Influenza a virus antiparallel helical nucleocapsid-like pseudo-atomic structure

Florian Chenavier, Eleftherios Zarkadas, Lily-Lorette Freslon, Alice J Stelfox, Guy Schoehn, Rob W H Ruigrok, Allison Ballandras-Colas, Thibaut Crépin

Nucleic Acids Research, 2024, 1–10
https://doi.org/10.1093/nar/gkae1211
Molecular and Structural Biology

Influenza a virus antiparallel helical nucleocapsid-like pseudo-atomic structure

Florian Chenavier [1], Eleftherios Zarkadas [2], Lily-Lorette Freslon [1], Alice J Stelfox [1], Guy Schoehn [1], Rob WH Ruigrok [1], Allison Ballandras-Colas [1,*] and Thibaut Crépin [1,*]

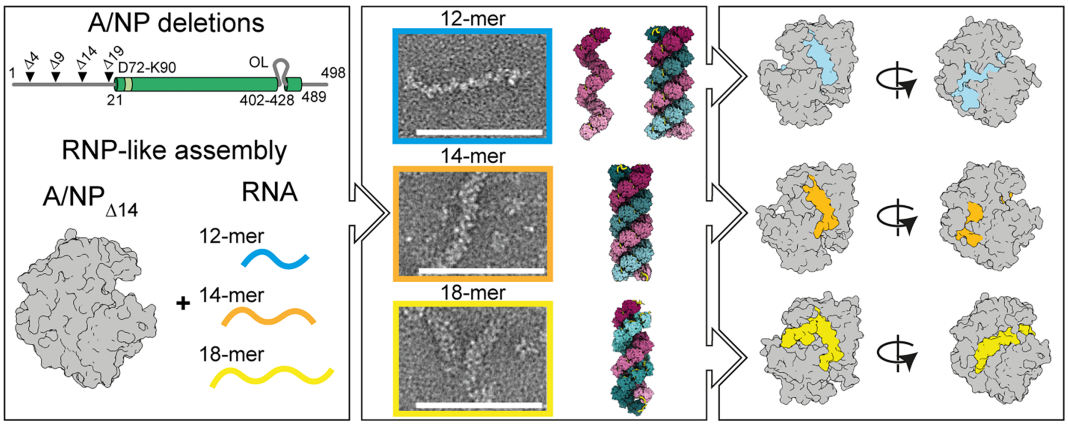
[1]Univ. Grenoble Alpes, CNRS, CEA, IBS, 71 avenue des Martyrs, F-38000 Grenoble, France
[2]Univ. Grenoble Alpes, CNRS, CEA, EMBL, ISBG, 71 avenue des Martyrs, F-38000 Grenoble, France

*To whom correspondence should be addressed. Tel: +33 476 209 439; Fax: +33 476 209 400; Email: thibaut.crepin@ibs.fr
Correspondence may also be addressed to Allison Ballandras-Colas. Tel: +33 457 428 686; Fax: +33 476 209 400; Email: allison.ballandras-colas@ibs.fr

Abstract

Influenza A viruses are responsible for human seasonal epidemics and severe animal pandemics with a risk of zoonotic transmission to humans. The viral segmented RNA genome is encapsidated by nucleoproteins (NP) and attached to the heterotrimeric polymerase, forming the viral ribonucleoproteins (vRNPs). Flexible helical vRNPs are central for viral transcription and replication. In this study, we present an advanced biological tool, the antiparallel helical RNP-like complex, assembled from recombinant N-terminally truncated NP and short synthetic RNA. The 3.0 Å cryo-electron microscopy structure details for the first time the whole RNA pathway across NP as well as NP-NP interactions that drive the antiparallel helical assembly accommodating major and minor grooves. Our findings show that the surface of the protein can harbour several conformations of the RNA, confirming that the number of nucleobases that binds to NP is not fixed, but ranges probably between 20 and 24. Taking all together, our data provide details to further understand the genome encapsidation and explain the inherent flexibility of influenza A virus vRNPs.

Graphical abstract



Introduction

Influenza A virus (IAV) causes annual epidemics and occasional pandemics of respiratory disease, the most recent example being the 2009 pandemic caused by low-pathogenic H1N1 viruses. Presently, there is a risk of a catastrophic pandemic precipitating if adaptions of the widespread high-pathogenic avian H5N1 viruses enable efficient human-to-human transmission. IAV viral genome contains eight single-stranded negative-polarity RNAs assembled into ribonucleoprotein particles (vRNPs). Each vRNP comprises a single heterotrimeric polymerase bound to both complementary 3′ and 5′ RNA termini and multiple copies of the viral nucleoprotein (NP) that encapsidates each of the eight viral RNA (vRNA) segments over the entire length without any sequence specificity (1,2). The vRNP resembles a large loose loop, supercoiled into a two antiparallel stranded helical filament (2). The vRNPs are largely flexible, and thus, all cryo-electron microscopy (cryo-EM) structures available to date are only of nanometer resolution (3–5), and neither allow a precise visualization and/or the orientation of the different proteins, nor

Received: August 21, 2024. **Revised:** October 29, 2024. **Editorial Decision:** November 18, 2024. **Accepted:** November 21, 2024

the vRNA. At the same time, the last two decades were highly prolific for the structural analysis of the individual components of the vRNP. The analysis of the heterotrimeric polymerase greatly benefited from the advances in cryo-EM, in particular for the structural study of the molecular mechanisms of replication and transcription (6). However, despite multiple efforts to detail the encapsidation of the viral genome, the understanding of NP dynamics within the nucleocapsid (NP-RNA part of the vRNP) was limited for a long time to the analysis of the protein alone and its propensity to self-assemble in small oligomers, whatever the virus strain (7–10). IAV NP is folded into two alpha-helical domains arranged in a bean-shaped configuration with short intrinsically disordered N- and C-terminal extremities and an internal loop that promotes the oligomerization (OL) between neighbouring protomers (Figure 1A). *In vitro*, it was shown that the OL can be controlled: the monomer is stabilized at low salt concentration (11) but any salinity change and/or ligand addition induces the formation of several kinds of oligomers, from trimers to small rings (7,12–16). In 2021, Tang and coworkers circumvented the problem using largely modified NP constructs. By truncating the internal OL loop, they crystallized the complex between NP and a 9-mer poly-uridine (17). Only three nucleotides were visible in the final structure, the first two lying in a pocket at the surface of NP and the third one making crystal contacts between two adjacent asymmetric units. Recently, we established that long recombinant nucleocapsid-like particles similar to vRNPs could be obtained by incubating *in vitro* purified full-length monomeric NP with a short synthetic $^{5'}$P-$(UC)_6$-fluorescein$^{3'}$ RNA ($^{5'}$P-$(UC)_6$-FAM$^{3'}$, called 12-mer for clarity) (15). This protocol was then used to obtain a subnanometer cryo-EM structure of a helical NP-RNA complex. Although this helical nucleocapsid-like structure showed a parallel double-stranded conformation, it allowed for the first time the visualization of both NP-NP and NP-RNA interactions. During binding to the RNA, several rearrangements of NP were observed, in particular the loop spanning residues D72 to K90 (D72-K90 loop) that folded into an α-helix. In this structure, the RNA is located at the interface of neighboring NP protomers with a local resolution between 4.5 and 6.8 Å only, suggesting that these nucleocapsid-like particles were still flexible. From that, we were only able to model a 12-mer phospho-ribose backbone in the density, without any detail for the nucleobases, the 5′ to 3′ direction of the RNA, nor the role of the FAM molecule in the assembly (18). In the present study, we continued with the same strategy, but by slightly modifying NP and modulating the RNA length, we have identified conditions enabling the formation of an improved double-stranded antiparallel helical particle mimicking the vRNP and allowing the detailed description of the molecular basis for the RNA encapsidation.

