Review

Structural perspective on the formation of ribonucleoprotein complex in negative-sense single-stranded RNA viruses

Honggang Zhou[1,2*], Yuna Sun[3*], Yu Guo[1], and Zhiyong Lou[2]

[1] State Key Laboratory of Medicinal Chemical Biology and College of Pharmacy, Nankai University, Tianjin 300071, China
[2] Laboratory of Structural Biology and Ministry of Education (MOE) Key Laboratory of Protein Sciences, School of Medicine, Tsinghua University, Beijing 100084, China
[3] National Laboratory of Macromolecules, Institute of Biophysics, Chinese Academy of Science, Beijing 100101, China

Negative-sense single-stranded RNA viruses (NSRVs) possess a ribonucleoprotein (RNP) complex composed of viral polymerase and genomic RNA surrounded by viral nucleoprotein. The RNP facilitates virus replication, transcription, and assembly. To date, a large body of structural work, through crystallography and electron microscopy (EM) analysis, has been performed to aid understanding the molecular mechanism of RNP formation in NSRVs, and provides great potential for the discovery of antiviral agents targeting viral RNP formation.

RNP complexes in NSRVs

NSRVs constitute a large group of viral pathogens responsible for causing severe infectious diseases in humans, animals, and plants. NSRVs are distinguished by their genome which consists of one or several antisense RNA segment(s), and are mainly divided into the order Mononegavirales and several viral families that are not assigned to an order [1] (Box 1).

To ensure that the NSRV genome will be immediately transcribed into mRNA by the viral RNA-dependent RNA polymerase (RdRp) upon entry into host cells, viral RdRp and genomic RNA must be included in the virions. Throughout the entire life cycle of NSRVs, the genome-length RNA is encapsidated by a virus-encoded nucleoprotein (NP) instead of a naked RNA, and is further associated with RdRp (or polymerase complex) to form a stable ribonucleoprotein (RNP) complex which is responsible for virus replication, transcription, and assembly [2]. During this process the NP can protect the RNA against exogenous nucleases or the innate immune system in the host cell. Therefore, knowledge of RNP formation and its function is essential to understand the mechanisms of the NSRV life cycle and to aid the discovery of new antiviral therapeutics.

Structural knowledge of NSRV RNPs was initiated 10 years ago by studying non-segmented NSRVs (nsNSRVs), including Borna disease virus (BDV), rabies virus, vesicular stomatitis virus (VSV), and respiratory syncytial virus (RSV), and this knowledge has been greatly enhanced by recent investigations of segmented NSRVs (sNSRVs), including arenaviruses [3–5], bunyaviruses [6–12], and influenza virus [13–17]. Not only did these structural achievements indicate how NP, genomic RNA, and RdRp form a highly ordered RNP, but they also revealed unexpected enzymatic functions of viral NPs. In particular, advances in the visualization of native or authentic RNP through EM, combined with the structures of RNP components at atomic resolution, led to an understanding of the dynamic processes of RNP formation at the molecular level. This review focuses on the structural aspects of NSRV RNP formation, particularly sNSRV RNP, and on current efforts in a new strategy to discover antivirals that target RNP formation.

RNP in nsNSRVs

Initial understanding of the mechanism by which NP encapsidates RNA was accomplished through the crystal structures of the NP–RNA complexes of VSV [18], rabies virus [19], and RSV [20] (Figure 1A). The recombinant NP–RNA complexes from VSV, rabies virus, and RSV form helical or heterogeneous ring-shaped structures with 9–15 NP protomers per ring, together with the bound RNA [18–20]. Although these NPs show similarities in their molecular folding, the most significant difference appears in the position of the bound RNA. The RNA-binding grooves of VSV, rabies virus, and RSV NPs are located at the interface of the N- and C-lobes, but the lateral NP contacts in the VSV and rabies virus NP–RNA complexes are positioned such that the curvature is opposite to that of the RSV ring. This condition results in an inside-out nucleocapsid ring, with the RNA inside and the NP molecule oriented outside-in [20].

Corresponding author: Lou, Z. (louzy@mail.tsinghua.edu.cn).
Keywords: negative-sense single-stranded RNA virus; ribonucleoprotein complex; structure; formation; antiviral drugs.
* These authors contributed equally.

0966-842X/$ – see front matter
 http://dx.doi.org/10.1016/j.tim.2013.07.006

Box 1. Classification of NSRVs

NSRVs are grouped into one order and three viral families that are not assigned to an order. Depending on the number of genome segments, these are classified as non-segmented NSRVs (nsNSRVs) or segmented NSRVs (sNSRVs), together with three genera that are not assigned to a family.

The order Mononegavirales, which is characterized by long non-segmented RNA genomes, consists of four families: the Rhabdoviridae family (e.g., rabies virus and VSV); Paramyxoviridae, with two subfamilies, Paramyxovirinae (e.g., measles and Sendai viruses) and Pneumovirinae (e.g., RSV); Bornaviridae (e.g., BDV); and Filoviridae (e.g., Marburg virus and Ebola virus).

By contrast, three viral families contain segmented RNA genomes and are grouped into sNSRVs. Among them, the genomes of Arenaviridae (e.g., LAFV) and Bunyaviridae (e.g., RVFV, CCHFV, and BUNV) have two [5] and three [10] segments, respectively, whereas Orthomyxoviridae (e.g., influenza virus) have six to eight genome segments [17].

Moreover, three viral genera in NSRV are not assigned into a viral family. The *Deltavirus* genus contains a circular negative-sense ssRNA genome, folding in native conditions into a nearly complementary rod-like structure [51]. The *Emaravirus* genus contains a four-segmented linear negative-sense ssRNA genome [52]. The *Tenuivirus* genus has a close phylogenic relationship with members of the *Tospovirus* genus in the Bunyaviridae family but has four or five genome segments [53].

In an exciting development, Ge *et al*. reported the structure of the trunk portion of the VSV virion by visualization through cryo-EM [21] (Figure 1B). This structure revealed that the VSV virion includes an outermost lipid bilayer, a middle layer composed of the matrix protein (M), and an interior of helically organized NP–RNA strings [21]. In the trunk portion, a complete repeat of the NP–RNA and M helices indicates that there are 37.5 units for each in two turns (Figure 1C), which is obviously larger than those observed in crystal structures. Moreover, the crystal structure of individual NPs can be fitted into this EM map, and the highest-density regions at the interface of the N- and C-lobes reflect the binding of viral genome [21] (Figure 1D). Furthermore, no density for RdRp was found in the EM map, which may be caused by reduced symmetric or asymmetric organization of RdRp in the virions [21]. In a recent result, Desfosses *et al*. further elucidated how VSV NP folds into conical tips and how further growth into the whole bullet-shaped virion takes place through adjusting pH and ionic strength conditions [22]. This work demonstrated that NP plays a dominant role in the formation of the bullet-shaped virion without the help of other viral components, including matrix protein.

