

**The following resources related to this article are available online at  
[www.sciencemag.org](http://www.sciencemag.org) (this information is current as of November 27, 2009 ):**

**Updated information and services**, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/cgi/content/full/326/5957/1279>

**Supporting Online Material** can be found at:

<http://www.sciencemag.org/cgi/content/full/326/5957/1279/DC1>

This article **cites 18 articles**, 11 of which can be accessed for free:

<http://www.sciencemag.org/cgi/content/full/326/5957/1279#otherarticles>

This article appears in the following **subject collections**:

Biochemistry

<http://www.sciencemag.org/cgi/collection/biochem>

Information about obtaining **reprints** of this article or about obtaining **permission to reproduce this article** in whole or in part can be found at:

<http://www.sciencemag.org/about/permissions.dtl>

Education, Culture, Sports, Sciences, and Technology of Japan (to T.T.), the Strategic Japan-UK Cooperation Program of the Japan Science and Technology Agency (to T.T.), a grant from Takeda Science Foundation (to Y.Y.), and a grant from Chungbuk BIT Research-Oriented University Consortium (to S.J.L.). Coordinates and

structure factors have been deposited to the PDB under accession code 3A6P.

### Supporting Online Material

www.sciencemag.org/cgi/content/full/326/5957/1275/DC1  
Materials and Methods

Figs. S1 to S11  
Tables S1 to S3  
References

6 July 2009; accepted 9 October 2009  
10.1126/science.1178705

# Crystal Structure of a Nucleocapsid-Like Nucleoprotein-RNA Complex of Respiratory Syncytial Virus

Rajiv G. Tawar,<sup>1\*</sup> Stéphane Duquerry,<sup>1,2\*</sup> Clemens Vornrhein,<sup>3</sup> Paloma F. Varela,<sup>1†</sup> Laurence Damier-Piolle,<sup>1</sup> Nathalie Castagné,<sup>4</sup> Kirsty MacLellan,<sup>5‡</sup> Hugues Bedouelle,<sup>6</sup> Gérard Bricogne,<sup>3</sup> David Bhella,<sup>5</sup> Jean-François Eléouët,<sup>4§</sup> Félix A. Rey<sup>1§</sup>

The respiratory syncytial virus (RSV) is an important human pathogen, yet neither a vaccine nor effective therapies are available to treat infection. To help elucidate the replication mechanism of this RNA virus, we determined the three-dimensional (3D) crystal structure at 3.3 Å resolution of a decameric, annular ribonucleoprotein complex of the RSV nucleoprotein (N) bound to RNA. This complex mimics one turn of the viral helical nucleocapsid complex, which serves as template for viral RNA synthesis. The RNA wraps around the protein ring, with seven nucleotides contacting each N subunit, alternating rows of four and three stacked bases that are exposed and buried within a protein groove, respectively. Combined with electron microscopy data, this structure provides a detailed model for the RSV nucleocapsid, in which the bases are accessible for readout by the viral polymerase. Furthermore, the nucleoprotein structure highlights possible key sites for drug targeting.

**H**uman respiratory syncytial virus (RSV) is an important viral agent of pediatric respiratory tract disease worldwide, causing pneumonia and bronchiolitis in infants (1). No vaccine is currently available, and effective treatments have yet to be developed (2). RSV is a nonsegmented, negative-strand RNA virus of the *Paramyxoviridae* family in the *Mononegavirales* order (3), which also includes the *Rhabdoviridae*, *Bornaviridae*, and *Filoviridae* families. RSV is further classified into the *Pneumovirus* genus within the *Pneumovirinae* subfamily of the *Paramyxoviridae* (4). The 15.2-kb genomic RNA contains 10 genes, 4 of which code for intracellular proteins that are involved in genome transcrip-

tion, replication, and particle budding and have orthologs in all *Mononegavirales*, namely N (nucleoprotein), P (phosphoprotein), M (matrix protein), and L (“large” protein, containing the RNA polymerase catalytic motifs).

The RSV genomic RNA forms a nuclease-resistant (1) helical ribonucleoprotein (RNP) complex with the N protein, termed nucleocapsid, which is used as template for RNA synthesis by the viral polymerase complex (5). Studies on Sendai virus [a paramyxovirus (6)] and on the vesicular stomatitis virus [VSV, *Rhabdoviridae* family (7)] have shown that the polymerase in replication mode includes an L-P-N core complex, whereas in transcription mode, it is a complex of L, P, and cellular proteins. This pattern is believed to hold true for all members of the *Mononegavirales*. Electron microscopy (EM) observations of authentic RSV nucleocapsids derived from infectious virions (8), as well as recombinant nucleocapsid-like RNP complexes (9), revealed flexible helical filaments 15 nm in diameter, with spikes protruding at 45°, resembling a head of wheat.

