGQQMGRDLYDDDDKDPSSR' at the N terminus. The expressed nucleoproteins were purified using Ni-NTA affinity (Qiagen), a HiTrap heparin HP column and a Superdex 200 gel filtration column (Amersham Biosciences). Purified nucleoprotein was at least 95% pure according to Coomassie SDS–polyacrylamide gel electrophoresis and had a A_{280nm}/A_{260nm} ratio of at least 1.8.

For gel-shift experiments, a 24-nucleotide ssRNA (6.5 μM) and various amounts of nucleoprotein were incubated in 15 μl binding buffer (containing 50 mM Tris-HCl pH 7.5, 200 mM NaCl, 1 mM dithiothreitol). The RNA substrate was a 24-nucleotide ssRNA (5′-UUUGUUACACACACACACGCUGUG-3′) obtained by *in vitro* transcription. Reaction mixtures were incubated at 37 °C for 30 min before being subjected to non-denaturing electrophoresis on a 3–9% polyacrylamide gel.

For electron microscopy, nucleoprotein samples were stained with freshly prepared 0.75% uranyl formate solution. Images were taken using a JEOL 2010 electron microscope operating at 120 kV, with a magnification of ×50,000 and a defocus of 2 to 3 μm. The sample of nucleoprotein:RNA complex was prepared by incubating nucleoprotein (650 nM) with the 24-nucleotide ssRNA at 1:2 molar ratio.

Crystals of nucleoprotein were grown using the vapour diffusion method with the drop made of equal volume of protein solution (5 mg ml^{-1}) and well solution (containing 100 mM Tris-HCl pH 7.5, 3% PEG8000, 10% glycerol, 10 mM dithiothreitol). Diffraction data were collected from single frozen crystals. A Se derivative was obtained using SeMet-substituted nucleoproteins. Other three heavy atom derivatives were prepared by soaking native crystals overnight in stabilizing solution containing 0.5 mM heavy atom compounds, followed by back-soaking for at least 2 h. Heavy atom derivative data were collected at the *f″* peak wavelength determined from fluorescence scan.

Diffraction data were collected at the Brookhaven National Light Source (BNLS) and the Cornell High Energy Synchrotron Source (CHESS). All diffraction data were processed using the program HKL2000 (ref. 24).

The nucleoprotein structure was determined by multiple isomorphous replacement with anomalous scattering (MIRAS) (Supplementary Table 1). Heavy atom sites were initially determined from the anomalous signal using the program SOLVE[25] and refined using SHARP[26]. Experimental phases were further improved by solvent flipping, histogram matching and three-fold NCS averaging. The NCS symmetry operator was determined manually and refined using the program DM[27]. The protein model, which was built using the program O[28], was refined using CNS[29] with NCS restraint and later using REFMAC[30] without restraints. The polypeptide tracing was confirmed using heavy atom sites. For instance, 18 of the 19 modelled methionines coincide with strong peaks ($\sigma > 2.5$) in the anomalous difference Fourier map of the SeMet derivative. All figures were prepared using the program PYMOL (W. L. Delane, http://www.pymol.org) unless noted otherwise.

Received 3 October; accepted 26 October 2006.
Published online 6 December 2006.