RNP in sNSRVs

In the past 5 years, structural investigations of arenavirus, bunyavirus, and influenza virus have greatly advanced our understanding of RNP formation and the biological function of NPs in sNSRVs, and have revealed surprising features compared with nsNSRVs.

Formation of arenavirus RNP

The structure of the Lassa fever virus (LAFV) NP provided a new structural insight into the NSRV RNP [5]. LAFV is a prototypic member of the Arenaviridae family, which causes severe hemorrhagic fever in humans with high morbidity and mortality. Similarly to other arenaviruses, the genome of LAFV consists of two segments encapsidated by NP throughout the viral life cycle [23,24]. Qi *et al*. first reported the structure of full-length LAFV NP and revealed distinct N- (NTD) and C-terminal (CTD) domains, in contrast to previously reported structures of VSV, RSV, rabies virus, Rift Valley fever virus (RVFV), and influenza virus NPs that do not contain separate N- and C-terminal domains [5] (Figure 2A). Although the RNA-binding site was not experimentally confirmed in this study, a highly positively charged groove located at the interface of the NTD and the CTD was predicted to be the genomic RNA-binding site based on the trimeric structure of LAFV NP, which is consistent with a small-angle X-ray scattering analysis [25] (Figure 2B). Meanwhile, Hastie *et al*. solved the structures of NTD of LAFV NP in complex with ssRNA, and revealed that ssRNA binds in a deep crevice located entirely within the NTD, under the control of a gating mechanism [4] (Figure 2C,D). In this model, the trimeric NP proposed by Qi *et al*. may represent an initial formation of NP polypeptide after translation [4] (Figure 2B). During RNP formation, the CTD moves slightly away from its position in the RNA-free trimer, allowing helices α5 and α6 to open away from the RNA-binding crevice. Here, the monomer-sized NP–RNA complex would form a continuous head-to-tail polymer and thus create the filamentous RNP in the virion [4] (Figure 2D). These studies suggest that LAFV possesses a novel mechanism of RNP formation.

Formation of bunyavirus RNP

The Bunyaviridae family is the largest NSRV family and is composed of over 350 members that cause severe infectious diseases. The Bunyaviridae family consists of five genera, in other words *Hantavirus* (e.g., Hantaan virus), *Nairovirus* [e.g., Crimean-Congo hemorrhagic fever virus (CCHFV)], *Orthobunyavirus* [e.g., Bunyamwera virus (BUNV)], *Phlebovirus* (e.g., RVFV), and *Tospovirus* (e.g., tomato spotted wilt virus). All Bunyaviridae have typical tripartite genomes (L, M, and S segments), which are encapsidated in the form of individual RNPs encoding RdRp (L), the glycoproteins Gc and Gn, and NP. Although all Bunyaviridae share high genomic similarity, their RNPs have significant differences, either in the primary sequence or in their biological functions. For example, CCHFV NP has the highest molecular weight, at 52 kDa, and Hantaan virus has a smaller NP with a molecular weight of 45 kDa, whereas the molecular weights of NPs from the other three genera range from 26 kDa to 30 kDa (summarized in Figure S1 of [10]).

The first structural determination of a bunyavirus-encoded NP was obtained using phlebovirus NP. RVFV is a prototypic member of the *Phlebovirus* genus and is the causative agent of Rift Valley fever. The first structure of RVFV NP was solved in monomeric form using a denature/refold method [12]. The monomeric RVFV NP is a novel compact structure that lacks a positively charged crevice for RNA binding and has no protruding portions for NP oligomerization [12] (Figure 3A). Ferron *et al*. used a different purification method to solve the hexameric structure of RVFV NP, which has a highly positively charged inner perimeter as the RNA-binding site (Figure 3B) [7].