Structural data are necessary to understand the mechanism of RNA synthesis used by these viruses. The structure of L (>200 kD) is not known for any of the *Mononegavirales*. Three-dimensional (3D) structures of fragments of P are available for a number of them, but not for RSV. The crystal structure of N was determined for the Borna disease virus [BDV, *Bornaviridae* family (10)] and for the rhabdoviruses VSV (11) and rabies virus [RV (12)]. The latter two structures revealed

decamer and hendecamer N rings, respectively, which were bound nonspecifically to cellular RNA. Extrapolation from these structures to understand how the RNA in the nucleocapsid is presented to the VSV and RV polymerases for viral RNA synthesis was not straightforward, however, because the RNA was occluded inside the rings.

We obtained crystals of recombinant decameric RSV N-RNA rings diffracting to 3.3 Å resolution (13). Phase extension from a low-resolution EM reconstruction (14) using 20-fold real-space averaging yielded a very clear 3.3 Å-resolution electron density map in which the atomic model was built and refined (table S1). Despite the relatively limited resolution, the 20-fold redundancy made the positioning of atoms sufficiently accurate for an unambiguous interpretation of their interactions (hydrogen bonds, for instance), which would not be the case in a less-over-determined structure at 3.3 Å resolution.

The RNA chain runs within a basic surface groove, surrounding the periphery of the N protein ring, as a belt (Fig. 1 and fig. S2). Each N subunit interacts with seven ribonucleotides, resulting in an RNA chain of 70 bases in total. The RNA electron density is well defined, with the density of the bases corresponding to the mean of the four bases of the genetic code, reflecting the random composition of the bound RNA (15). We modeled the bases as cytosines. The RNA adopts an extended conformation with phosphate-sugar torsion angles and sugar pucker similar to those of A-DNA (table S2), except at two “switches.”

Each N subunit is organized as a core region containing two domains, N- and C-terminal (NTD and CTD), that are connected through a hinge region. The RNA groove is formed at the NTD/CTD interface, the interdomain connection lines its internal side, and the  $\alpha_{C3}$ -to- $\alpha_{C4}$  loop (Fig. 1) forms a lid partially covering the exposed side of the RNA. This NTD/CTD core of the molecule has N- and C-terminal extensions, termed N arm and C arm, respectively, that appear folded only in the context of the quaternary interactions in the ring. A search of the Protein Data Bank with the Dali server (16) indicated that the closest relatives are, as expected, the BDV, VSV, and RV N protein orthologs (table S3 and Fig. 2). Both the NTD and CTD interact laterally in the ring with their counterparts from the adjacent subunits but do not make tight contacts. The interacting surfaces are highly hydrated, primarily involving intermittent van der Waals contacts. The ring is stabilized by the RNA belt and by the N chain that results from the insertion of the N arm (residues 1 to 35) of one subunit into the

<sup>1</sup>Institut Pasteur, Unité de Virologie Structurale, Département de Virologie and CNRS Unité de Recherche Associée (URA) 3015, 25 Rue du Dr Roux, 75724 Paris Cedex 15, France. <sup>2</sup>Université Paris-Sud, Faculté d'Orsay, 91405 Orsay Cedex, France. <sup>3</sup>Global Phasing, Sheraton House, Castle Park, Cambridge CB3 0AX, UK. <sup>4</sup>L'Institut National de la Recherche Agronomique, Unité de Virologie et Immunologie Moléculaires, Domaine du Vilvert, 78350 Jouy-en-Josas, France. <sup>5</sup>Medical Research Council Virology Unit, University of Glasgow, Church Street, Glasgow G11 5JR, UK. <sup>6</sup>Institut Pasteur, CNRS URA 3012, 25 Rue du Dr Roux, 75724 Paris Cedex 15, France.

\*These authors contributed equally to this work.

†Present address: Synchrotron Soleil, Laboratoire de Biologie, L'Orme des Merisiers, Saint Aubin, Boite Postale 48, 91192 Gif-sur-Yvette Cedex, France.

‡Present address: National Institute for Biological Standards and Control, Blanche Lane, South Mimms, Potters Bar, Hertfordshire EN6 3QG, UK.

§To whom correspondence should be addressed. E-mail: jean-francois.eleouet@jouy.inra.fr (J.-F.E.); rey@pasteur.fr (F.A.R.)

compact fold of the adjacent one (Fig. 1B). This lateral connectivity helps explain the observed malleability of the N-N interactions in the flexible, yet very stable, RSV nucleocapsid. The C arm lies above the CTD in the ring, occupying the space that would be between consecutive turns of the helical nucleocapsid. It contains residues 361 to 391, the last 12 to 20 amino acids being disordered in the crystal (13).

