



Cryo-EM structure of the Ebola virus nucleoprotein–RNA complex

Robert N. Kirchdoerfer,^a Erica Ollmann Saphire^{b*} and Andrew B. Ward^{a*}

^aDepartment of Integrative Structural and Computational Biology, The Scripps Research Institute, La Jolla, CA 92037, USA, and ^bDepartment of Immunology and Microbiology, The Scripps Research Institute, La Jolla, CA 92037, USA.

*Correspondence e-mail: erica@scripps.edu, andrew@scripps.edu

Received 13 February 2019

Accepted 1 April 2019

Edited by K. K. Kim, Sungkyunkwan University School of Medicine, Republic of Korea

Keywords: Ebola virus; nucleoprotein; NP; helix; cryo-electron microscopy.

PDB reference: Ebola virus NP–RNA helical filament, 6nut

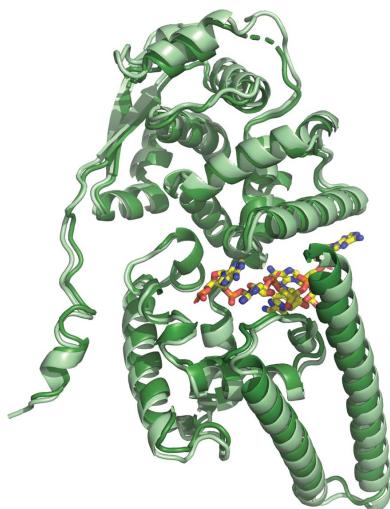
Supporting information: this article has supporting information at journals.iucr.org/f

Ebola virus is an emerging virus that is capable of causing a deadly disease in humans. Replication, transcription and packaging of the viral genome are carried out by the viral nucleocapsid. The nucleocapsid is a complex of the viral nucleoprotein, RNA and several other viral proteins. The nucleoprotein forms large, RNA-bound, helical filaments and acts as a scaffold for additional viral proteins. The 3.1 Å resolution single-particle cryo-electron microscopy structure of the nucleoprotein–RNA helical filament presented here resembles previous structures determined at lower resolution, while providing improved molecular details of protein–protein and protein–RNA interactions. The higher resolution of the structure presented here will facilitate the design and characterization of novel and specific Ebola virus therapeutics targeting the nucleocapsid.

1. Introduction

Ebola virus is capable of causing a disease manifesting as a severe hemorrhagic fever with case fatality rates as high as 90% (*Ebola haemorrhagic fever in Zaire, 1976, 1978*). Although generally outbreaks are geographically isolated, the outbreak of 2014–2015 was responsible for over 28 000 cases and spread between countries (Hersey *et al.*, 2015). As a non-segmented negative-sense RNA virus, the RNA genome of Ebola virus is encapsidated by a viral nucleoprotein (NP) throughout the virus life cycle. This protein–RNA complex serves as a scaffold for additional viral proteins, including the viral polymerase, to carry out viral transcription and replication. The complex of RNA, nucleoprotein and these additional viral proteins forms the nucleocapsid and is a structural component of Ebola virions. Full-length Ebola virus nucleoprotein is 739 amino acids in length. However, the N-terminal 450 amino acids (NP 1–450) are necessary and sufficient for protein oligomerization and RNA binding (Watanabe *et al.*, 2006), and this N-terminal region of the protein forms helical filaments resembling NP in virions (Bharat *et al.*, 2012).

The determinants of NP oligomerization and RNA binding have been the focus of several structural studies. Crystal structures and hydrogen/deuterium-exchange mass spectrometry have shown that the N-terminal region of NP possesses a folded core flanked by disordered regions (Kirchdoerfer *et al.*, 2015; Leung *et al.*, 2015). The folded core of NP possesses N- and C-terminal lobes and is primarily α -helical in structure. The available crystal structures of Ebola NP were all determined using monomeric NP (Dong *et al.*, 2015; Kirchdoerfer *et al.*, 2015; Leung *et al.*, 2015). In two of these structures, NP is bound to a peptide of viral protein 35 (VP35; Kirchdoerfer *et al.*, 2015; Leung *et al.*, 2015). The binding of this VP35 peptide



is thought to be important for maintaining a monomeric state of NP and preventing the premature and nonspecific encapsidation of host RNA (Kirchdoerfer *et al.*, 2015; Leung *et al.*, 2015; Leyrat *et al.*, 2011).

After the VP35 peptide has been removed, NP oligomerizes into a linear polymer capable of encapsidating RNA. Oligomerization of NP is mediated by protein regions N- and C-terminal to the two-lobed folded NP 1–450 core (Kirchdoerfer *et al.*, 2015; Leung *et al.*, 2015). This oligomerization of NP causes a conformational change between the two lobes, creating a high-affinity binding site for RNA (Kirchdoerfer *et al.*, 2015; Sugita *et al.*, 2018). Cryo-electron tomography (cryo-ET) studies of Ebola virus nucleocapsids as well as Ebola virus NP 1–450 demonstrated the left-handed helical nature of the NP filaments, as well as the first structural evidence of NP oligomerization regions mediating contacts between NP monomers (Wan *et al.*, 2017). However, the limited resolution of this earlier cryo-ET work precluded clear elucidation of the molecular interactions among NPs and with RNA. A more recent study of NP 1–450 using single-particle cryo-electron microscopy (cryo-EM) determined the structure of this helical filament at 3.6 Å resolution (Sugita *et al.*, 2018). This structure revealed molecular interactions as well as conformational changes between the monomeric, chaperoned NP and the oligomeric, RNA-bound NP. Here, we have used single-particle cryo-EM to determine the structure of the NP 1–450 helix at an improved resolution of 3.1 Å.