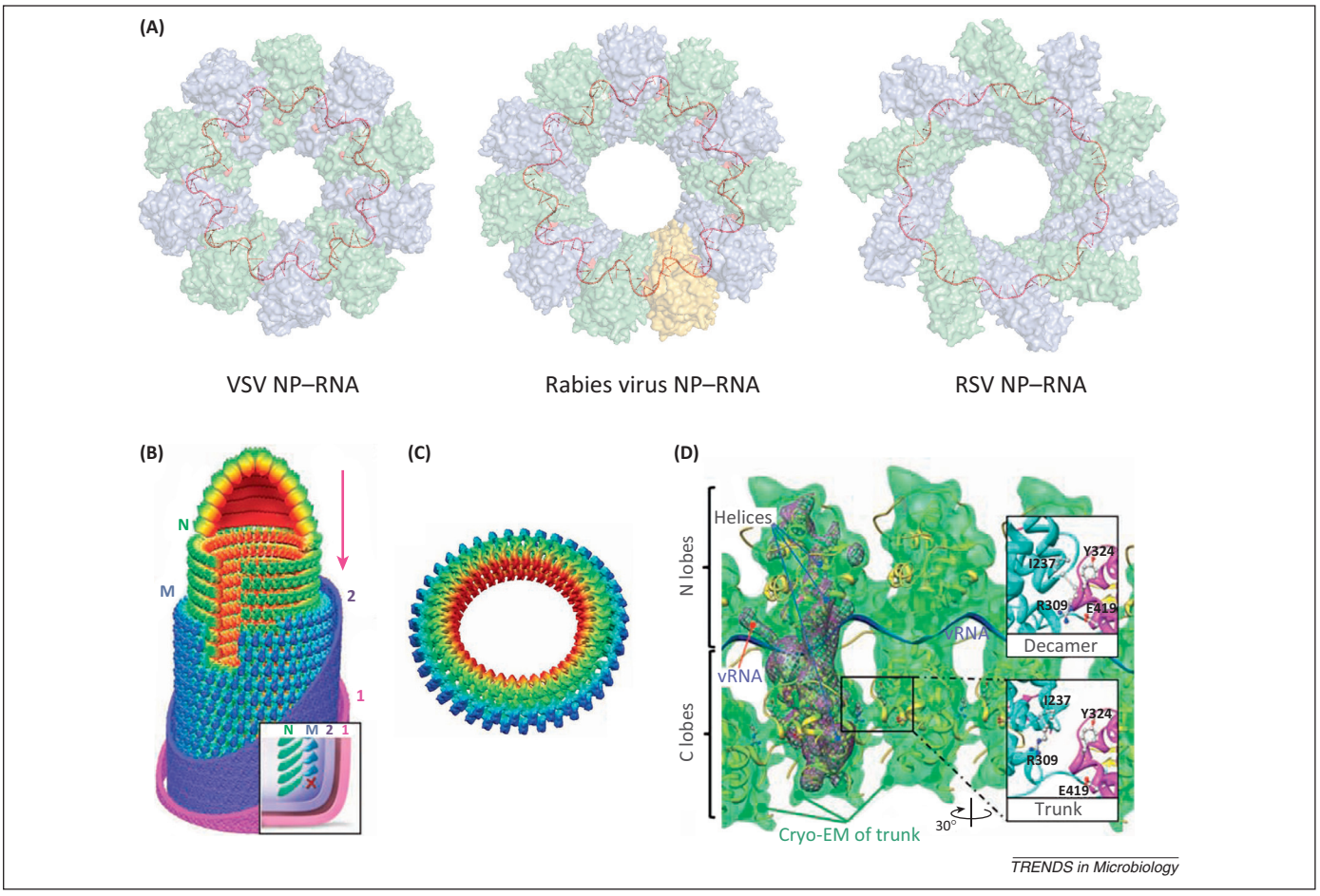

Figure 1. Ribonucleoprotein (RNP) formation in non-segmented negative-sense ss RNA viruses (NSRVs). **(A)** Crystal structures of the nucleoprotein (NP)–RNA complexes from vesicular stomatitis virus (VSV) (PDB code: 2GIC), rabies virus (PDB code: 2GTT), and respiratory syncytial virus (RSV) (PDB code: 2WJ8) are aligned and displayed in the same orientation. Because 11 NP protomers are in one NP–RNA ring-shaped complex of the rabies virus, one additional protomer is covered by a yellow surface, whereas the other protomers are covered with blue and green surfaces. **(B)** A model of the VSV virion based on the cryo-EM study. NP and M (matrix protein) are colored green and blue, respectively, whereas the inner and outer leaflets of the membrane are purple and pink, respectively. The enlarged figure shows the arrangement of NP and M in virions. The 'X' marks the absence of M below the lowest turn of the N helix. **(C)** A complete repeat of the NP–RNA and M in one turn. **(D)** Fitting of the crystal structure of the VSV NP–RNA complex into the cryo-EM density from the virion trunk. The purple wire frames represent the highest-density regions of the cryo-EM structure, which colocalize with a helix and the genomic RNA of the crystal structure. The insets indicate the change in the distances between the key residues for NP–NP interactions after flexible docking of the structure from the decamer into the cryo-EM density map of the trunk; (B–D) are reproduced/adapted with permission [21].

Although the body regions of the two structures are identical, there is significant conformational difference at the N-terminal arm (N-arm), which consists of the first 30 amino acids (Figure 3C). In the monomeric structure, the N-arm interacts with the body region of the same molecule and shadows the RNA-binding groove [12]. In sharp contrast, the N-arm in the hexameric structure extends out and interacts with an adjacent protomer to form a ring-shaped hexamer, allowing the RNA-binding site to be exposed. The distinct positions of the N-arm reflect structural flexibility during RNP formation, in which the monomeric structure may represent a 'waiting' conformation before oligomerization and binding with RNA [7]. Subsequently, Raymond *et al*. reported the structures of the NP–RNA complexes of the RVFV in tetrameric, pentameric, and hexameric forms, and a hexameric form for NP of Toscana viruses and other members of the *Phlebovirus* genus (Figure 3D) [26]. These structures confirmed that contacts between NP protomers, responsible for RNP formation, are mediated by the highly flexible N-arm (Figure 3E) [26]. Although different NP–RNA oligomers were found in crystal structures and solution conditions, EM visualization suggests that the monomeric NP–RNA complex most likely matches the size of the repeating unit in viral RNP (Figure 3F). This monomeric building block and the flexibility of the NP–NP interaction in the oligomer formation allow RVFV RNP to pack into virus particles with higher structures and density [26]. Interestingly, the RNA-binding behavior of RVFV NP differs significantly from other nsNSRVs, in which RVFV NP sequesters four RNA bases in a narrow hydrophobic binding slot, and all RNA bases face the NP protein [26]. Furthermore, two studies reported the crystal structures of NP of thrombocytopenia syndrome virus (SFTSV), another member of the *Phlebovirus* genus that causes newly emerging infectious diseases in China [27]. They also verified a conserved mechanism of NP multimerization within the *Phlebovirus* genus [9,28].

Following this structural research on phlebovirus, the structure of CCHFV (*Nariovirus* genus) NP, the largest NP encoded by a Bunyaviridae family member, indicated a novel mechanistic feature for RNP formation. The