Both core domains are  $\alpha$ -helical bundles, with 10  $\alpha$  helices in the NTD and 4 in the CTD. The NTD has 218 residues (36 to 253), with a long  $\beta$  hairpin projecting away from the molecule at the most distal end (Fig. 2). This region is not conserved in amino acid sequence within the *Pneumovirinae* (fig. S3) and is also most variable in 3D structure when compared with the other mononegavirus N proteins (Fig. 2). Mutations conferring in vitro resistance to an anti-RSV compound, RSV604 (17), map to this insertion in strand  $\beta_{12}$  and in helix  $\alpha_{12}$  (Fig. 2 and fig. S3).

The CTD has 107 residues (254 to 360), spanning the most conserved region in primary structure within the *Pneumovirinae*. Furthermore, the structural alignment shows that its four  $\alpha$

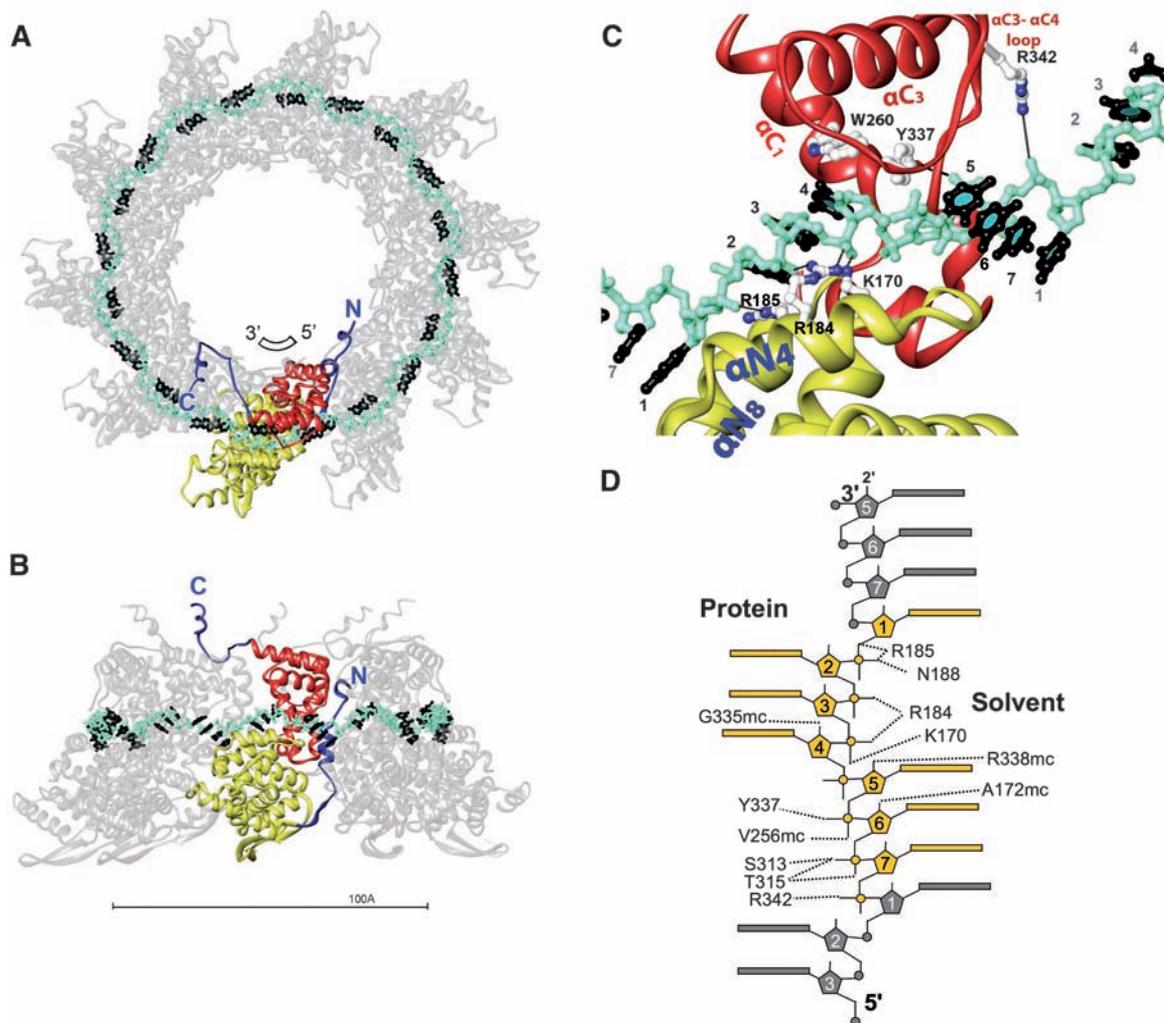
helices pack identically in the available N structures of the various mononegaviruses, in spite of the lack of sequence conservation. The  $\alpha_{C3}$ - $\alpha_{C4}$  loop (Fig. 1C) is responsible for many of the RNA contacts, with the side chain of Arg<sup>338</sup> arching above the RNA to make a salt bridge with Asp<sup>175</sup> in the NTD (fig. S2B).