Materials and methods

Cloning and plasmid preparation

Primers were designed for the amplification of the different constructs of A/NP (A/$NP_{\Delta 4}$: 5–498; A/$NP_{\Delta 9}$: 10–498; A/$NP_{\Delta 14}$: 15–498 and A/$NP_{\Delta 19}$: 20–498). Each DNA fragment was cloned into pET22b bacterial expression vector (EMBL) using 5′ end Nde1 and 3′ end Xho1 restriction sites introduced by PCR amplification. All constructs were expressed as C-terminally His-tagged proteins. Sequencing verifications were performed by Eurofins Genomics.

Sample preparation

The recombinant NP (A/WSN/1933) purification and RNP-like assembly were performed as previously described by Labaronne *et al.* (15). Briefly, wild-type and mutant NPs were produced in *Escherichia coli* BL21 RIL (Agilent) and purified following a three-step procedure, Ni^{2+} affinity chromatography (Ni-NTA; Qiagen), followed by purification on a heparin column (Cytiva) and size-exclusion chromatography on a Superdex increase S75 10/300 GL column (Cytiva). The protein in the final buffer 20 mM HEPES, pH 7.5, 50 mM NaCl, 5 mM β-mercaptoethanol were concentrated to 100 μM and stored at −80°C. The nucleocapsid-like particles were assembled by incubating each RNA overnight at room temperature with a 1:1 ratio of A/$NP_{\Delta 14}$ in 20 mM HEPES pH 7.5, 150 mM NaCl, 5 mM β-mercaptoethanol. The UC-repeated sequence was used for its inability to form secondary structures as well as its high affinity for A/NP (18).

Negative stain grid preparation and visualization

A 4 μl drop of freshly assembled nucleocapsid-like diluted to 0.05 mg/ml was applied to the clean carbon side of a carbon/mica interface. The sample on carbon was then immediately stained on a 200 μl drop of 2% (w/v) sodium silicotungstate (pH 7.0), fished with a 400-mesh copper grid (Quantifoil) and air-dried at room temperature. Images were collected on a Tecnai F20 electron microscope (FEI Tecnai, Hillsboro, OR, USA) operating at 200 kV at a nominal magnification of ×55000 and equipped with an FEI Ceta camera.

Cryo-grid preparation and data collection

The grids were prepared as previously described by Chenavier *et al.* (18), except using UltrAuFoil grids R1.2/1.3 300 mesh Au (Quantifoil). Briefly, the sample was incubated 30 min on ice with 2 mM methyl-PEG8–*N*-hydroxysuccinimide ester reagent (Sigma-Aldrich) before plunge freezing using a Mark IV Vitrobot (Thermo Fisher Scientific). Data collection were performed on a Talos Glacios 200 kV FEG microscope (Thermo Fisher Scientific) equipped with a K2 summit direct electron detector camera (Gatan). A total of 9143, 2417 and 5697 movies of 40–60 frames each were collected using SerialEM (19), with a dose rate of 1–0.9 $e^-/Å^2$ per frame, which corresponds to a total dose of 40–53 $e^-/Å^2$. Movies were collected at a nominal magnification of ×36000, corresponding to a pixel size of 1.145 Å per pixel with a defocus ranging from −0.5 to −2 μm.

Image processing and model building

A similar image processing strategy was used for all the datasets. Movies were imported in Relion 4.0 (20) in different optics group. MotionCorrection was performed using their own implementation of motioncorr2 algorithm. The aligned and dose weighted micrographs have been imported into CryoSPARC v4.5.1 (21) for further processing. Patch CTF estimation was performed, followed by curation of the exposures to select those with a CTF fit below 4–5 Å resolution and to exclude those with ice too thick. The particles were selected by template picking using the previously determined structure (EMDB-11862) as a template and a minimal