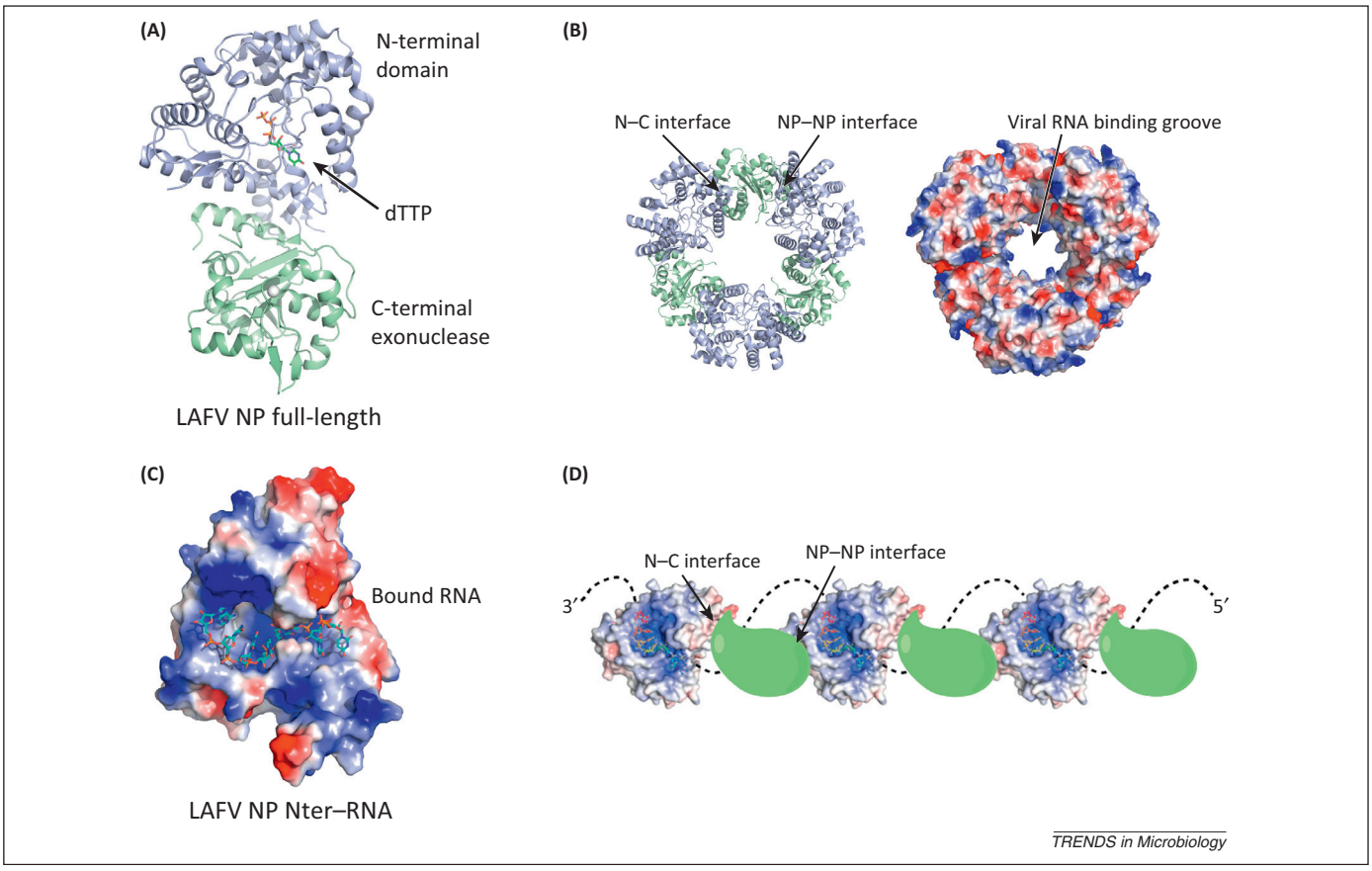

Figure 2. Structural implications for Lassa fever virus (LAFV) RNP formation. **(A)** The crystal structure of full-length LAFV NP (PDB code: 3MX2) is shown as a colored cartoon diagram in the left panel, in which the N-terminal domain and C-terminal exonuclease domain are colored blue and green, respectively. The bound dTTP is represented as a colored stick. **(B)** The structure of the RNA-free trimer of LAFV NP (PDB code: 3MWP) in a cartoon representation and a potential surface view (positive and negative charges are colored blue and red with limits $\pm 20\ k_bT/e_c$). N- and C-terminal domains are colored blue and green, respectively. **(C)** The crystal structure of the N-terminal domain of LAFV NP (LAFV NP Nter-RNA) in complex with RNA (PDB code: 3T5Q). The polypeptide of the N-terminal domain of LAFV NP is covered with a potential surface, whereas the bound RNA molecule is shown as colored sticks. **(D)** A gate-mechanism-controlled model for the LAFV RNP formation is shown, in which the C-terminal domain is shown as a cartoon model; (D) is reproduced with permission [4].

recombinant CCHFV (strain YL04057) NP exists in monomeric form and possesses a racket-shaped structure with head and stalk domains (Figure 4A) [8]. Although the most interesting finding in CCHFV NP is its unexpected nuclease activity, a positively charged groove clamped by the head and stalk domains has been suggested to be responsible for RNA binding [8]. Subsequently, Carter *et al*. reported the structure of NP from CCHFV strain Baghdad-12, and revealed a significant transposition of the stalk domain through rotation of 180° and translation of 40 Å, suggesting structural flexibility to switch between alternative NP conformations during RNA binding or oligomerization (Figure 4B) [29]. Furthermore, Wang *et al*. studied the structure of CCHFV (strain IbAr10200) NP pretreated with ssRNA and revealed an oligomeric form composed of double antiparallel superhelices with a similar conformation shift on the stalk domain, which was identical to that reported by Carter *et al*. (Figure 4B) [30]. Based on these models, a head-to-tail interaction of the stalk domain with the base of the head domain of the adjacent subunit is proposed to stabilize the helical organization of the oligomeric form of CCHFV NP during RNP formation (Figure 4C). In addition, the stalk domain tends to adopt a large shift during its transition from monomer to oligomer in the RNP (Figure 4D) [30].

Recent progress was achieved toward understanding RNP formation in the *Orthobunyavirus* genus, which encodes the smallest NP within the Bunyaviridae family [10]. Several groups analyzed the structures of BUNV [10], Leanyer virus (LEAV) [11], La Crosse orthobunyavirus (LACV) [31] NP–RNA complexes, Schmallenberg virus (SBV) NP alone [6], and the SBV NP–RNA complex [32]. All these structures show that orthobunyaviruses possess a highly conserved NP structure with N- and C-terminal extensions for inter-protomer interaction, in addition to N- and C-lobes for RNA binding (Figure 5A). However, the oligomerization and RNA-binding features of orthobunyaviruses showed variations. SBV NP formed a tetrameric structure in native conditions, and a hexamer structure following denaturation and refolding [6]. Although both N- and C-termini were demonstrated to be crucial for viral replication, the C-terminus of SBV NP is free in the native tetramer and is proposed to rotate when binding with RNA to form the RNP, which is different from other reported orthobunyaviral NPs [6] (Figure 5A). Interestingly, the N-terminal extension in the tetrameric structure folds backwards to reach the RNA-binding groove that is clamped by N- and C-lobes, while extending out in the hexameric form (Figure 5A). This structural variation is similar to the flexibility of the N-arm of RVFV NP and is proposed to