The 7 nucleotides (nt) interacting with each N subunit have bases 2 to 4 stacked and facing the protein in a cavity within the groove, at the N-N interface (Fig. 1 and fig. S2C). Base 1 is sandwiched between the upstream base 7 and the backbone of helix  $\alpha_{N8}$  (Fig. 1C). This  $\alpha$  helix has two glycine residues in consecutive turns (Gly<sup>241</sup> and Gly<sup>245</sup>), which align to form a flat face of the helix on which base 1 packs. The packing of base 1 on the preceding base 7 results in a row of four stacked bases (5-6-7-1) facing solvent, away from the protein. The 2-3-4 base stack in the cavity is such that base 2 contacts Asn<sup>249</sup> and Arg<sup>185</sup>, and base 4, on the other side, contacts Trp<sup>260</sup> and Val<sup>256</sup>. Base 3, which is in the middle, makes no direct protein contacts. The RNA conformation switches from base out to base in at phosphate 1, and vice versa at phosphate 4 (going 3' to 5'), constrained by the presence of helix

$\alpha_{N8}$  at switch 1, and that of Tyr<sup>337</sup>, which packs against the ribose ring of nucleotide 4, at switch 2 (Fig. 1C). Mutation of Tyr<sup>337</sup> was reported to abolish nucleocapsid formation (18), highlighting the key role of its aromatic side chain in imposing the required conformation on the RNA chain. As indicated in Fig. 1D, all the observed hydrogen bonds are directed to the RNA backbone. In particular, the 2' OH of riboses 4, 5, and 6 donate hydrogen bonds to main-chain carbonyls, explaining the specificity for RNA instead of DNA.

The 3D superposition of the VSV N subunit onto its RSV counterpart reveals that the three-base stack of RNA facing the protein binds in the same way in the two complexes in a cavity at the N-N interface, in spite of the presence of a bulged-out base in the VSV complex (Fig. 3C). Although the RNA groove is in the same location in both cases, the lateral N contacts in the VSV and RV rings are such that the curvature is opposite to that of the RSV ring (Fig. 3). The result is an inside-out nucleocapsid ring, with the RNA inside and the N molecule oriented outside-in. Because each rhabdovirus N subunit contacts nine ribonucleotides instead of seven,

**Fig. 1.** A decameric ribonucleoprotein ring complex. One N subunit is colored according to domains: yellow and red indicate NTD and CTD, respectively, and the N and C arms are blue. The RNA is displayed with the backbone in cyan and the bases in black, except in (D). (A) View down the ring axis. The RNA polarity is indicated within the ring. (B) Side view. The bar indicates 100 Å. (C) Close-up of the protein-RNA interactions. (D) Details of the interactions with RNA. The schematic diagram is numbered from 3' to 5'. Pentagons, rectangles, and small circles represent riboses, bases, and phosphates, respectively. The relevant contacts with the protein are indicated. Main-chain contacts are indicated by mc.



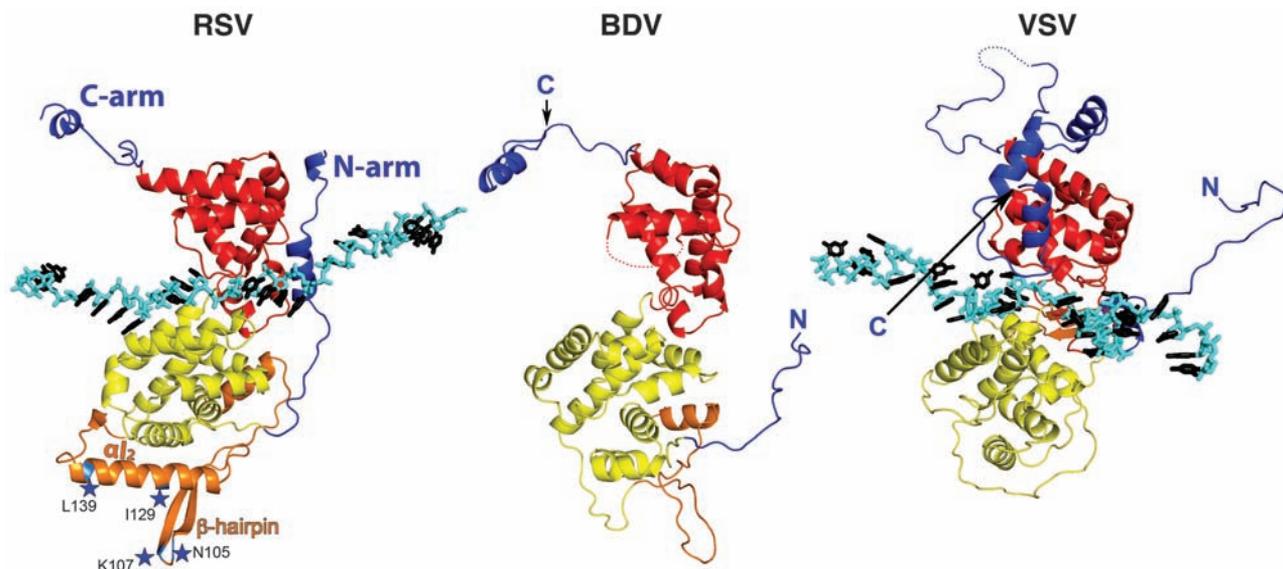
the RNA ring follows a more convoluted path to fit 90 nt in a smaller diameter than the 70-nt ring of the RSV counterpart (Fig. 3).