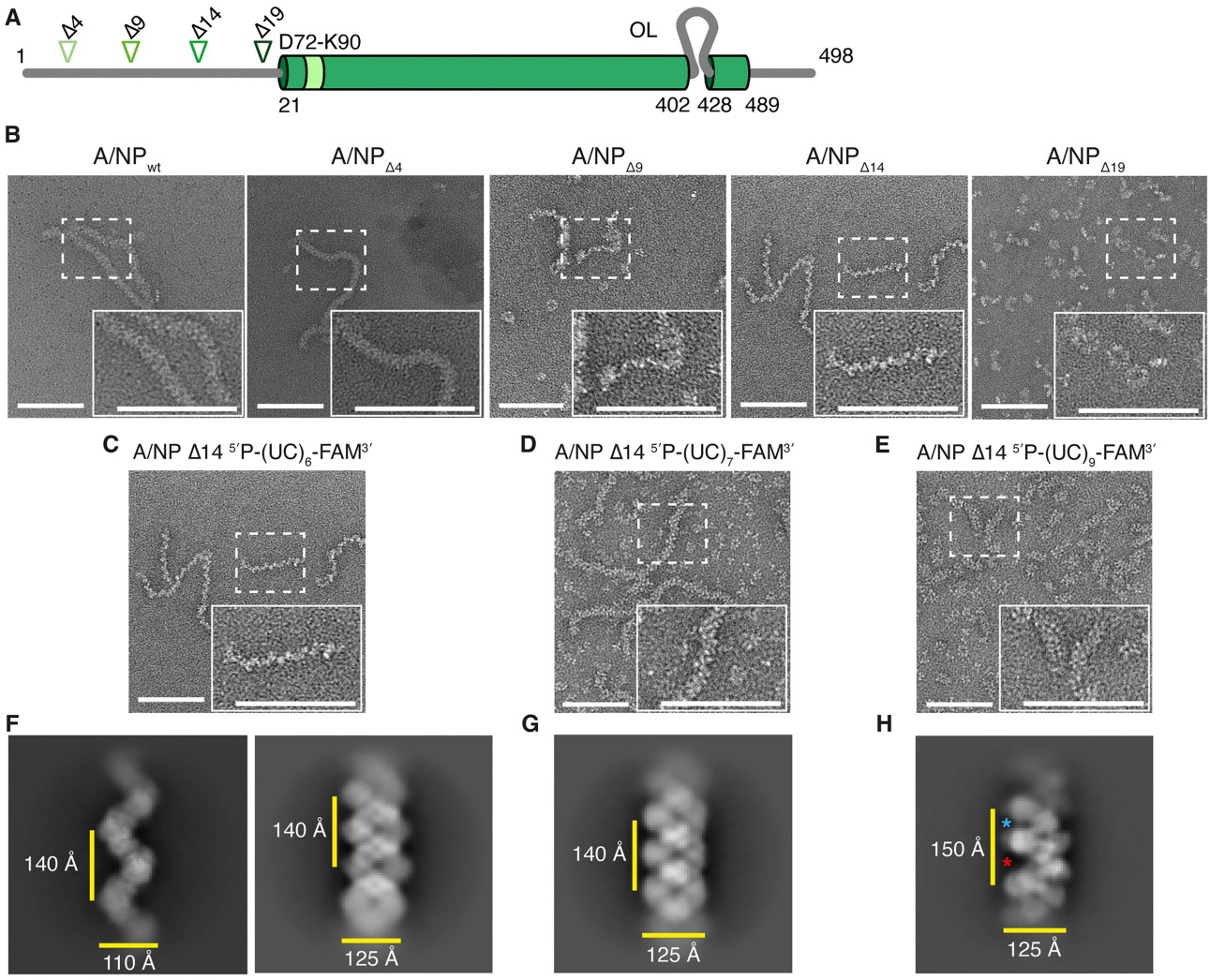


Figure 1. Impact of the N-terminal deletion on the nucleocapsid-like assembly. (**A**) Schematic representation of A/NP with the position of the N-terminal deletions (arrows). The structured core is represented as a cylinder with the unfolded terminals parts as a grey line. The position of the D72-K90 (pale-green) and OL loops are indicated. (**B**) Electron microscopy observations (negative staining) of the A/NP-RNA complexes obtained with the wild-type protein or the different deletion mutants after an overnight incubation with the $^{5'}$P-$(UC)_6$-FAM$^{3'}$. (**C**–**E**) Electron microscopy observations (negative staining) of the nucleocapsid-like particles assembled with by incubating $A/NP_{\Delta 14}$ with RNA molecules of various length. The RNA sequence remains a poly-UC with lengths of 12, 14 and 18 nucleobases for (C), (D) and (E) respectively. The scale bar represents 50 nm. (**F**–**H**) 2D classes from cryo-EM data collections from samples prepared for (C–E). The vertical and horizontal yellow bars correspond to the estimated helical pitch and diameter, respectively. The red and blue stars respectively indicate the major and minor groove position.

separation distance close to the estimated rise (27–30 Å). The particles were extracted into 400 × 400 pixels^2 boxes four times binned and subjected to several rounds of 2D classification to remove junk particles. The first *ab-initio* models were generated and further refined using either homogeneous or non-uniform refinement with a central mask encompassing 30% of the total box. By inspecting the resulting maps, the helical parameters were estimated using HI3D in real space (22). Based on the helical parameters, the particles with a separation distance less than three times the estimated rise were removed to avoid duplication of signal during further masked processing.

Regarding the single-stranded conformation using the 12-mer RNA (Supplementary Figure S1), a mask encompassing three neighbouring protomers was used for 3D classification (Relion 4.0) to obtain the best RNA density. Finally, the remaining particles were locally refined using the same mask and reached 3.3 Å resolution. For the parallel double-stranded conformation using the 12-mer (Supplementary Figure S1) or the 14-mer RNAs (Supplementary Figure S4), a C2 symmetry has been applied to average the information of both strands. Symmetry expanded particles were then subtracted, using a mask encompassing three neighbouring protomers on one strand, and further locally refined. The resolution estimated with CryoSPARC indicates 4.0 and 3.6 Å for the parallel double-stranded helical particle obtained with the 12- and 14-mer, respectively. For the antiparallel double-stranded conformation using the 18-mer RNA, the image processing was slightly different (Supplementary Figure S5). Based on the initial map, the particles separated by <90 Å were removed to avoid duplication on both strands. Helical parameters have been estimated using HI3D in real space and applied. A first high-resolution map has been obtained by locally refining the central part of the helix, revealing differences in arrangement and resolution on both strands. Local refinement focused on one or two protomers has been performed on the

two strands to achieve the highest resolution (3.04 and 3.08 Å, respectively). For each final map, the local resolution was estimated using CryoSPARC algorithm and the maps post-processed with a manually determined B-factor and with the EMReady algorithm (23).

The atomic models were built from the B-factor sharpened maps with the help of the EMready modified maps using COOT (24) and PHENIX (25). The atomic model (PDB: 8PZQ) was fitted in the central NP density. The RNA molecule was manually built as well as the C_{18}-FAM modified nucleotide. Each model was refined using real-space refinement against non-modified EM maps. Atomic model validation has been performed using MolProbity (26) and the PDB validation server. Finally, the central protomer was duplicated, fitted and refined within the external NP densities to model the NP-NP helical arrangement. The high-resolution model from the single-stranded conformation (12-mer RNA) was rigid fitted in the EM map of the double-stranded conformation (12-mer RNA) without further real space refinement.

All the figures were prepared using ChimeraX version 1.8 (27) or PyMOL (28).

Results

Partial deletion of NP N-terminal tail enables stable nucleocapsid-like assembly *in vitro*

As shown with measles virus, for which a more rigid version of the recombinant nucleocapsid-like particle was obtained by deleting the large intrinsically disordered C-terminal end of the nucleoprotein (29,30), we focused here on the twenty first N-terminal residues of NP (Figure 1). We expressed, purified variants that were truncated systematically every five residues (A/$NP_{\Delta 4}$, A/$NP_{\Delta 9}$, A/$NP_{\Delta 14}$ and A/$NP_{\Delta 19}$ corresponding to residues 5–498, 10–498, 15–498 and 20–498 of the wild-type NP, respectively) and tested their ability to form nucleocapsid-like particles *in vitro* with the 12-mer RNA (Figure 1B). The systematic shortening of the N-terminal extremity led to the progressive loss of NP ability to assemble spontaneously into regular thick helical filaments. Whereas A/$NP_{\Delta 4}$ formed long nucleocapsid-like filaments similar to the wild-type, we observed a proportional decrease in the thickness of the long helical filaments with A/$NP_{\Delta 9}$ and A/$NP_{\Delta 14}$. By negative staining electron microscopy, only various small closed oligomers in rings but no more helical filaments were observed with A/$NP_{\Delta 19}$ (Figure 1B). The electron micrographs visual analysis indicates that the fine helical assemblies observed with A/$NP_{\Delta 14}$ looked similar to single-stranded nucleocapsid of unwound vRNPs diluted in low salt conditions (2).

A cryo-EM dataset of A/$NP_{\Delta 14}$ incubated with the 12-mer RNA was acquired using a Talos Glacios transmission electron microscope (TEM) and two structures were obtained after image analysis: a right-handed single-stranded helix (major species) and a right-handed parallel double-stranded (minor species) at 3.3 and 4.0 Å resolution, respectively (Figure 1C and F; Supplementary Figure S1, Supplementary Table S1). Both have helical parameters close to the subnanometer structure obtained previously (18). Nonetheless, as A/$NP_{\Delta 14}$ appeared to have a greater propensity to assemble into regular and less flexible helices than the wild-type protein, we probed more RNA substrates by systematically increasing the length from 12 to 24 nucleotides by increments of 2 and checked the result with EM (Figure 1D, E, G and H; Supplementary Figure S2). It appeared that only two other RNA lengths enabled the assembly in nucleocapsid-like particles, and the resulting filaments appear wider in diameter than the single-stranded helical structure obtained by mixing A/$NP_{\Delta 14}$ with the 12-mer RNA. With the $^{5'}$P-$(UC)_7$-FAM$^{3'}$ (14-mer) RNA, the filaments looked similar in length and diameter to the original filaments obtained by incubating the wild-type NP and the 12-mer RNA. Using the $^{5'}$P-$(UC)_9$-FAM$^{3'}$ (18-mer) RNA, the filaments were systematically shorter yet comparable in diameter to the double-stranded filaments observed so far. For all the other RNA lengths tested, we observed mixtures of open and close rings. Cryo-EM datasets were acquired using the Glacios TEM and image analysis performed, resulting in two additional structures, a right-handed parallel double-stranded helix with the 14-mer RNA and a right-handed antiparallel double-stranded helix with the 18-mer RNA, at 3.6 and 3.0 Å resolution, respectively (Supplementary Figure S3 and Supplementary Table S1).