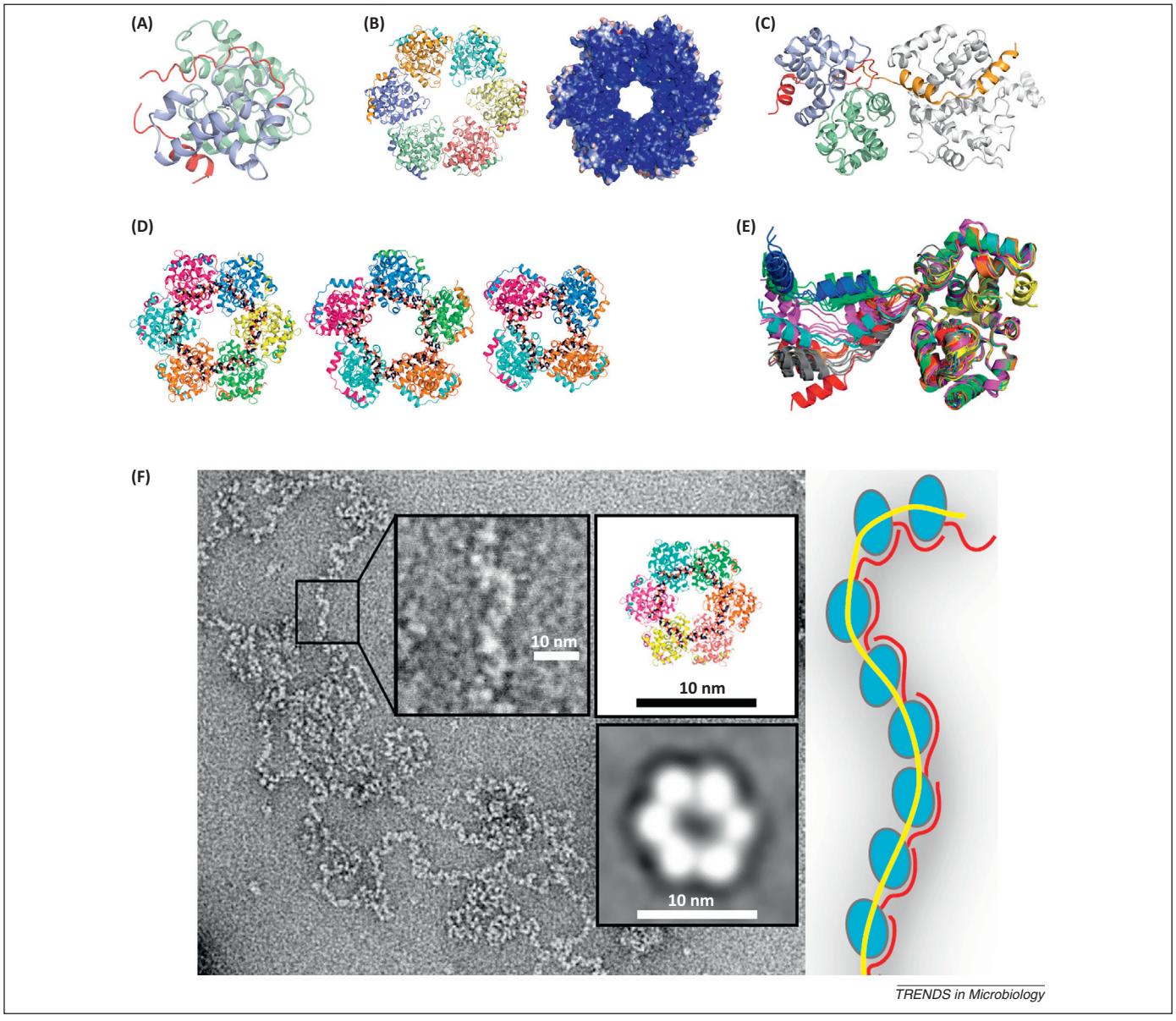

Figure 3. Structural implications for Rift Valley fever virus (RVFV) RNP formation. **(A)** Crystal structures of the monomeric RVFV NP achieved by denaturation/refolding (PDB code: 3LYF). The N-arm, N-lobe, and C-lobe are colored red, blue, and green, respectively. **(B)** Structure of the hexameric ring of the recombinant RVFV NP obtained in native purification conditions shown in a cartoon diagram and a potential surface image (PDB code: 3OU9). **(C)** Comparison of RVFV NP structure in the monomeric and hexameric states. The N- and C-lobes of one protomer are colored blue and green, whereas an adjacent protomer in the hexameric ring is shown as a white cartoon. The N-arms of the NP monomer and one protomer in the hexameric form are colored red and gold, respectively. **(D)** The structures of ring-shaped RVFV NP–RNA complexes in hexameric, pentameric, and tetrameric forms. **(E)** Alignment of RVFV NP in the NP6–RNA35 complex (PDB code: 4H6F), RVFV NP in the NP5–RNA35 complex (PDB code: 4H5O), RVFV NP in the NP4–RNA28 complex (PDB code: 4H5P), RVFV NP in the NP6–RNA30 complex (PDB code: 4H5Q), and a hexameric RVFV NP ring (PDB code: 3OV9). The corresponding colors of the N-arms are orange, green, blue, grey, cyan, yellow, magenta, and red. **(F)** Negative-stain EM visualization of the RVFV RNP isolated from infected cells with enlarged images to compare with the hexameric structure of RVFV NP (left). The monomer-sized NP–RNA complex is likely to be the building block of RVFV RNP (right); (D–F) are reproduced/adapted with permission [26].

reflect the conformational shift during RNP formation [6]. The structures of BUNV and LEAV, as well as SBV, NP–RNA complexes revealed that both N- and C-terminal extensions contribute to NP oligomerization in RNP formation through a head-to-head mode, in which the N-terminal extension interacts with the N-lobe of an adjacent NP protomer in a counter-clockwise direction, and the C-terminal extension contacts the C-lobe of a neighboring NP protomer in a clockwise direction [10,11,32]. The high structural similarity of these structures revealed a conserved mechanism for RNP formation. In addition, EM visualization of the recombinant BUNV NP–RNA complex suggested the coexistence of tetramer, pentamer, and hexamer in solution conditions, but the monomer-sized NP–RNA complex is likely to be the building block of native RNP [10] (Figure 5B,C). Furthermore, Reguera *et al.* reported the monomeric, tetrameric, and filamentous states of LACV NP alone or in complex with RNA [31]. In the tetrameric structure of LACV NP–RNA complex, the N- and C-terminal extensions of the two adjacent protomers fold back or reorient toward the core, but extend away from the core in the other two protomers, which significantly differs from their positions in LACV NP monomer [31] (Figure 5A). The authors further solved the structure