2. Materials and methods

2.1. Expression of NP helical filaments

NP filaments were prepared using similar methods to those described previously (Bharat *et al.*, 2012). 0.5 mg NP 1–450-pDisplay plasmid without any additional peptide tags was diluted into 20 ml Opti-MEM and then filtered through a 0.22 µm filter. To this solution was added 1.5 mg poly-ethylenimine (PEI) in 5 ml Opti-MEM. The DNA–PEI complexes were allowed to incubate for 20 min at room temperature and were then added to 11 293F cells at 1.1×10^6 cells ml⁻¹ growing in 293Freestyle medium in a 2 l baffled flask. Transfected cells were incubated at 37°C with 8% CO₂ at 130 rev min⁻¹ for 48 h.

2.2. Purification of NP helical filaments

The cells were collected by centrifugation at 1000g for 15 min and were then resuspended in 8 ml 25 mM Tris–HCl, 150 mM sodium chloride, 1 mM calcium chloride, 5 mM β-mercaptoethanol. The cells were then lysed by the addition of 10 ml 25 mM Tris–HCl pH 7.4, 150 mM sodium chloride, 1 mM calcium chloride, 0.2% Igepal CA-630, 5 mM β-mercaptoethanol. The lysate was clarified by centrifugation at 3200g for 20 min.

The NP helical filaments were isolated by loading the cleared lysate onto a double cushion of 5 ml 20% sucrose and 5 ml 90% sucrose. The sucrose cushions were spun in an ultracentrifuge using a Beckman SW28 rotor at 28 000 rev min⁻¹

for 3 h. The sucrose cushions were fractionated by bottom puncture and manually collected fractions of ~1 ml were analyzed by SDS–PAGE. Fractions containing NP 1–450 were dialyzed into 500 ml 25 mM Tris–HCl, 150 mM sodium chloride, 1 mM calcium chloride, 5 mM β-mercaptoethanol using 1 MDa molecular-weight cutoff dialysis tubing overnight with one buffer change to remove sucrose. 3 ml of the dialyzed protein was then loaded onto one of two discontinuous caesium chloride gradients containing 24 ml 25%(w/v) caesium chloride and 10 ml 60%(w/v) caesium chloride. The caesium chloride gradients were spun in a SW28 rotor at 28 000 rev min⁻¹ for 4 h. The gradients were then manually fractionated after bottom tube puncture and the fractions were analyzed with SDS–PAGE before overnight dialysis of fractions containing NP 1–450 to remove caesium chloride.

NP 1–450 was concentrated by pelleting the dialyzed protein (~3 ml) through 400 µl of a 20% sucrose cushion using a Beckman TLA100.3 rotor at 48 000 rev min⁻¹ for 1 h. The pellet was resuspended in 400 µl 25 mM Tris–HCl pH 7.4, 150 mM sodium chloride, 1 mM calcium chloride, 5 mM β-mercaptoethanol. Measuring the UV absorbance of the resuspended sample yielded an A_{280} of 3.8 with an A_{260}/A_{280} ratio of 1.8.

2.3. Cryo-electron microscopy data collection and preprocessing

UltraAuFoil R 1.2/1.3 Au 300 mesh grids (Quantifoil) were plasma-cleaned for 7 s with an Ar/O₂ gas mixture. 4 µl sample was blotted onto the grids and the grids were blotted for 4 s using a Vitrobot IV (Thermo Fisher) before plunge-freezing into liquid ethane. Data were collected using *Leginon* (Suloway *et al.*, 2005) on a Talos Arctica (Herzik *et al.*, 2017) (Thermo Fisher) operating at 200 kV and a K2 Summit detector (Gatan) in counting mode using a magnification of 43 478 and at a dose rate of 5.06 e per pixel per second, collecting 200 ms frames for a 13 s exposure and a total dose of 49.7 e Å⁻². Frame alignments were performed using *MotionCorr2* (Zheng *et al.*, 2017). CTF estimates were performed with *Gctf* (Zhang, 2016) and images with estimated resolutions lower than 5 Å were removed.

2.4. Data processing and refinement

Filaments were picked manually, and particles were extracted from filaments using a box size of 360 pixels at a pixel size of 1.15 Å per pixel and an interbox distance of 100 Å (He & Scheres, 2017). The resulting stack of particles was cleaned by 2D classification in *RELION* 2.1 (Kimanius *et al.*, 2016), where poorly aligned particles were excluded from subsequent processing. 3D refinement was initiated with a bead model produced using *relion_helix_toolbox* (He & Scheres, 2017) and helical parameters estimated from a previous reconstruction of the Ebola virus nucleocapsid (Bharat *et al.*, 2012; Wan *et al.*, 2017). The same helical parameters were used as starting helical parameters for 3D refinement and these parameters were allowed to optimize during refinement.