The structure obtained with the 14-mer RNA displays a parallel double-stranded helical organization (Supplementary Figures S3 and S4), nearly identical to the 12-mer double-stranded structure and at higher resolution.

CryoEM structure of the antiparallel helical nucleocapsid-like assembly

For the dataset of A/$NP_{\Delta 14}$ in complex with the 18-mer RNA, we followed the same data processing as previously described for the wild-type NP in complex with the 12-mer RNA (Supplementary Figure S5) (18). On the 2D classes, two types of grooves were observed (Figure 1H). The NP-RNA complex adopts two different conformations on the opposing strands, and each has been locally refined to 3.0 Å resolution (Supplementary Figures S6 and S7). Thus, a right-handed antiparallel double-stranded helix is obtained by incubating A/$NP_{\Delta 14}$ with the 18-mer RNA, in which a minor groove and major groove can be observed (Figure 2A).

However, the two strands are not related by a strict 180° rotation around the axis perpendicular to the helical axis. Indeed, NPs on the first strand (Strand-1) are slightly tilted toward the central helical axis while NPs on the second strand (Strand-2) are tilted away from the axis (Supplementary Figure S8), making the RNA molecule of Strand-1 less exposed to the solvent than Strand-2. Though NPs have different orientations, they are similarly spaced on both strands and thus they have the same pseudo-helical symmetry parameters. This architecture is stabilized at several levels (Figure 3). First, the cohesion of each strand is ensured by the internal loop known to promote the OL of NPs (the loop of NP_0 enters the NP_{-1} body that also interacts with the RNA through S413; Figure 3B), and is reinforced by the RNA located at the interface of neighbouring protomers (Figure 3A). Second, the integrity of the double-stranded structure is maintained by inter-strand interactions between one protomer (NP_0) from Strand-2 and two protomers (NP_0 and NP_{-1}) from Strand-1. In detail, R204, R208 and R213 of Strand-2/NP_0 make salt-bridges with D127, D128 and E80 of Strand-1/NP_0 (Figure 3C) and Q241, E254 and D439 of Strand-2/NP_0 make salt-bridges with R208, R213, R246 from Strand-1/NP_{-1} (Figure 3D). The RNA from one strand makes no contact with any protomers from the opposite strand nor the RNA, demon-

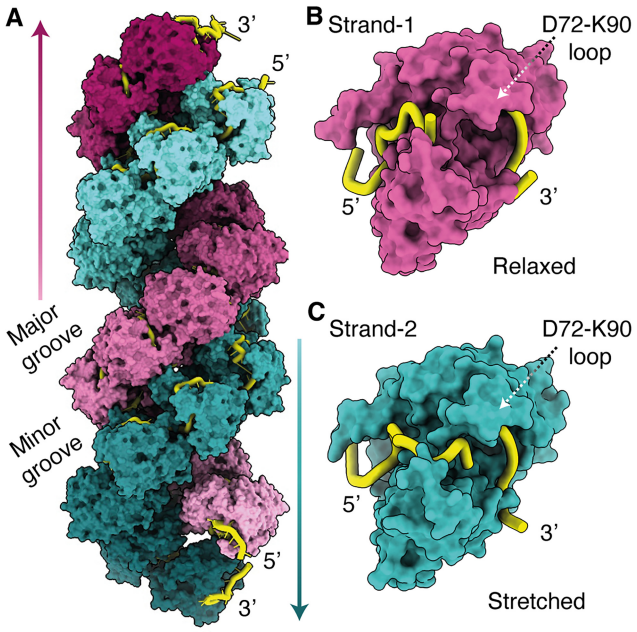


Figure 2. The antiparallel double-stranded nucleocapsid-like particle. (**A**) Helical symmetrized representation from the atomic model of one protomer per strand. Each protomer is shown in surface representation coloured as gradients of pink or cyan for Strand-1 and Strand-2, respectively, that follow the 5′ to 3′ direction of the RNA. The RNA is shown as a yellow ribbon. (**B**) and (**C**) RNA conformation on the two strands. A single protomer of A/NP$_{\Delta 14}$ is shown in surface with the RNA in ribbon. The colour code is the same as (A).

strating that the antiparallel helix is only stabilized by protein–protein contacts.

The RNA molecules of each opposite strands also adopt two different conformations, a relaxed (RNA in Strand-1) and a stretched (RNA in Strand-2) conformation (Figures 2B,C and 4), in correlation with the NP tilt per strand (Supplementary Figure S8). The quality of the cryo-EM map enabled modelling of the entirety of the two RNA molecules from residue U1 to C18, including the FAM molecules. For both, the main NP-RNA contacts are driven by the phospho-ribose backbone that is stabilized by numerous positively-charged residues of NP. Several nucleobases are stacked altogether, making only few specific base-contacts with the protein (detailed on Figure 4B and C), supporting the absence of sequence specificity. When superposing the NP$_0$ of the two opposite strands, the 5′ phosphate groups of U1 interact with S69 and R75 of protomers NP$_{-1}$, (Figure 4A and B). From U1 to C6, the two RNA molecules adopt exactly the same spatial conformation, stabilized by the same residues of NP$_{-1}$. From U7 onwards, the two RNAs follow globally the same pathway in the NP central basic groove, but their spatial positions differ, one stabilized before the D72-K90 loop (Strand-1; Figure 4B) and the other one passing below (Strand-2; Figure 4C). This flexible loop folds into an α-helix upon RNA binding, as described previously (18). In Strand-1, the N-terminal α-helix of NP$_{+1}$ enters the positively-charged groove of NP$_0$, preventing C18 from further entering the groove (Supplementary Figure S9A). The position C18 of the relaxed RNA in Strand-1 is thus constrained by the global helical conformation and stabilized before the D72-K90 loop. Regarding the stretched RNA in Strand-2, the N-terminal α-helix of NP$_{+1}$ does not interact with the neighbouring NP$_0$ and the RNA molecule occupies entirely the central basic groove (Supplementary Figure S9B). The RNA is engaged under the D72-K90 loop by an arginine-quadruplet made by two patches of two consecutive arginines (R74-R75 and R174-R175) that bends the phospho-ribose backbone of the RNA (Supplementary Figure S9C). The phosphate group of C16 is clamped between R74 and R174, where the latter residue also makes hydrogen bonds with the phosphate groups of U15 and C18 (Figure 4C). In this position, the 3′ OH group of C18 is located only 12 Å away from the 5′ phosphate group (stabilized by R75) of the next U1, R175 being positioned in the middle. By comparison, this distance is >20 Å for Strand-1. The FAM molecule neither promotes nor constrains either conformation of the RNA. In both cases, it lies on the surface of the protein in two distinct places, mainly stabilized by a single contact with an arginine, R121 or R75 for Strand-1 and Strand-2, respectively.

The structures of the single-stranded and the parallel double-stranded helices obtained with 12- or the 14-mer present an RNA pathway similar to that obtained with the relaxed conformation (Strand-1) of the antiparallel double-stranded helix (Supplementary Figure S10A). The position of all 5′ ends are identical, starting with the phosphate groups of U1 interacting with S69 and R75 of protomers NP$_{-1}$ and U1 to C10 adopting exactly the same spatial conformation (Supplementary Figure S10B and S10C). The position of U11-C12 of the 12-mer RNA is similar to that of the 18-mer RNA in Strand-1, with its FAM molecule lying at the surface of NP in a third position. For the structure obtained with the 14-mer RNA, the density is visible up to U13 while C14 and the FAM molecules are absent, and are likely too flexible to be built (Supplementary Figure S10C).