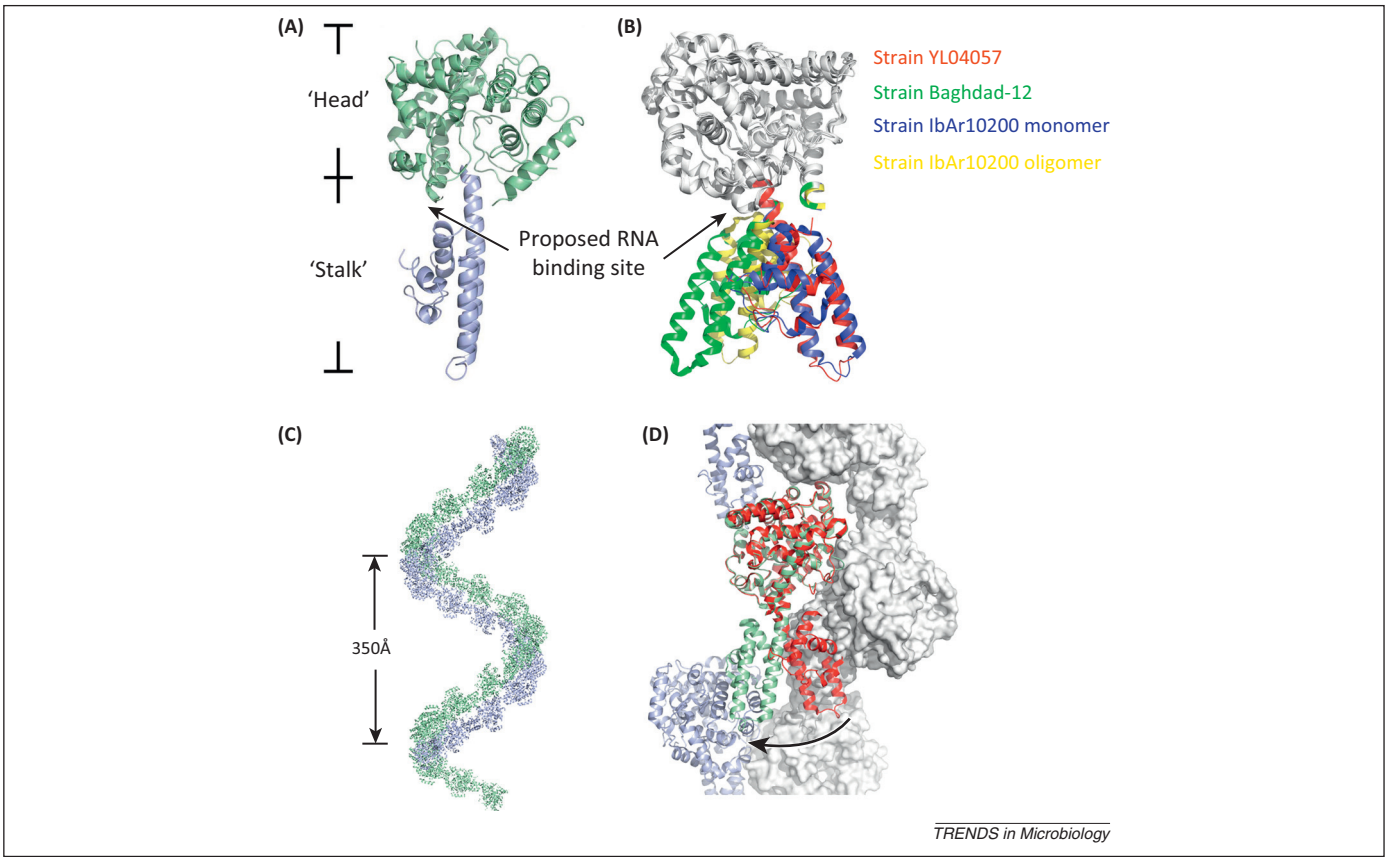

Figure 4. Structural implications for Crimean-Congo hemorrhagic fever virus (CCHFV) RNP formation. **(A)** The crystal structure of full-length CCHFV NP from strain YL04057 is shown as a cartoon diagram. The head and stalk domains are colored green and blue, respectively. **(B)** Comparison of the four different CCHFV NP structures. The head domains are colored white, whereas the stalk domains of CCHFV strain YL04057, strain Baghdad-12, strain IbAr10200 (monomeric form), and strain IbAr10200 (oligomeric form) NP are colored red, green, blue, and yellow, respectively. **(C)** The antiparallel double superhelix polymer of CCHFV NP, which was obtained using the CCHFV (strain IbAr10200) NP. **(D)** The conformational shift of the stalk domains reflects the dynamic process of CCHFV NP from the monomeric state to the oligomerized RNP. The CCFHV NP molecules in oligomeric and monomeric forms are shown in green and red, respectively. Other NP molecules in the superhelical string are shown as blue cartoons or as a white surface. The arrow indicates the movement of the stalk domain from monomeric form to oligomer form. The PDB codes for the structures used in this figure are: CCHFV strain YL04057, 3U3I; strain Baghdad-12, 4AKL; strain IbAr10200 (monomeric form), 4AQG; and strain IbAr10200 (oligomeric form), 4AQF.

of LACV NP in the $P4_1$ space group. In this structure, LACV NP forms a filamentous structure along the crystallographic 4_1 axis, in which both N- and C-terminal extensions have flexibilities similar to those observed in the tetramers (Figure 5D). This structure fits well with the highly ordered helical structure seen in the EM visualization of the native RNP structure (Figure 5E). In a recent result, Ariza *et al*. reported a similar structure for BUNV NP–RNA complex and presented an EM visualization of the native RNP, showing the helical architecture made of the tetrameric NP–RNA building block [33]. It is likely that the difference in purification methods for the native RNP may lead to variations in EM performance. Collectively, all these structures confirmed that the dynamic N- and C-terminal extensions are responsible for orthobunyaviral RNP formation.