Although the N RNA contacts in the groove are not base-specific, the cavity appears tailored to bind a set of three stacked bases, a feature that appears to be conserved across the *Mononegavirales* order. Because the bases are aver-

aged out in our crystals, it is not possible to tell from the structure whether certain particular nucleotide sequences would make stronger or weaker interactions within the cavity. The overall arrangement of the RNA around the ring is reminiscent of the structure of the *trp* RNA-binding attenuation protein (TRAP)/RNA complex from *Bacillus subtilis* (19), which also forms a ring of 11 TRAP

subunits, with the RNA running at the periphery, and three stacked bases are inserted into a cavity located at the subunit interface. However, in contrast to the TRAP/RNA complex, the contacts with RSV N are not base-specific but rather RNA backbone-specific.

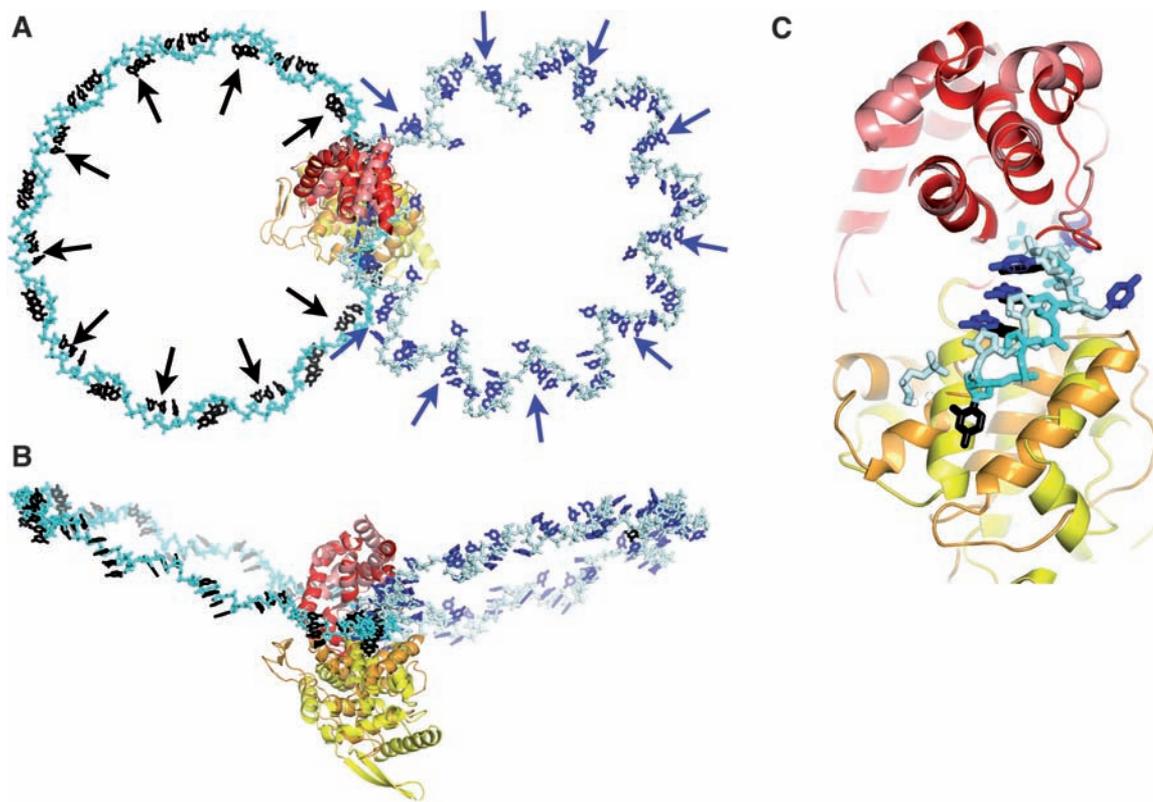
We calculated a 26 Å-resolution EM reconstruction (13) from cryo-negative stain images