Discussion

We present here a biochemical approach to assemble IAV NP in right-handed helices *in vitro*, and more importantly in antiparallel nucleocapsid-like particles, for which a 3 Å cryo-EM structure was obtained. This provides for the first time, the molecular basis for the stabilization of the RNA along the whole IAV NP protein, and regarding the helical assembly we demonstrated the importance of the unfolded N-terminal tail of NP in the dynamics of this elaborate architecture.

Previously, we showed that the FAM attached at the 3′ of the $^{5'}$P-(UC)6 was necessary for assembly of the nucleocapsid-like particles *in vitro*, but no affinity between NP and the FAM alone was observed (15). When comparing all the structures in this manuscript, the FAM is observed on the surface of NP, at three different positions, without any extensive contact with the protein (Supplementary Figure S11). Therefore, this suggests that the FAM is not a driving force for the helical assembly but rather, it helps stabilizing the last nucleobase of the RNA on the surface of NP.

Our data shows for the first time that the whole RNA pathway crosses over IAV NP through the basic groove (Figure 2C) and details all the protein-RNA contacts. Regarding the sense of the RNA, our results are consistent with the work published by Tang and coworkers, where a C-terminally truncated monomeric NP was crystallized with three nucleobases methylated at the 2′-*O* positions (17). The RNA direction (5′ to 3′) is consistent between the X-ray and cryo-EM structures (Supplementary Figure S12). The second nucleobase (mU2) of

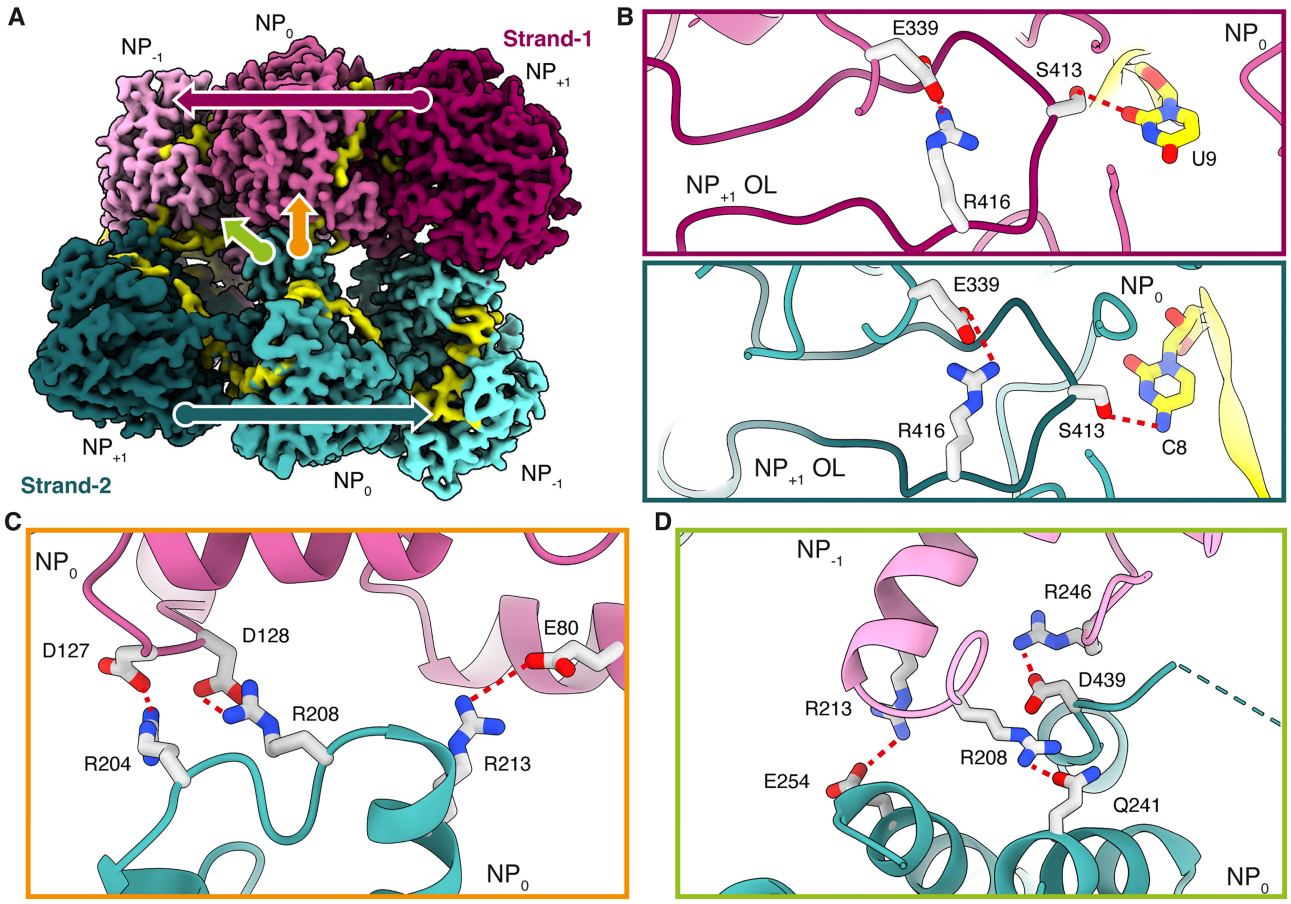


Figure 3. NP-NP interactions that stabilize the double-stranded helix. (**A**) Cryo-EM experimental map of the antiparallel double-stranded helical conformation. The two strands are coloured as gradients of pink or cyan for Strand-1 and Strand-2, respectively, with the RNA in yellow. The arrows represent the NP-NP and NP-RNA interactions coloured in orange/green or pink/blue, respectively. The EMready map is used for illustration purpose. (**B**–**D**), Residues involved in NP-NP and NP-RNA interactions that stabilizes the two strands. Each NP protomer is shown in cartoon representation coloured according to their position in the helix. The H-bonds are represented as red dotted lines. The colour code is identical to (A).

the X-ray model occupies the position of the U1 in the cryo-EM structures. However, the first nucleobase (mU1) has no equivalent and the third (mU3), being stabilized by crystal contacts, is displaced from the real RNA pathway observed in the cryo-EM structures. Other differences in the mechanism of RNA stabilization can also be observed. In the crystal structure, R74 contacts only the phosphate group of mU1, whereas in the cryo-EM structure it forms a clamp with R174 to stabilize the phosphate group of C16 when the RNA occupies the basic groove of NP. In the cryoEM structure, the D72-K90 loop is completely folded into an α-helix, and so R75 can interact with the phosphate groups of the RNA (of U1 in the cryo-EM models), which is not observed in the crystal structure.

Our findings on the RNA pathway across NP support previous reports suggesting that the number of nucleobases that bind to the protein is not fixed, but ranges between 20 and 24 (31,32). All four structures present the same spatial conformation of the RNA pathway on the right side of NP, compared to a higher mobility tolerance on the left side (Supplementary Figure S13). Indeed, the left side can accommodate a variable number of nucleobases and thus, one could propose that RNA loops involved in intersegment RNA interactions (33–35) may protrude from that position, and probably from the most exposed RNA position in the minor groove, as shown on the Strand-2 (Supplementary Figures S3C and S8).