Formation of influenza virus RNP

Unlike the large RdRp encoded by the other NSRVs, influenza virus encodes three polymerase proteins, PA, PB1, and PB2, that form its polymerase complex [34,35] and its RNP (with NP and genomic RNA). Therefore, structural investigations of influenza virus RNP formation focus on the formation of NP–RNA and polymerase complexes (reviewed in [36,37]) as well as on the architecture of RNP.

There was a breakthrough in the understanding of the structural details of influenza virus NP when two independent groups reported the crystal structure of the RNA-free NP from influenza A virus [13,14]. Both structures show the crescent-shaped influenza A virus NP with head, body, and tail portions, and a deep positively charged RNA-binding groove clamped by the head and the body regions, whereas the tail loop mediates oligomerization by interacting with the neighboring molecule (Figure 6B). The tail loop enables the NP to form a homotrimer with the RNA-binding groove exposed to the exterior of the NP oligomer [13,14]. However, the higher-order structure of the influenza virus NP has a distinct difference. Ng *et al*. reported the structure of influenza B virus NP (Figure 6C) and indicated a homotetrameric architecture (Figure 6E) instead of the homotrimer structure formed by the influenza A virus NP (Figure 6D), although both NPs show high primary sequence and structural similarities with only a few conformational shifts in their extended portions [38]. This result indicates structural reorganization of the helix–loop–helix motif and maintenance of some conserved hydrophobic contacts in higher-order NP oligomer

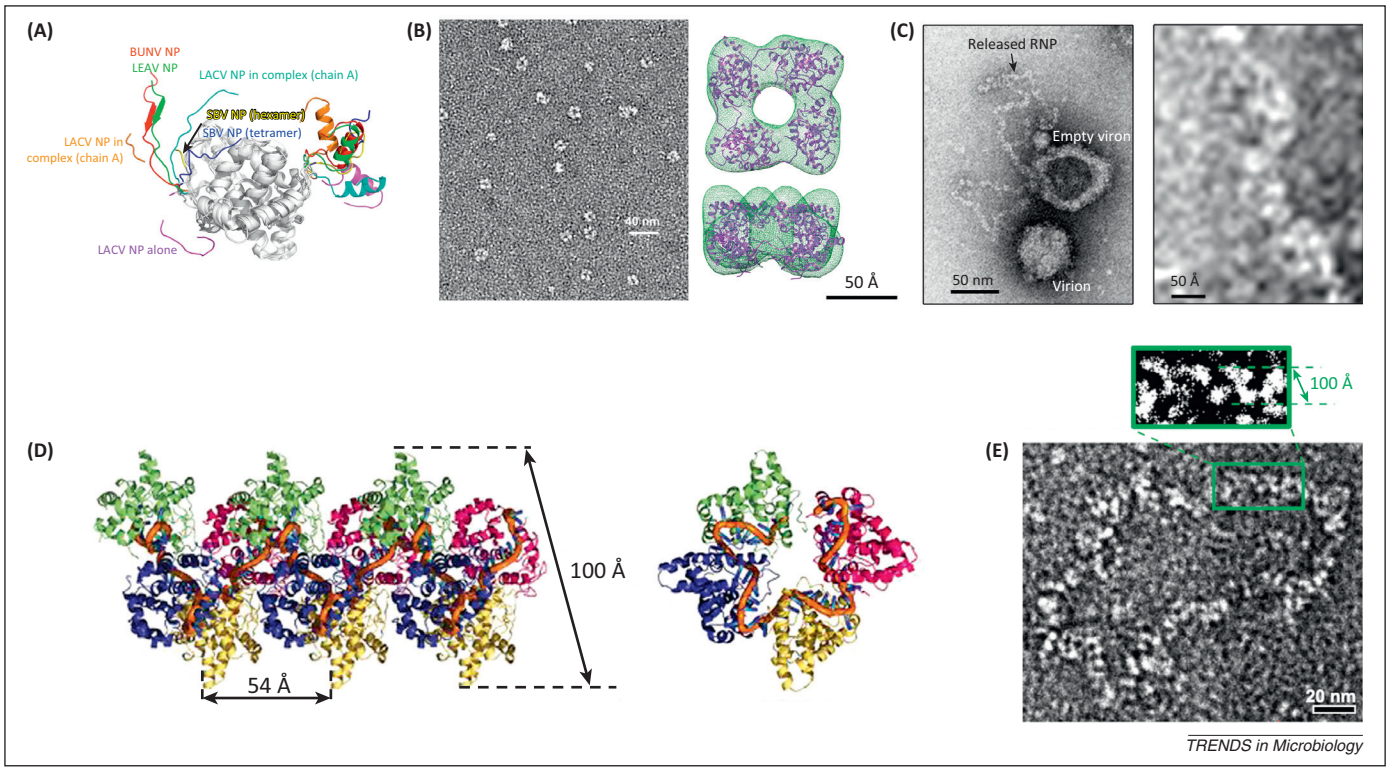

Figure 5. Structural implications for orthobunyavirus RNP formation. **(A)** Structural comparison of the reported structures of the orthobunyavirus NPs. The structures of Schmallenberg virus (SBV) NP in tetrameric form (PDB code: 4IDU), SBV NP in hexameric form (PDB code: 4IDX), Bunyamwera virus (BUNV) NP (PDB code: 4IJS), Leanyer virus (LEAV) NP (PDB code: 4J1G), La Crosse orthobunyavirus (LACV) NP in monomeric form (PDB code: 4BGP), and chain A and chain B of LACV NP in tetrameric form (PDB code: 4BHH) are colored blue, yellow, red, green, magenta, cyan, and orange, respectively. **(B)** EM image of the negatively stained recombinant BUNV NP–RNA complex (left) and docking of the crystal structure of tetrameric complex into the EM density map (right). **(C)** Negatively stained EM image of the native BUNV RNP extracted from virions. **(D)** A filamentous pseudo RNP model that was built based on the LACV NP–RNA complex in the $P4_1$ space group from a side view (left) and a top view (right). **(E)** Negatively stained EM image of the native LACV RNP extracted from virions. The RNPs are flexible but, in some areas, display regions with apparent helical characteristics as highlighted in green and shown in close-up; (D,E) are reproduced with permission [31].

formation [38]. Moreover, Chenavas *et al*. solved the monomeric structure of influenza NP using an obligate monomeric R416A mutant [16] (Figure 6A). In this monomeric structure, the folding of residues 386–401, the domain exchange loop (residues 402–428), and residues 429–498, present a significant shift compared to that in the oligomer. Interestingly, the regions of residues 386–390 and 392–407 suggest that the RNA-binding surface has a lower positive charge in this mutant monomer. These data explain the mechanism of influenza virus NP polymerization through a conformation change in the domain exchange loop.