Table 1

Data-collection and coordinate model-refinement statistics.

Data and reconstruction statistics	
EMDB code	EMD-0522
Microscope	Talos Actica
Voltage (kV)	200
Detector	Gatan K2 Summit
Dose rate (e per pixel per second)	5.06
Exposure (s)	13
Dose (e Å ⁻²)	49.7
Frames	65
Defocus range (µm)	0.6–2.0
Filaments	2143
Particles	24609
<i>B</i> factor (Å ²)	–94
Resolution (Å)	3.1
Coordinate model refinement	
PDB code	6nut
No. of residues	
Amino acids	379
Nucleotides	6
R.m.s.d., bonds (Å)	0.016
R.m.s.d., angles (°)	1.35
Ramachandran statistics	
Favored (%)	99.2
Allowed (%)	0.8
Outliers (%)	0.0
Rotamer outliers (%)	0.3
Clashscore	2.89
<i>MolProbity</i> score	1.08
<i>EMRinger</i> score	3.15

Atomic coordinates were built into density accounting for the asymmetric unit of the helix using *Coot* (Emsley *et al.*, 2010) and a monomeric Ebola virus NP structure, PDB entry 4ztg (Kirchdoerfer *et al.*, 2015), as a starting model. The previously published 3.6 Å resolution model with PDB code 5z9w was used for comparison (Sugita *et al.*, 2018). The coordinates were refined against the map using *RosettaRelax* (DiMaio *et al.*, 2015). The final coordinates were evaluated using *MolProbity* (Williams *et al.*, 2018) and *EMRinger* (Barad *et al.*, 2015). Data-collection and processing as well as model-refinement statistics are presented in Table 1.

3. Results

The NP helical filaments have a barbed appearance in the aligned, dose-weighted micrograph movies (Fig. 1). 2D alignment of particles revealed classes with clear secondary-structure features. The refined 3D particle orientations demonstrate a predominance of views perpendicular to the helical axis, as expected given the orientations of the helical filaments picked in the raw micrographs (Fig. 2*a*). The refined helical parameters for the final map are –14.71° twist and 2.84 Å rise per NP protomer. The reconstructed map has a resolution of 3.1 Å estimated using a gold-standard Fourier shell correlation cutoff of 0.143. Local resolution estimation shows a resolution range of 3.0–3.8 Å (Fig. 2), with the folded core of the NP exhibiting the highest resolution, while peripheral, surface-exposed regions were reconstructed to lower resolution, possibly indicating local mobility or conformational pleiomorphism.

The high resolution of the map enabled the building and refinement of an atomic model of the Ebola virus NP bound

to RNA. Side-chain density for most amino acids is clearly evident and we observe density for six nucleotides of RNA, similar to previous studies (Sugita *et al.*, 2018; Fig. 3). Because NP is believed to bind to RNA without sequence specificity, the RNA nucleotides seen in the map represent the average of bound nucleotides. We modeled this density as a polymer of six adenosine residues since, as expected, none of these residues makes sequence-specific contacts.