**Fig. 2.** 3D fold of RSV N and comparison with BDV and VSV N. (Left) RSV N is colored according to domains as in Fig. 1. The variable region is in orange, with the  $\beta$  hairpin highlighted and labeled with the RSV604 resistance

mutation sites. The BDV (center) and VSV (right) N structures, respectively, oriented and colored identically to RSV N. The conserved domains (used for the alignment) are red and yellow. The RNA is colored as in Fig. 1.

**Fig. 3.** Comparison of RSV with the VSV decameric ring. Overlap of the RNA rings (RSV on the left, 70 bases, and VSV on the right, 90 bases) after superposition of the conserved domains (red and yellow in Fig. 2) of one N subunit. The arrows point to the sites where each N protein binds three stacked bases in each ring. Blue and black arrows indicate the binding sites of VSV and RSV N, respectively. (A) Top view, (B) side view. (C) Close-up showing the remarkable superposition of the three bases buried within the N protein groove, considering that in VSV, an intervening base is looped out. In the last three panels, the bases in the VSV RNA are blue instead of black, for clarity; VSV N is colored pale red and orange for the CTD and NTD, respectively (N and C arms were removed for clarity).



of nucleocapsid-like helical assemblies of recombinant N complexed with cellular RNA (Fig. 4B). The reconstruction showed that the repeating units form lateral contacts resembling those observed in the ring. Because of the limited resolution, we did not attempt a direct fitting of the individual N subunits into the EM reconstruction, but we used the least possible distortion to the ring contacts to generate the corresponding helix, as described in (13). The packing of the subunits in the ring indeed suggests a simple way of modeling the helical nucleocapsid with minor slippage about the lateral contacts. The model showed that the NTD is easily recognized as forming the spikes projecting at roughly 45° from the nucleocapsid axis (Fig. 4, C and D). The loose contacts between N subunits in the ring can readily adapt to the distortion introduced by enforcing a helical axis instead of the 10-fold ring axis, while maintaining the RNA connectivity, as shown in movies S1 to S3. The RNA can easily follow the contacts, making a helix of a pitch varying between 69 Å (which is the minimum pitch possible to avoid clashes with the subsequent turn) to more than 100 Å, with 10 to 11 N proteins per turn, accounting for the observed flexibility of the nucleocapsid. Also, the region between helical turns is occupied by the mobile C arm, which may play a functional

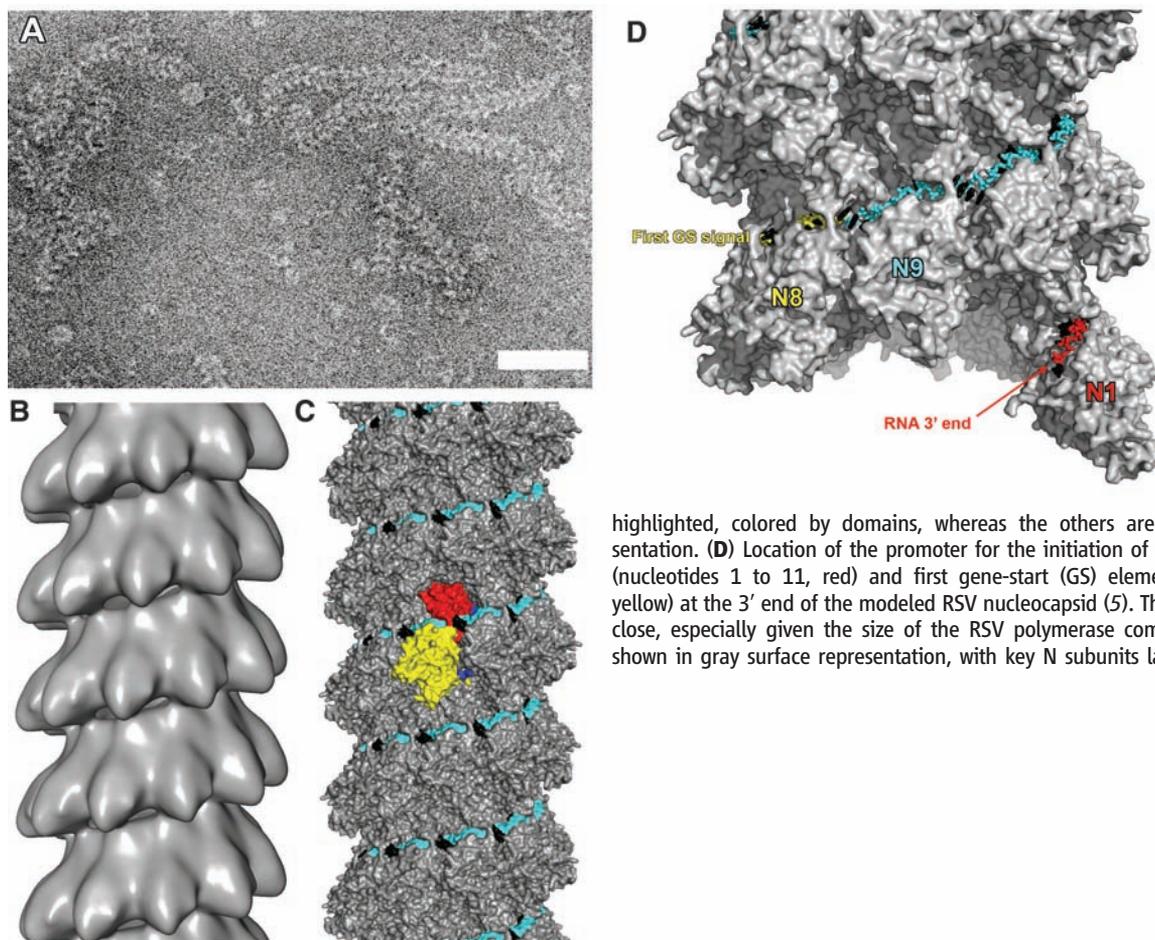
role by providing added flexibility to the nucleocapsid. RSV P was shown to bind the C arm of N (20), an interaction that may allow the polymerase complex to distort the helical conformation of the nucleocapsid during RNA synthesis.