We show that in presence of the 12-mer RNA, $NP_{\Delta 14}$ by itself can spontaneously resume a single-stranded helical conformation, with a pitch of ∼140 Å, which is double the size of the values observed for known single-stranded nucleocapsids of non-segmented RNA viruses (30,36,37). By increasing the size of the RNA, NP without its N-terminal tail, can also assemble in either parallel or antiparallel double-stranded conformations. We do not know the reasons for this parallel to antiparallel transition. We observe that by increasing the number of bases, the RNA induces the spreading of the protomers (Supplementary Figure S14), by displacing first, the head domain of protomer NP_{-1}, and second the N-terminal α-helix of NP_{+1}. When the RNA occupies the basic groove, the N-terminal α-helix of NP_{+1} is evicted from site (Supplementary Figure S9A and S9B). Being positioned between the NP protomers, the RNA acts as both the cement and the spacer of the helix and thus contributes to the flexibility of the helices. Furthermore, the deletion of the unfolded N-terminal tail of NP enables the generation of a more rigid helical nucleocapsid-like particles, leading to a 3 Å resolution structure. By comparison, our previous structure obtained with the full-length NP (and solely in presence of the 12-mer RNA), was of middle range resolution (18), while those of

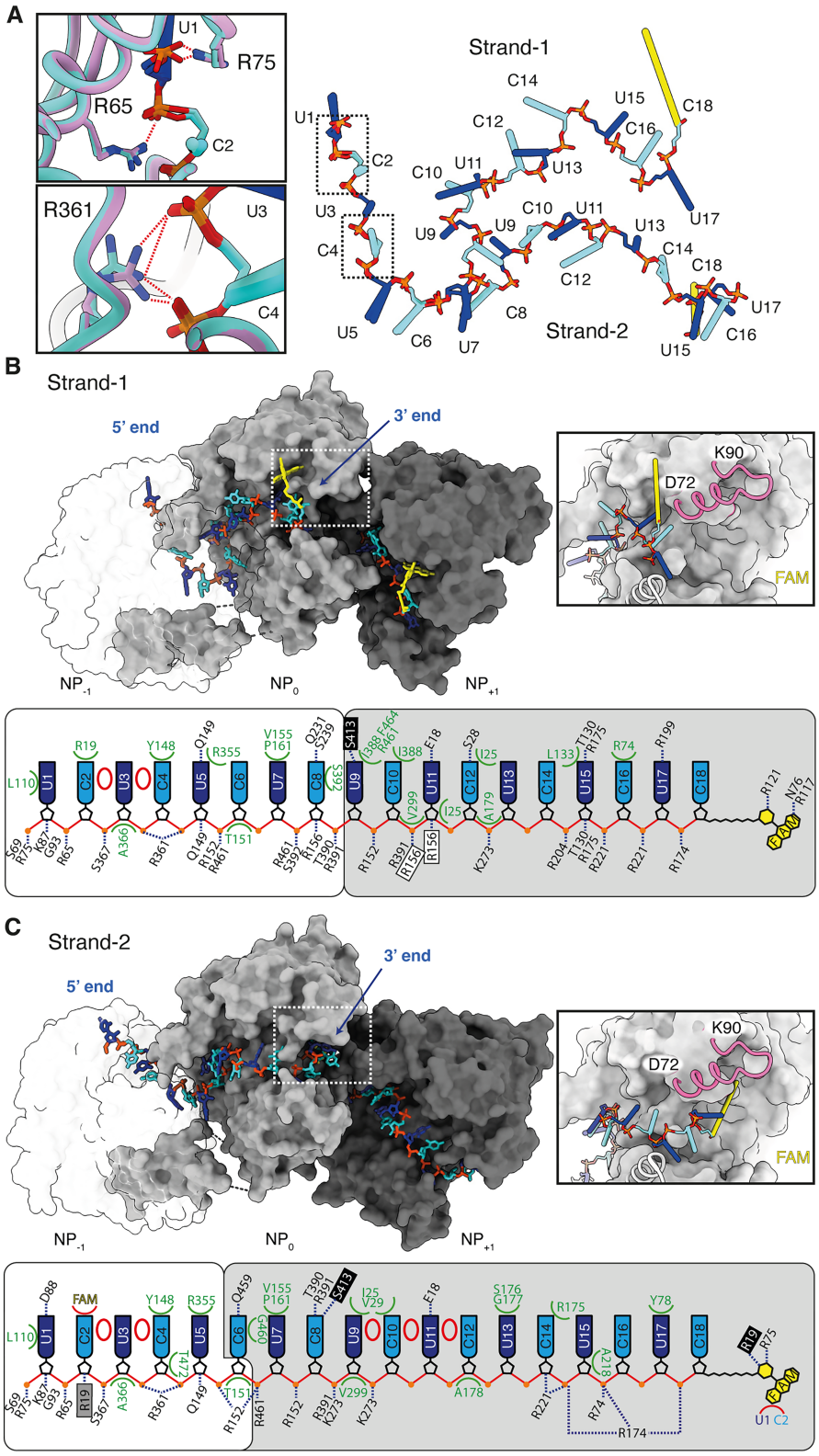


Figure 4. The 18-mer RNA in the double-stranded helix. (**A**) Alignments of the two RNA conformations after the superimposition of NP_0 from the two strands. The two RNAs are shown in stick representation. The positions of the six first nucleobases are aligned and the two conformational changes are visible from U7. The inserts on the left correspond to the close views of the stabilization of U1 and C2 (top) and U3 and C4 (bottom). (**B**–**C**) Details of the NP-RNA interactions. The surface representation of the NP_0, NP_{-1} and NP_{+1} in grey, white and black, respectively, with RNA as sticks, are displayed for Strand-1 (B) and Strand-2 (C). The dotted-line square represents close views of the D72-K90 loop shown on the right insets. Below the surface representations are schematized the details of the interactions between the proteins (NP_{-1} and NP_0 as a white and grey squares, respectively) and the RNA. The uracyl and cytosyl nucleobases are coloured in dark blue and cyan respectively with the FAM in yellow. NP residues with white, grey or black background indicate interactions with RNA coming from NP_{-1}, NP_0 or NP_{+1}, respectively. The blue dotted lines, green semi-circles and red ellipses represent the H-bonds interactions, the polar contacts and the nucleobases stackings, respectively.

vRNPs extracted from viruses, being even more flexible, are even less resolved (3–5). We then suspect that the presence of the fourteen first residues in full-length NP may prevent such stabilization and also contribute to a higher flexibility within the helix. Our results thus show that the fourteen first amino acids of A/NP concentrate at least two important functions for the biology of influenza virus, by (i) conferring a certain degree of flexibility to the vRNPs and (ii) ensuring their nuclear trafficking as well as that of the neosynthesised NPs (38–40).

With this approach to assemble influenza A antiparallel nucleocapsid-like particles *in vitro*, we now have the most advanced tool to understand (i) the encapsidation process of the RNA in the context of the vRNP and (ii) the dynamic protein-RNA architecture. The RNA would never be involved in the inter-strand interaction, but rather the antiparallel nucleocapsid helical structure would only be stabilized by protein–protein contacts. This is consistent with a previous study showing that, under specific conditions, RNA-free NP purified from influenza virus A/PR/8/34 forms RNP-like polymers (41). All residues involved in inter-strand protein–protein interactions have strictly conserved charges among all IAV strains, including the clade 2.3.4.4b H5N1 recently identified in the dairy cattle in the US and the first fatal human H5N2 virus (Supplementary Figure S15). The only exception regards position 437, which is a strictly conserved threonine, except in bat IAV NPs where there is an isoleucine. The orientation of NP observed within the antiparallel nucleocapsid-like structure is consistent with the native vRNP, and supports most of the biological data published so far. For example, NP residues (R204, W207 and R208) are solvent exposed at one extremity of the helical structure, and therefore available to interact with the polymerase (42). In the same way, the residues (48, 52, 98, 99 and 313) of NP targeted by human restriction factors MxA and BTN3A3 to prevent host-species jumps (43,44) are also solvent exposed all along the helical nucleocapsid-like. Therefore, we demonstrate that the RNA and NP are both actors in the helicity and the flexibility of the vRNPs, while the RdRp, by joining both 5′ and 3′ ends of the vRNA segments, dictates their antiparallelity.