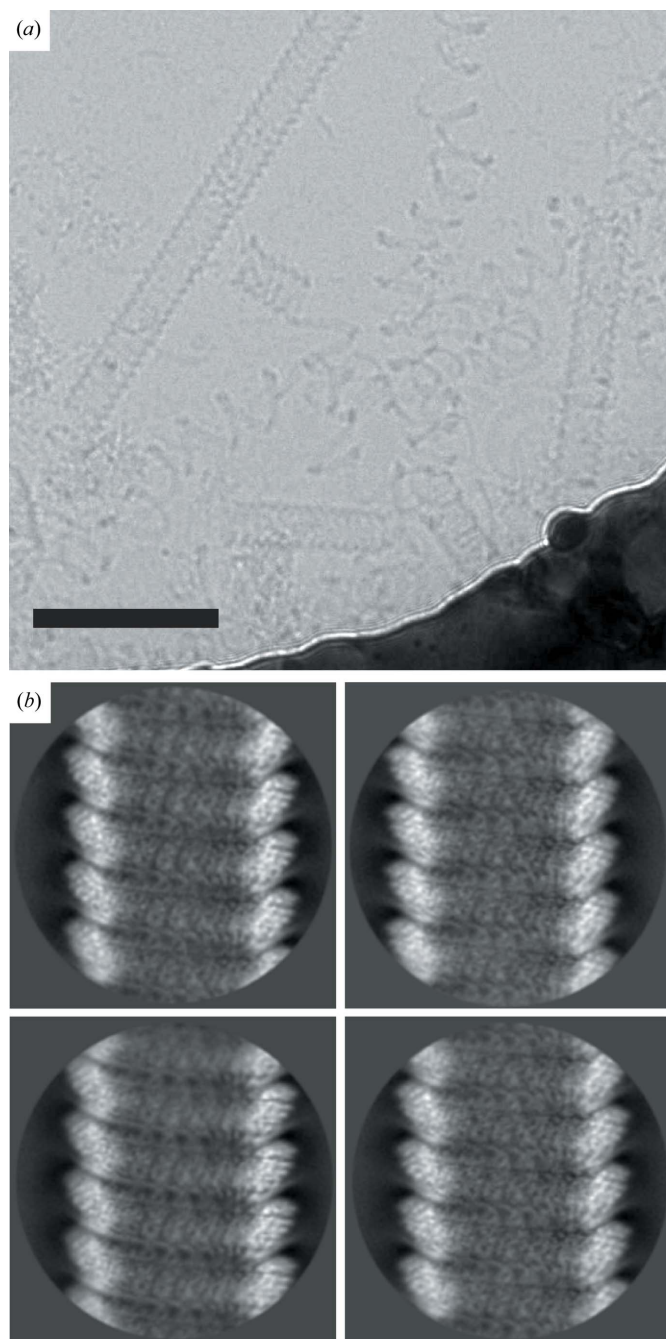
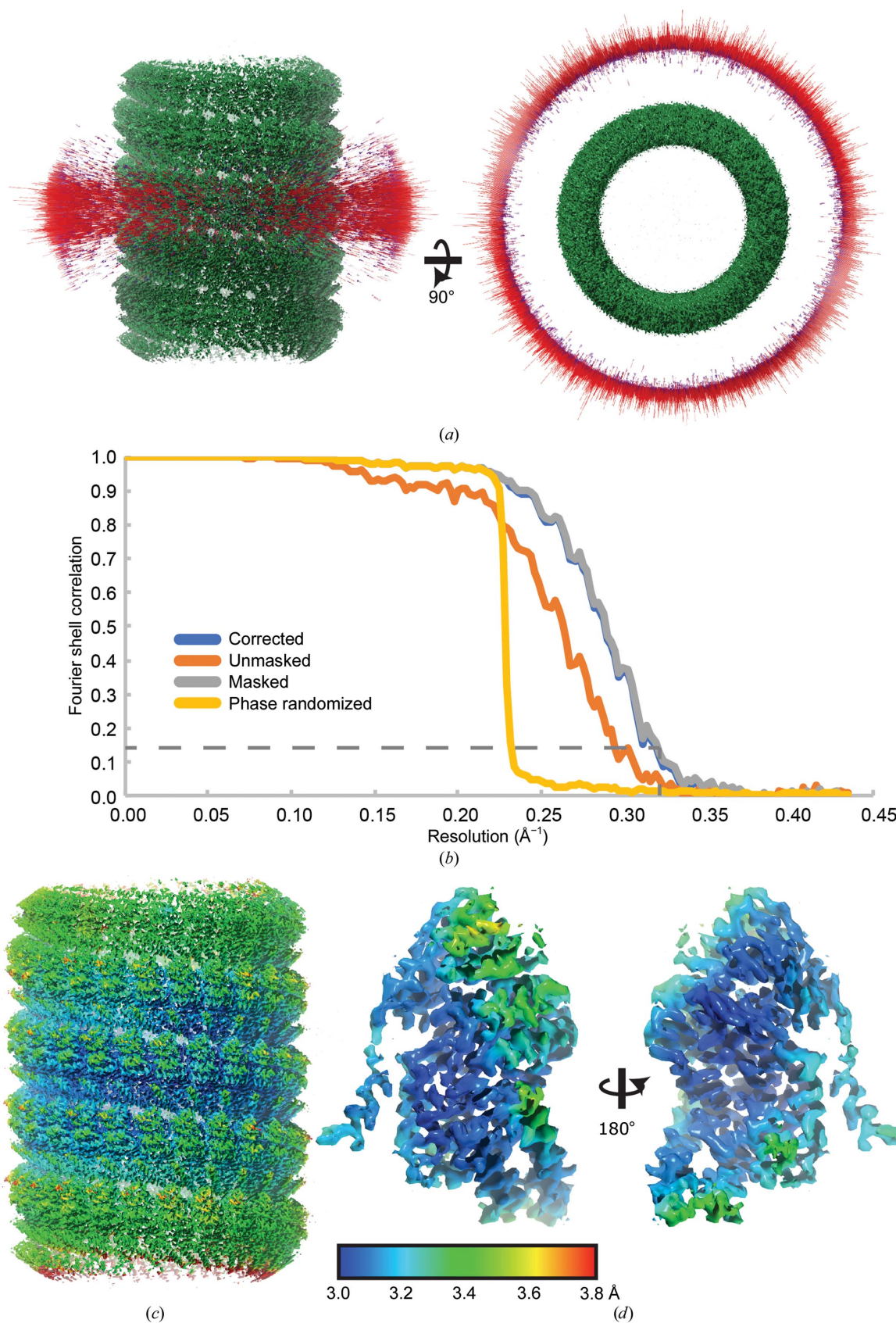


Figure 1 Image acquisition and processing by two-dimensional alignment. (a) Raw micrograph movies were aligned and dose weighted. The scale bar is 100 nm. (b) 2D alignment of picked particles shows the helical nature of the particles with evidence of secondary structure.


Figure 2

Particle angular distribution and map resolution. (a) The angular distribution of refined particle orientations indicates a predominance of views (red) perpendicular to the filament axis. (b) Fourier shell correlation plots indicated a resolution of 3.1 Å using a gold-standard FSC cutoff of 0.143 as indicated by the dashed line. The local resolution of the NP helical filament (c) and an isolated NP protomer (d) was estimated in *RELION* (Kucukelbir *et al.*, 2014) and indicates a high-resolution folded core of NP with lower-resolution peripheral regions.