Our model for the RSV nucleocapsid reveals a plausible way for the polymerase to thread through the RNA, reading the bases without needing to disassemble the nucleocapsid helix. Indeed, the domain organization of N suggests that the polymerase can induce a hinge movement of the NTD with respect to the CTD. The elongated NTD would act as a lever, with the polymerase contacting at its distal end (Fig. 2, orange) and causing the hinge movement, which would result in a transient opening of the groove during RNA readout. The location of the resistance mutations to the RSV604 compound (17) are thus quite likely to point to an interaction site of N with the polymerase complex. The hinge movement can make the three buried bases flip out, resulting in 11 bases in a row available for readout (5671-234-5671, in the numbering of Fig. 1). Support for this interpretation comes from studies using a phosphorylation mutant of RSV P that is impaired in transcription elongation, leading to the accumulation of abortive transcripts between 9 and 11 nt long (21). Given the large size of the polymerase complex, it is plausible

that during elongation, the complex can easily maintain at least 3 or 4 consecutive N subunits [as suggested for rhabdoviruses (22)] in an open-hinge conformation, allowing the release of 21 to 28 nucleotides from the nucleoprotein grip, such that these nucleotides can be accommodated within the polymerase active site for synthesis of the complementary strand. In this model, the N protein would act as a helicase, dissociating the transient double-stranded RNA segment during procession of RNA synthesis along the genome.

Several studies of RSV RNA synthesis [reviewed in (5)] indicate that both the structure of the 3' end of the RSV nucleocapsid, together with the specific 3' terminal RNA sequence, are recognized by the viral polymerase for initiation of viral RNA synthesis. Our model shows that the 3' promoter sequence and the first gene-start signal for transcription are spatially close to each other, at the first turn of the nucleocapsid helix (Fig. 4D).

The structure of the RSV RNP ring, together with the derived atomic model of the helical nucleocapsid, suggests important common features of the template for transcription and replication of viruses in the *Mononegavirales* order, which also includes other human pathogenic viruses such as those causing measles, mumps, Ebola



**Fig. 4.** The helical nucleocapsid. (A) Cryogenic negative-stain electron micrograph of recombinant RSV nucleocapsid-like helices. Scale bar, 50 nm. (B) 26 Å-resolution 3D reconstruction calculated from cryogenic negative-stain images. The helix comprises 9.8 N subunits per turn and has a pitch of 69 Å. (C) The helical nucleocapsid modeled from the contacts in the ring, using the same pitch (fig. S1), which results in 10.35 N subunits per turn (13). One subunit is

highlighted, colored by domains, whereas the others are in gray, in surface representation. (D) Location of the promoter for the initiation of replication and transcription (nucleotides 1 to 11, red) and first gene-start (GS) elements (nucleotides 45 to 54, yellow) at the 3' end of the modeled RSV nucleocapsid (5). The two sites are spatially very close, especially given the size of the RSV polymerase complex. The protein moiety is shown in gray surface representation, with key N subunits labeled from the 3' end.

fever, and rabies. Furthermore, in the case of RSV, these results reveal important interaction sites, for instance, the cavity for three stacked bases, the site of insertion of the N arm, or the tip of the NTD where the resistance mutations arise, which can be specifically targeted for the development of therapeutic treatments, interfering with encapsidation or other roles of N.