Data availability

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Extended Data. Cryo-EM density maps of RNP-like resolved in this study and the associated atomic coordinates have been, respectively, deposited in the Electron Microscopy Data Bank (EMDB) and in the PDB under the following accession codes: 9GAN (EMD-51184), 9GAP (EMD-51185), 9GAQ (EMD-51186), 9GAV (EMD-51189), 9GAS (EMD-51187) and 9GAT (EMD-51188).

Supplementary data

Supplementary Data are available at NAR Online.

Acknowledgements

We thank Léna Louracia for technical assistance, Leandro Estrozi for discussion on helices, Francois-Xavier Briand, Beatrice Grasland and Mariette Ducatez for the sequences. We also thank FC's thesis committee, Serena Bernacchi, Catherine Bougault, and Juan Reguera, and Hélène Malet and Quentin Durieux Trouilleton for discussions. We also acknowledge the molecular graphics and analyses performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization and Informatics at the University of California, San Francisco, with support from National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases. This work used the platforms of the Grenoble Instruct-ERIC center (ISBG ; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB), supported by FRISBI (ANR-10-INBS-0005-02) and GRAL, financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). The IBS-ISBG EM facility is supported by the Auvergne-Rhône-Alpes Region, the Fondation Recherche Medicale (FRM), the fonds FEDER and the GIS-Infrastructures en Biologie Sante et Agronomie (IBISA).

Florian Chenavier: Conceptualization, Methodology, Formal analysis, Validation, Visualization, Writing—original draft and Writing—review & editing. Eleftherios Zarkadas: Methodology and Formal analysis. Freslon Lily-Lorette: Methodology and Formal analysis. Alice J Stelfox: Methodology and Writing—review & editing. Guy Schoehn: Conceptualization, Methodology, Formal analysis, Validation, Visualization, Writing—review & editing, Supervision and Funding acquisition. Rob WH Ruigrok: Conceptualization and Writing—review & editing. Allison Ballandras-Colas: Conceptualization, Methodology, Formal analysis, Validation, Visualization, Writing—review & editing and Funding acquisition. Thibaut Crépin: Conceptualization, Methodology, Formal analysis, Validation, Visualization, Writing—original draft, Writing—review & editing, Supervision and Funding acquisition.

Funding

Chemistry Biology Health Graduate School of University Grenoble Alpes [ANR-17-EURE-0003]. Agence Nationale pour la Recherche [ANR-23-CE11-01]. Funding for open access charge: Agence Nationale pour la Recherche [ANR-23-CE11-01].

Conflict of interest statement

The authors declare that they have no competing interests.

References

1. Baudin,F., Bach,C., Cusack,S. and Ruigrok,R.W. (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.
2. Klumpp,K., Ruigrok,R.W.H. and Baudin,F. (1997) Roles of the influenza virus polymerase and nucleoprotein in forming a functional RNP structure. *EMBO J.*, **16**, 1248–1257.
3. Arranz,R., Coloma,R., Chichón,F.J., Conesa,J.J., Carrascosa,J.L., Valpuesta,J.M., Ortín,J. and Martín-Benito,J. (2012) The structure of native influenza virion ribonucleoproteins. *Science*, **338**, 1634–1637.