The coordinate model presented here overlays well with the previously published 3.6 Å resolution structure of NP bound to RNA (Sugita *et al.*, 2018; Fig. 4). There does appear to be a 2.2% isotropic stretching between the previously published map and the map presented here. We attribute this to issues in the pixel-size calibration of the previously collected data. However, this effect is small and is unlikely to impact the conclusions drawn from the previous study. Future antiviral development efforts may benefit from making note of such calibration issues to produce more accurate models.

The interaction of NP with RNA is primarily mediated by electrostatic interactions, where Lys160, Arg174, Arg298, His310 and Arg401 interact directly with negatively charged phosphates of the RNA backbone (Fig. 5*a*). Additional residues, Val178, Leu245, Leu331 and Val334, create surfaces for the hydrophobic packing of the nucleotide bases. Turn-on-turn NP interactions appear to be minimal. Large gaps between turns of the helix (Fig. 5*b*) suggest flexible or conformationally heterogeneous side-chain interactions mediated by charged amino acids between turns, which have previously been hypothesized to impart flexibility to the NP filament (Sugita *et al.*, 2018). Oligomerization of NPs is mediated by side-to-side interactions as well as by N- and C-terminal regions flanking the folded core of the NP (Fig. 5*c*). Side-to-side NP interactions are mediated by a mixture of salt bridges and hydrophobic interactions (Fig. 5*d*). The N-terminal oligomerization region is a helix connected to the folded core of NP by a short linker. This helix buries Ile24 in a hydrophobic pocket on the adjacent NP and occupies the same site as the VP35 chaperoning peptide (Kirchdoerfer *et al.*, 2015; Leung *et al.*, 2015) (Fig. 5*e*). The C-terminal oligomerization region folds into a pair of helices which stack on the equivalent helices from the adjacent NP using primarily hydrophobic residues (Fig. 5*f*). The formation of these helices creates a platform onto which to place a hydrophobic surface of the C-terminal lobe of the folded NP core, supporting a model of oligomerization-induced NP conformational changes in the NP core to create the RNA-binding site.

4. Discussion

Similar to the nucleoproteins of other nonsegmented negative-sense RNA viruses, Ebola virus nucleoprotein begins its life