1. Lamb, R. A. & Krug, R. M. in *Fields Virology* (eds Knipe, D. M. & Howley, P. M.) 1487–1531 (Lippincott-Raven, Philadelphia, 2001).
2. Albertini, A. A. *et al.* Crystal structure of the rabies virus nucleoprotein-RNA complex. *Science* **313,** 360–363 (2006).
3. Green, T. J., Zhang, X., Wertz, G. W. & Luo, M. Structure of the vesicular stomatitis virus nucleoprotein-RNA complex. *Science* **313,** 357–360 (2006).
4. Elton, D., Medcalf, L., Bishop, K., Harrison, D. & Digard, P. Identification of amino acid residues of influenza virus nucleoprotein essential for RNA binding. *J. Virol.* **73,** 7357–7367 (1999).
5. Albo, C., Valencia, A. & Portela, A. Identification of an RNA binding region within the N-terminal third of the influenza A virus nucleoprotein. *J. Virol.* **69,** 3799–3806 (1995).
6. Kobayashi, M., Toyoda, T., Adyshev, D. M., Azuma, Y. & Ishihama, A. Molecular dissection of influenza virus nucleoprotein: deletion mapping of the RNA binding domain. *J. Virol.* **68,** 8433–8436 (1994).
7. Ortega, J. *et al.* Ultrastructural and functional analyses of recombinant influenza virus ribonucleoproteins suggest dimerization of nucleoprotein during virus amplification. *J. Virol.* **74,** 156–163 (2000).
8. Compans, R. W., Content, J. & Duesberg, P. H. Structure of the ribonucleoprotein of influenza virus. *J. Virol.* **10,** 795–800 (1972).
9. Area, E. *et al.* 3D structure of the influenza virus polymerase complex: localization of subunit domains. *Proc. Natl Acad. Sci. USA* **101,** 308–313 (2004).
10. Baudin, F., Bach, C., Cusack, S. & Ruigrok, R. W. 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 (1994).
11. Rudolph, M. G. *et al.* Crystal structure of the borna disease virus nucleoprotein. *Structure* **11,** 1219–1226 (2003).
12. Iseni, F., Barge, A., Baudin, F., Blondel, D. & Ruigrok, R. W. Characterization of rabies virus nucleocapsids and recombinant nucleocapsid-like structures. *J. Gen. Virol.* **79,** 2909–2919 (1998).
13. Wang, P., Palese, P. & O'Neill, R. E. The NPI-1/NPI-3 (karyopherin α) binding site on the influenza A virus nucleoprotein NP is a nonconventional nuclear localization signal. *J. Virol.* **71,** 1850–1856 (1997).
14. Neumann, G., Castrucci, M. R. & Kawaoka, Y. Nuclear import and export of influenza virus nucleoprotein. *J. Virol.* **71,** 9690–9700 (1997).
15. Cros, J. F., Garcia-Sastre, A. & Palese, P. An unconventional NLS is critical for the nuclear import of the influenza A virus nucleoprotein and ribonucleoprotein. *Traffic* **6,** 205–213 (2005).
16. Weber, F., Kochs, G., Gruber, S. & Haller, O. A classical bipartite nuclear localization signal on Thogoto and influenza A virus nucleoproteins. *Virology* **250,** 9–18 (1998).
17. Fontes, M. R., Teh, T. & Kobe, B. Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-α. *J. Mol. Biol.* **297,** 1183–1194 (2000).
18. Digard, P. *et al.* Modulation of nuclear localization of the influenza virus nucleoprotein through interaction with actin filaments. *J. Virol.* **73,** 2222–2231 (1999).
19. Medcalf, L., Poole, E., Elton, D. & Digard, P. Temperature-sensitive lesions in two influenza A viruses defective for replicative transcription disrupt RNA binding by the nucleoprotein. *J. Virol.* **73,** 7349–7356 (1999).
20. Honda, A., Ueda, K., Nagata, K. & Ishihama, A. RNA polymerase of influenza virus: role of NP in RNA chain elongation. *J. Biochem.* **104,** 1021–1026 (1988).
21. Shapiro, G. I. & Krug, R. M. 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 (1988).
22. Biswas, S. K., Boutz, P. L. & Nayak, D. P. Influenza virus nucleoprotein interacts with influenza virus polymerase proteins. *J. Virol.* **72,** 5493–5501 (1998).
23. Gabriel, G. *et al.* The viral polymerase mediates adaptation of an avian influenza virus to a mammalian host. *Proc. Natl Acad. Sci. USA* **102,** 18590–18595 (2005).
24. Otwinowski, Z. & Minor, W. (eds) *Processing of X-ray Diffraction Data in Oscillation Mode* (Academic Press, New York, 1997).
25. Terwilliger, T. C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallogr. D* **55,** 849–861 (1999).
26. Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr. D* **59,** 2023–2030 (2003).
27. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–763 (1994).

28. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47,** 110–119 (1991).
29. Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).
30. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53,** 240–255 (1997).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank V. Vikharia for providing the baculovirus for expressing nucleoprotein from influenza virus A/Hong Kong/1074/99 (H9N2); Z. Xie for electron microscopy analysis; P. Sliz for collecting the nucleoprotein native data; Y. Shamoo and G. Wu for providing the 24-nucleotide ssRNA; and Y. Shamoo, D. Mata and S. Harrison for discussions and critical reading of the manuscript. Nucleoprotein diffraction data were collected at the Brookhaven National Synchrotron Light Source and the Cornell High Energy Synchrotron Source. This work is supported by grants from the Welch Foundation (to Y.J.T and R.M.K), National Institutes of Health (to Y.J.T), and the Nanoscale Science and Engineering Initiative of the National Science Foundation.

Author Contributions Q.Y. crystallized the nucleoprotein. Q.Y. and Y.J.T solved the nucleoprotein crystal structure. Y.J.T, Q.Y. and R.M.K. wrote the paper. All authors discussed the results and commented on the manuscript.

Author Information The coordinates and the structure factor files have been deposited in the Protein Data Bank under accession code 2IQH. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to Y.J.T (ytao@rice.edu).