4. Moeller,A., Kirchdoerfer,R.N., Potter,C.S., Carragher,B. and Wilson,I.A. (2012) Organization of the influenza virus replication machinery. *Science*, **338**, 1631–1634.
5. Coloma,R., Arranz,R., de la Rosa-Trevín,J.M., Sorzano,C.O.S., Munier,S., Carlero,D., Naffakh,N., Ortín,J. and Martín-Benito,J. (2020) Structural insights into influenza A virus ribonucleoproteins reveal a processive helical track as transcription mechanism. *Nat. Microbiol.*, **5**, 727–734.
6. Zhu,Z., Fodor,E. and Keown,J.R. (2023) A structural understanding of influenza virus genome replication. *Trends Microbiol.*, **31**, 308–319.
7. Ye,Q., Krug,R.M. and Tao,Y.J. (2006) The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature*, **444**, 1078–1082.
8. Ng,A.K.-L., Zhang,H., Tan,K., Li,Z., Liu,J.-H., Chan,P.K.-S., Li,S.-M., Chan,W.-Y., Au,S.W.-N., Joachimiak,A., *et al.* (2008) Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. *FASEB J.*, **22**, 3638–3647.
9. Ng,A.K., Lam,M.K., Zhang,H., Liu,J., Au,S.W., Chan,P.K., Wang,J. and Shaw,P.C. (2012) Structural basis for RNA binding and homo-oligomer formation by influenza B virus nucleoprotein. *J. Virol.*, **86**, 6758–6767.
10. Donchet,A., Oliva,J., Labaronne,A., Tengo,L., Miloudi,M., C A Gerard,F., Mas,C., Schoehn,G., W H Ruigrok,R., Ducatez,M., *et al.* (2019) The structure of the nucleoprotein of Influenza D shows that all *Orthomyxoviridae* nucleoproteins have a similar NP_{CORE}, with or without a NP_{TAIL} for nuclear transport. *Sci. Rep.*, **9**, 600.
11. Chenavas,S., Estrozi,L.F., Slama-Schwok,A., Delmas,B., Di Primo,C., Baudin,F., Li,X., Crépin,T. and Ruigrok,R.W.H. (2013) Monomeric nucleoprotein of influenza A virus. *PLoS Pathog.*, **9**, e1003275.
12. Gerritz,S.W., Cianci,C., Kim,S., Pearce,B.C., Deminie,C., Discotto,L., McAuliffe,B., Minassian,B.F., Shi,S., Zhu,S., *et al.* (2011) Inhibition of influenza virus replication via small molecules that induce the formation of higher-order nucleoprotein oligomers. *Proc. Natl Acad. Sci. U.S.A.*, **108**, 15366–15371.
13. Tarus,B., Bakowiez,O., Chenavas,S., Duchemin,L., Estrozi,L.F., Bourdieu,C., Lejal,N., Bernard,J., Moudjou,M., Chevalier,C., *et al.* (2012) Oligomerization paths of the nucleoprotein of influenza A virus. *Biochimie*, **94**, 776–785.
14. Tarus,B., Chevalier,C., Richard,C.-A., Delmas,B., Di Primo,C. and Slama-Schwok,A. (2012) Molecular dynamics studies of the nucleoprotein of influenza A virus: role of the protein flexibility in RNA binding. *PLoS One*, **7**, e30038.
15. Labaronne,A., Swale,C., Monod,A., Schoehn,G., Crépin,T. and Ruigrok,R.W.H. (2016) Binding of RNA by the nucleoproteins of influenza viruses A and B. *Viruses*, **8**, E247.
16. Yang,F., Pang,B., Lai,K.K., Cheung,N.N., Dai,J., Zhang,W., Zhang,J., Chan,K.H., Chen,H., Sze,K.H., *et al.* (2021) Discovery of a novel specific inhibitor targeting influenza A virus nucleoprotein with pleiotropic inhibitory effects on various steps of the viral life cycle. *J. Virol.*, **95**, e01432-20.
17. Tang,Y.-S., Xu,S., Chen,Y.-W., Wang,J.-H. and Shaw,P.-C. (2021) Crystal structures of influenza nucleoprotein complexed with nucleic acid provide insights into the mechanism of RNA interaction. *Nucleic Acids Res.*, **49**, 4144–4154.
18. Chenavier,F., Estrozi,L.F., Teulon,J.M., Zarkadas,E., Freslon,L.L., Pellequer,J.L., Ruigrok,R.W.H., Schoehn,G., Ballandras-Colas,A. and Crepin,T. (2023) Cryo-EM structure of influenza helical nucleocapsid reveals NP-NP and NP-RNA interactions as a model for the genome encapsidation. *Sci. Adv.*, **9**, eadj9974.
19. Schorb,M., Haberbosch,I., Hagen,W.J.H., Schwab,Y. and Mastronarde,D.N. (2019) Software tools for automated transmission electron microscopy. *Nat. Methods*, **16**, 471–477.
20. Kimanius,D., Dong,L., Sharov,G., Nakane,T. and Scheres,S.H.W. (2021) New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochem. J.*, **478**, 4169–4185.
21. Punjani,A., Rubinstein,J.L., Fleet,D.J. and Brubaker,M.A. (2017) cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods*, **14**, 290–296.
22. Sun,C., Gonzalez,B. and Jiang,W. (2022) Helical indexing in real space. *Sci. Rep.*, **12**, 8162.
23. He,J., Li,T. and Huang,S.Y. (2023) Improvement of cryo-EM maps by simultaneous local and non-local deep learning. *Nat. Commun.*, **14**, 3217.
24. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta. Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
25. Afonine,P.V., Poon,B.K., Read,R.J., Sobolev,O.V., Terwilliger,T.C., Urzhumtsev,A. and Adams,P.D. (2018) Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D Struct. Biol.*, **74**, 531–544.
26. Chen,V.B., Arendall,W.B. 3rd, Headd,J.J., Keedy,D.A., Immormino,R.M., Kapral,G.J., Murray,L.W., Richardson,J.S. and Richardson,D.C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta. Crystallogr. D Biol. Crystallogr.*, **66**, 12–21.
27. Meng,E.C., Goddard,T.D., Pettersen,E.F., Couch,G.S., Pearson,Z.J., Morris,J.H. and Ferrin,T.E. (2023) UCSF ChimeraX: tools for structure building and analysis. *Protein Sci.*, **32**, e4792.
28. Schrodinger,LLC (2015) The PyMOL Molecular Graphics System, Version 1.8.
29. Schoehn,G., Mavrakis,M., Albertini,A., Wade,R., Hoenger,A. and Ruigrok,R.W.H. (2004) The 12 A structure of trypsin-treated measles virus N-RNA. *J. Mol. Biol.*, **339**, 301–312.
30. Gutsche,I., Desfosses,A., Effantin,G., Ling,W.L., Haupt,M., Ruigrok,R.W.H., Sachse,C. and Schoehn,G. (2015) Near-atomic cryo-EM structure of the helical measles virus nucleocapsid. *Science*, **348**, 704–707.
31. Ortega,J., Martín-Benito,J., Zürcher,T., Valpuesta,J.M., Carrascosa,J.L. and Ortín,J. (2000) Ultrastructural and functional analyses of recombinant influenza virus ribonucleoproteins suggest dimerization of nucleoprotein during virus amplification. *J. Virol.*, **74**, 156–163.
32. Hutchinson,E.C., Charles,P.D., Hester,S.S., Thomas,B., Trudgian,D., Martínez-Alonso,M. and Fodor,E. (2014) Conserved and host-specific features of influenza virion architecture. *Nat. Commun.*, **5**, 4816.
33. Gultyaev,A.P., Tsyganov-Bodounov,A., Spronken,M.I., van der Kooij,S., Fouchier,R.A. and Olsthoorn,R.C. (2014) RNA structural constraints in the evolution of the influenza A virus genome NP segment. *RNA Biol.*, **11**, 942–952.
34. Dadonaite,B., Gilbertson,B., Knight,M.L., Trifkovic,S., Rockman,S., Laederach,A., Brown,L.E., Fodor,E. and Bauer,D.L.V. (2019) The structure of the influenza A virus genome. *Nat. Microbiol.*, **4**, 1781–1789.
35. Jakob,C., Paul-Stansilaus,R., Schwemmle,M., Marquet,R. and Bolte,H. (2022) The influenza A virus genome packaging network—complex, flexible and yet unsolved. *Nucleic Acids Res.*, **50**, 9023–9038.
36. Jenni,S., Horwitz,J.A., Bloyet,L.M., Whelan,S.P.J. and Harrison,S.C. (2022) Visualizing molecular interactions that determine assembly of a bullet-shaped vesicular stomatitis virus particle. *Nat. Commun.*, **13**, 4802.
37. Watanabe,R., Zyla,D., Parekh,D., Hong,C., Jones,Y., Schendel,S.L., Wan,W., Castillon,G. and Saphire,E.O. (2024) Intracellular Ebola virus nucleocapsid assembly revealed by in situ cryo-electron tomography. *Cell*, **187**, 5587–5603.
38. Neumann,G., Castrucci,M.R. and Kawaoka,Y. (1997) Nuclear import and export of influenza virus nucleoprotein. *J. Virol.*, **71**, 9690–9700.
39. Cros,J.F., Garcia-Sastre,A. and Palese,P. (2005) An unconventional NLS is critical for the nuclear import of the influenza A virus nucleoprotein and ribonucleoprotein. *Traffic*, **6**, 205–213.
40. Sanchez,A., Guerrero-Juarez,C.F., Ramirez,J. and Newcomb,L.L. (2014) Nuclear localized Influenza nucleoprotein N-terminal

deletion mutant is deficient in functional vRNP formation. *Virol. J.*, **11**, 155.
41. Ruigrok,R.W. and Baudin,F. (1995) Structure of influenza virus ribonucleoprotein particles. II. Purified RNA-free influenza virus ribonucleoprotein forms structures that are indistinguishable from the intact influenza virus ribonucleoprotein particles. *J. Gen. Virol.*, **76**, 1009–1014.
42. Marklund,J.K., Ye,Q., Dong,J., Tao,Y.J. and Krug,R.M. (2012) Sequence in the influenza A virus nucleoprotein required for viral polymerase binding and RNA synthesis. *J. Virol.*, **86**, 7292–7297.
43. Petric,P.P., Schwemmle,M. and Graf,L. (2023) Anti-influenza A virus restriction factors that shape the human species barrier and virus evolution. *PLoS Pathog.*, **19**, e1011450.
44. Pinto,R.M., Bakshi,S., Lytras,S., Zakaria,M.K., Swingler,S., Worrell,J.C., Herder,V., Hargrave,K.E., Varjak,M., Cameron-Ruiz,N., *et al.* (2023) BTN3A3 evasion promotes the zoonotic potential of influenza A viruses. *Nature*, **619**, 338–347.

Received: August 21, 2024. **Revised:** October 29, 2024. **Editorial Decision:** November 18, 2024. **Accepted:** November 21, 2024