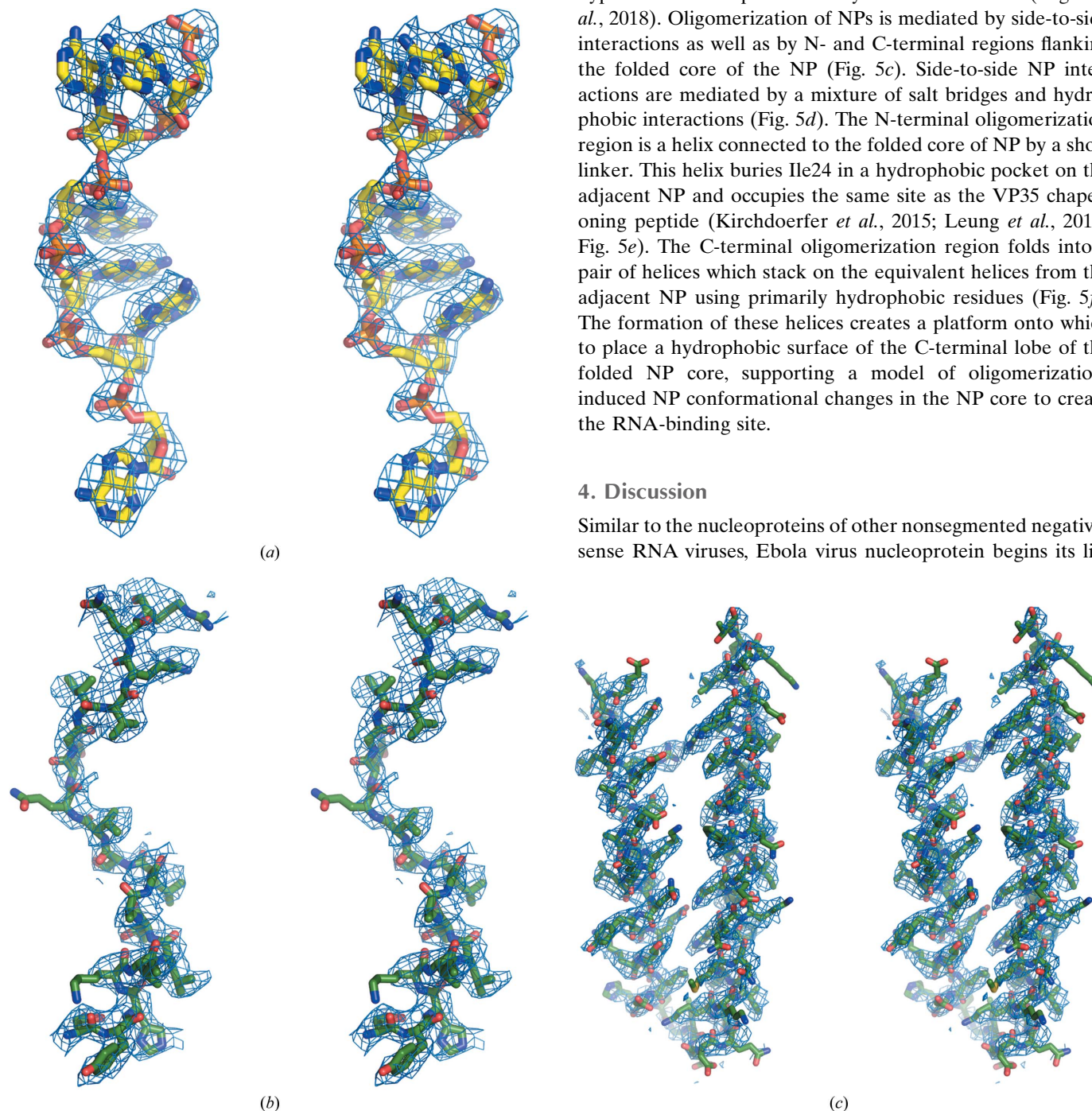


Figure 3 Fit of the model to the reconstructed density. Stereoviews of the NP–RNA helical filament focusing on (a) the bound RNA, (b) the N-terminal oligomerization arm and (c) the C-terminal oligomerization helices.

chaperoned by VP35 in a monomeric form, which then transitions to an oligomeric form bound to viral RNA. Here, we present a high-resolution cryo-EM structure of the oligomeric, RNA-bound Ebola virus nucleoprotein. The transition from the chaperoned monomer to the oligomeric, RNA-bound form involves the breakage and formation of extensive protein–protein interfaces which are coupled to conformational changes. Removal of the VP35 peptide from the chaperoned monomer is a prerequisite for these changes (Kirchdoerfer *et al.*, 2015), although how VP35 is removed *in vivo* remains unclear, and perhaps involves additional cellular or viral factors such as the viral polymerase L. After removal of VP35 from NP, the NP N-terminal oligomerization region and the lateral NP face of incoming monomers are able to

interact with adjacent NP. The subsequent folding of the C-terminal oligomerization helices creates a hydrophobic platform for binding incoming NP through their C-terminal lobes. This interaction of the C-terminal helices of NP_{*i*-1} with the C-terminal lobe of NP_{*i*} results in the movement of the C-terminal lobe of NP_{*i*} towards the N-terminal lobe, creating the RNA-binding site. The addition of monomeric NP to a growing oligomer in this manner would allow the encapsidation of viral RNA in a 5' to 3' direction, a mechanism which would lend itself to co-replicative encapsidation of viral RNA. One unknown related to this model is how this process is reversed to expose the viral RNA for RNA synthesis, as the nucleotide bases are inaccessible as templates for polymerase when encapsidated by NP.