#### References and Notes

1. P. L. Collins, J. E. Crowe, in *Fields Virology* (Lippincott, Philadelphia, PA, ed. 5, 2007), pp. 1601–1646.
2. M. P. Frogel, *Manag. Care* **17**, 7 (2008).
3. C. R. Pringle, *Order Mononegavirales*, in *Virus Taxonomy* (Elsevier/Academic Press, London, 2005), pp. 609–614.
4. A. A. Brunt *et al.*, *Family Paramyxoviridae*, in *Virus Taxonomy* (Elsevier/Academic Press, London, 2005), pp. 655–671.
5. V. M. Cowton, D. R. McGivern, R. Fearn, *J. Gen. Virol.* **87**, 1805 (2006).
6. S. M. Horikami, J. Curran, D. Kolakofsky, S. A. Moyer, *J. Virol.* **66**, 4901 (1992).
7. K. R. Qanungo, D. Shaji, M. Mathur, A. K. Banerjee, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 5952 (2004).
8. T. Bachi, C. Howe, *J. Virol.* **12**, 1173 (1973).
9. D. Bhella, A. Ralph, L. B. Murphy, R. P. Yeo, *J. Gen. Virol.* **83**, 1831 (2002).
10. M. G. Rudolph *et al.*, *Structure* **11**, 1219 (2003).
11. T. J. Green, X. Zhang, G. W. Wertz, M. Luo, *Science* **313**, 357 (2006).
12. A. A. Albertini *et al.*, *Science* **313**, 360 (2006).
13. See supporting material on Science Online.
14. K. MacLellan, C. Loney, R. P. Yeo, D. Bhella, *J. Virol.* **81**, 9519 (2007).
15. K. El Omari *et al.*, *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **64**, 1019 (2008).
16. L. Holm, S. Kaariainen, P. Rosenstrom, A. Schenkel, *Bioinformatics* **24**, 2780 (2008).
17. J. Chapman *et al.*, *Antimicrob. Agents Chemother.* **51**, 3346 (2007).
18. M. Juozapaitis, M. Coiras, J. Staniulis, K. Sasnauskas, *Biologija (Vilnius)* **3**, 79 (2006).
19. A. A. Antson *et al.*, *Nature* **401**, 235 (1999).
20. J. Murray, C. Loney, L. B. Murphy, S. Graham, R. P. Yeo, *Virology* **289**, 252 (2001).
21. L. C. Dupuy, S. Dobson, V. Bitko, S. Barik, *J. Virol.* **73**, 8384 (1999).
22. A. Albertini, G. Schoehn, W. Weissenhorn, R. Ruigrok, *Cell. Mol. Life Sci.* **65**, 282 (2008).
23. This work was initiated as part of the sixth European Union Framework Programme consortium (<http://cordis.europa.eu/life/src/control/qlk2-ct-2001-01225.htm>). We thank A. Albertini, J. Bernard, D. Gertler, A. Haouz, J. Lepault, M. Moudjou, C. Schulze-Briese, E. Stura, P. Weber, and R.P. Yeo for help and/or discussion; M. Backovic, D. Kolakovsky, J. Melero, L. Roux, and A. Tortorici for comments on the manuscript. Diffraction data collection was done at beamline X06SA of the Swiss Light Source and ID23-2 of the European Synchrotron Radiation Facility. R.G.T. benefits from a Marie-Curie RTN fellowship (consortium MRTN-CT-2006-035599, acronym EIHCV). This work was also supported by the French Fondation pour la Recherche Médicale (postdoctoral fellowship to P.F.V.), by a grant from the French Agence Nationale pour la Recherche (Programme MIME) to F.A.R., and by Merck-Serono. The coordinates and structure factors of the N-RNA rings were deposited in the Protein Data Bank with accession number 2wj8. The 3D EM reconstruction was deposited in the Electron Microscopy Data Bank with accession code EMD-1622.

#### Supporting Online Material

[www.sciencemag.org/cgi/content/full/326/5957/1279/DC1](http://www.sciencemag.org/cgi/content/full/326/5957/1279/DC1)

Materials and Methods

Figs. S1 to S3

Tables S1 to S3

References

Movies S1 to S3

12 June 2009; accepted 8 October 2009

10.1126/science.1177634