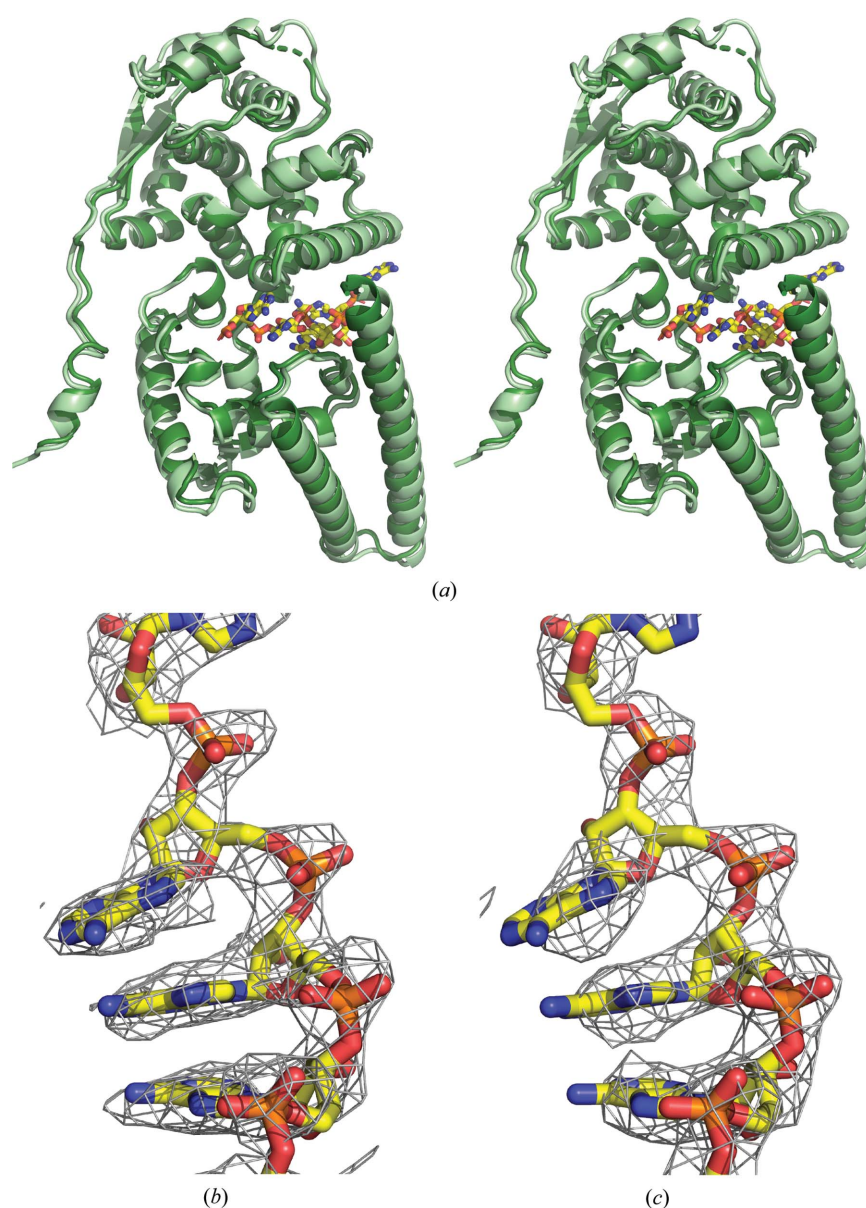


Figure 4

(a) Stereoview of the 3.1 Å resolution model presented here (dark green) superimposed with the previously published 3.6 Å resolution protein model (light green; PDB entry 5z9w; Sugita *et al.*, 2018). The bound RNA from the model presented here is shown as sticks. Comparison of the 3.1 Å resolution density presented here (b) with the 3.6 Å resolution density of Sugita *et al.* (2018) (c) shows a clearer definition of the RNA bases for the higher resolution data in (b).

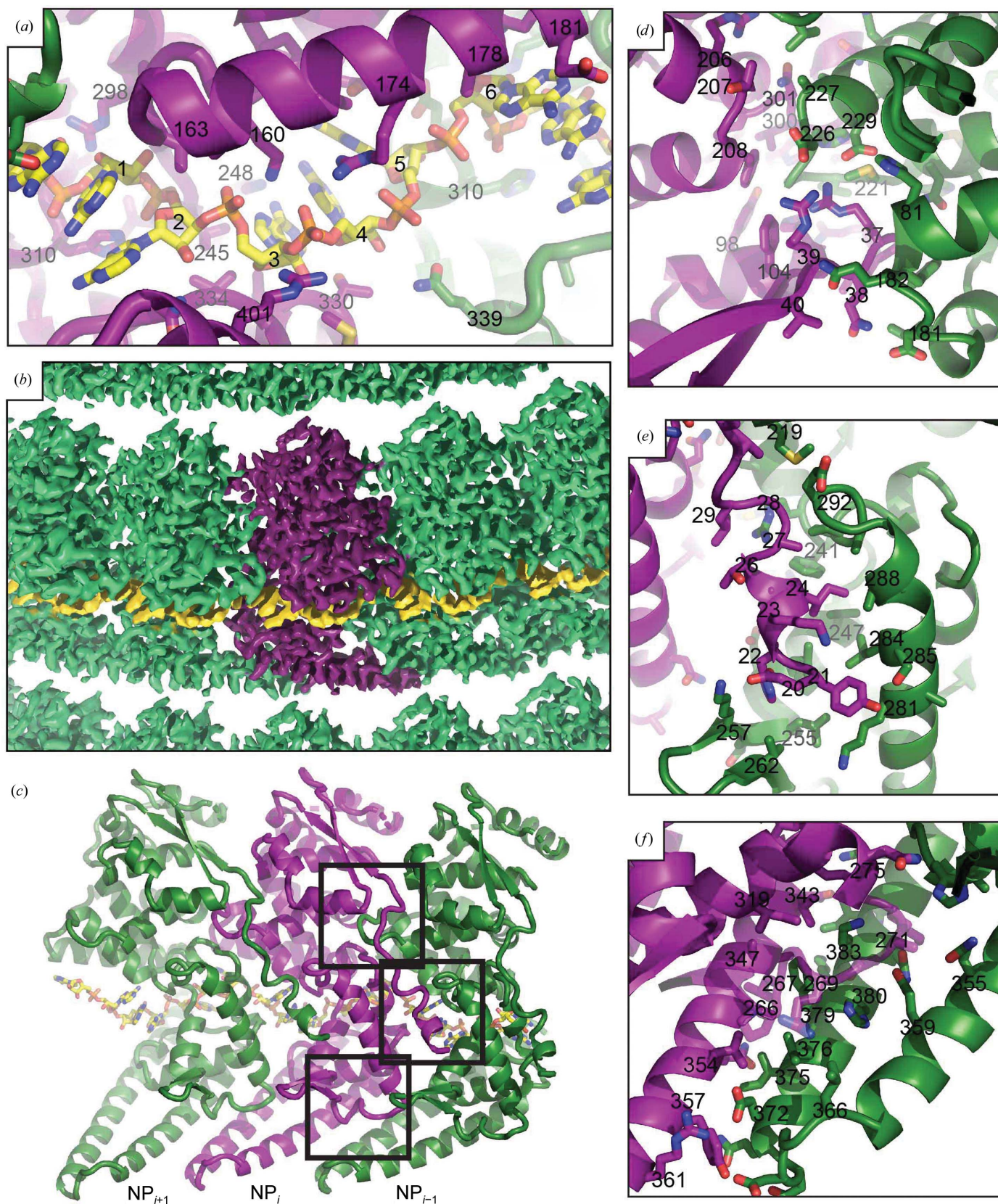


Figure 5
 Protein-protein interfaces involved in NP oligomerization and RNA binding. (a) RNA binds between the NP N- and C-terminal lobes of the folded core and is mediated by several basic amino acids as well as by the formation of hydrophobic surfaces for nucleotide base stacking. (b) Turn-on-turn interactions between NP are minimal and there are large gaps between helical turns. (c–f) NP protein-protein interactions are spread across side-to-side interactions (d), the N-terminal oligomerization region (e) and the C-terminal oligomerization region (f).

The 3.1 Å resolution structure presented here recapitulates the features observed in the previously published 3.6 Å resolution structure (Sugita *et al.*, 2018), including protein–protein and protein–RNA interactions. The improved resolution provides a clear delineation of protein side chains and nucleotide bases, clarifying the molecular contacts between NP protomers and the RNA. Ebola virus remains an emerging threat to world health security. Key to tackling this threat is the discovery and evaluation of novel therapeutics. The high-resolution structure of the Ebola virus NP bound to RNA will better enable the characterization of potential antiviral drugs targeting viral nucleocapsid assembly.

Acknowledgements

We gratefully acknowledge Mark Herzik and Bill Anderson for their help in microscope alignment and data collection. A special thanks to Sebastian Raemisch for his help in refining coordinate models using *RosettaRelax*. We also thank Charles Bowman for computational support.

Funding information

This work was funded by NIH/NIAID grants AI123498 to RNK and AI118016 to EOS and ABW.

References

- Barad, B. A., Echols, N., Wang, R. Y.-R., Cheng, Y., DiMaio, F., Adams, P. D. & Fraser, J. S. (2015). *Nature Methods*, **12**, 943–946.
- Bharat, T. A. M., Noda, T., Riches, J. D., Kraehling, V., Kolesnikova, L., Becker, S., Kawaoka, Y. & Briggs, J. A. G. (2012). *Proc. Natl Acad. Sci. USA*, **109**, 4275–4280.
- DiMaio, F., Song, Y., Li, X., Brunner, M. J., Xu, C., Conticello, V., Egelman, E., Marlovits, T., Cheng, Y. & Baker, D. (2015). *Nature Methods*, **12**, 361–365.
- Dong, S., Yang, P., Li, G., Liu, B., Wang, W., Liu, X., Xia, B., Yang, C., Lou, Z., Guo, Y. & Rao, Z. (2015). *Protein Cell*, **6**, 351–362.
- Ebola haemorrhagic fever in Zaire, 1976 (1978). *Bull. World Health Organ.* **56**, 271–293.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- He, S. & Scheres, S. H. W. (2017). *J. Struct. Biol.* **198**, 163–176.
- Hersey, S., Martel, L. D., Jambai, A., Keita, S., Yoti, Z., Meyer, E., Seeman, S., Bennett, S., Ratto, J., Morgan, O., Akyeampong, M. A., Sainvil, S., Worrell, M. C., Fitter, D. & Arnold, K. E. (2015). *MMWR Morb. Mortal. Wkly Rep.* **64**, 981–984.
- Herzik, M. A. Jr, Wu, M. & Lander, G. C. (2017). *Nature Methods*, **14**, 1075–1078.
- Kimanius, D., Forsberg, B. O., Scheres, S. H. W. & Lindahl, E. (2016). *Elife*, **5**, e18722.
- Kirchdoerfer, R. N., Abelson, D. M., Li, S., Wood, M. R. & Saphire, E. O. (2015). *Cell. Rep.* **12**, 140–149.
- Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. (2014). *Nature Methods*, **11**, 63–65.
- Leung, D. W., Borek, D., Luthra, P., Binning, J. M., Anantpadma, M., Liu, G., Harvey, I. B., Su, Z., Endlich-Frazier, A., Pan, J., Shabman, R. S., Chiu, W., Davey, R. A., Otwinowski, Z., Basler, C. F. & Amarasinghe, G. K. (2015). *Cell. Rep.* **11**, 376–389.
- Leyrat, C., Yabukarski, F., Tarbouriech, N., Ribeiro, E. A. Jr, Jensen, M. R., Blackledge, M., Ruigrok, R. W. & Jamin, M. (2011). *PLoS Pathog.* **7**, e1002248.
- Sugita, Y., Matsunami, H., Kawaoka, Y., Noda, T. & Wolf, M. (2018). *Nature (London)*, **563**, 137–140.
- Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., Stagg, S., Potter, C. S. & Carragher, B. (2005). *J. Struct. Biol.* **151**, 41–60.
- Wan, W., Kolesnikova, L., Clarke, M., Koehler, A., Noda, T., Becker, S. & Briggs, J. A. G. (2017). *Nature (London)*, **551**, 394–397.
- Watanabe, S., Noda, T. & Kawaoka, Y. (2006). *J. Virol.* **80**, 3743–3751.
- Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., Verma, V., Keedy, D. A., Hintze, B. J., Chen, V. B., Jain, S., Lewis, S. M., Arendall, W. B., Snoeyink, J., Adams, P. D., Lovell, S. C., Richardson, J. S. & Richardson, J. S. (2018). *Protein Sci.* **27**, 293–315.
- Zhang, K. (2016). *J. Struct. Biol.* **193**, 1–12.
- Zheng, S. Q., Palovcak, E., Armache, J.-P., Verba, K. A., Cheng, Y. & Agard, D. A. (2017). *Nature Methods*, **14**, 331–332.