Two recent EM analyses visualized authentic influenza virus RNPs which were extracted from native virions or produced in expression cells [15,17] (Figure 6F,G). Moeller *et al*. presented a structure at ~20 Å resolution using RNPs reconstructed in expression cells [17]. This model indicated that the RNP adopts a double-helical structure with two antiparallel strands leading to and away from the polymerase, which is located at one end of the RNP. The double-helical stem region has a rise of 32.6 Å between two neighboring NPs with 4.9 NP molecules per turn [17]. Arranz *et al*. presented a reconstruction of the native RNPs from influenza virions [15]. This structure showed that the rise step per monomer was constant at 28.4 Å and the rotation angle between monomers ranged from –57° to –64°, forming helices with approximately 12 NP monomers per turn [15]. In both RNP reconstructions, the monomer-sized NP is likely to arrange in twisted and antiparallel strands to form the RNPs. However, these two structures have significant variations in helical parameters and NP orientations. In particular, Arranz *et al*. presented a left-handed helix, whereas Moeller *et al*. reported a right-handed helix. As a result, Arranz *et al*. proposed that the body domain of NP mainly facilitates the inter-protomer interaction, whereas Moeller *et al*. indicated that the head domain of NP is likely to stabilize the RNP. This difference may have been caused by the different methods used to obtain the purified RNPs. Although the exact molecular architecture of NP in RNP remains controversial, both structures revealed that the polymerase complex is located at the open end of the RNP hairpin, thereby encircling the entire RNP.

Function of viral NP beyond RNA encapsidation

The structural role of viral NP in RNP formation, wherein the viral NP encapsidates genomic RNA and mediates RdRp–RNA interactions, is generally accepted. However, a few recent studies have extended knowledge of the function of virus-encoded NPs beyond a structural role.

The structure of LAFV NP provided, for the first time, structural evidence that a virus-encoded NP can have enzymatic activity [5]. In LAFV NP, the C-terminal domain has high structural similarity to the 3′–5′ exonuclease/exoribonuclease superfamily. Further analysis confirmed its metal-dependent exoribonuclease activity, and demonstrated that this activity is responsible for suppressing host immune

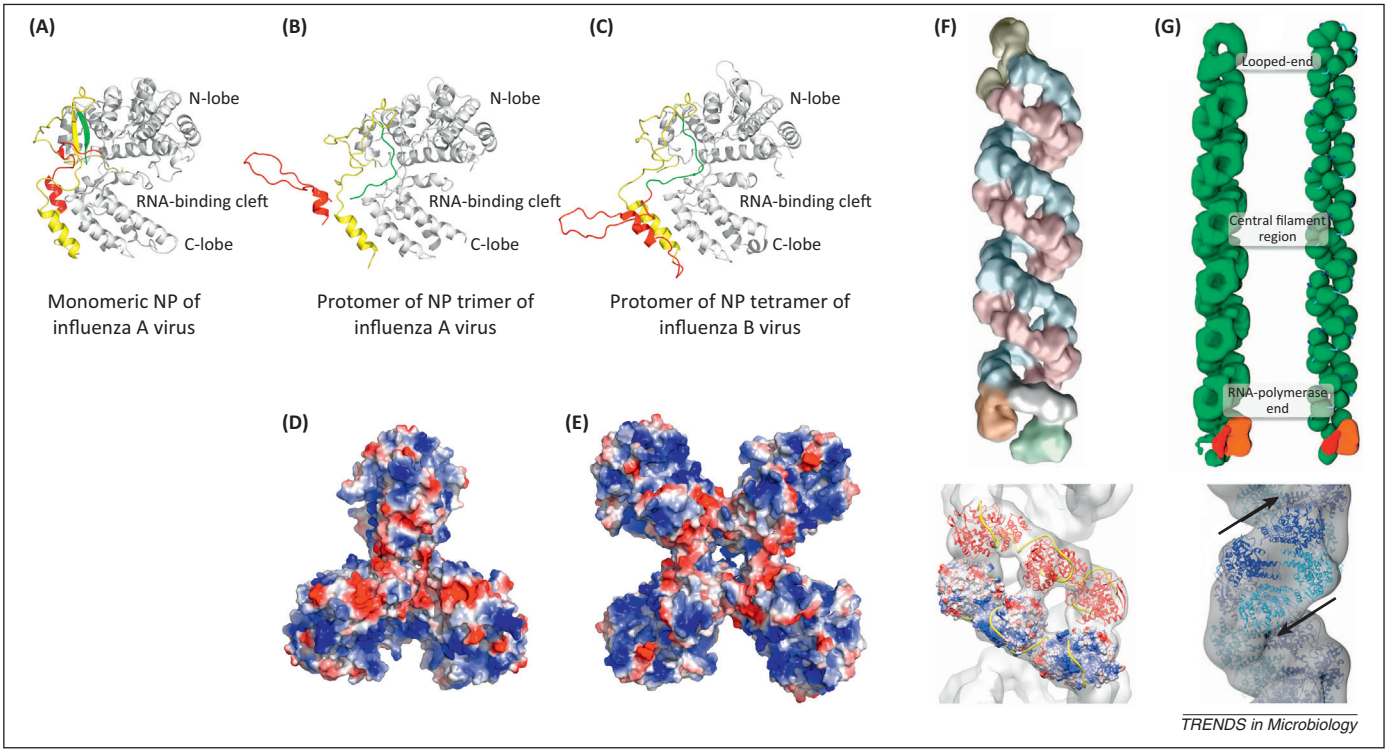

Figure 6. Molecular architecture of influenza virus RNP. The crystal structures of the NP of the influenza A virus in monomeric form (PDB code: 3ZDP) **(A)** and trimeric form (PDB code: 2IQH) **(B)** and the protomer of influenza B virus NP tetramer (PDB code: 3TJ0) **(C)** are aligned and shown in the same orientation. The core structures are shown as white ribbons, whereas the variable regions covering 386–401, 402–428, and 429–498 in the influenza A virus NP and their counterparts in the influenza B virus are colored green, red, and yellow, respectively. The homotrimer and homotetramer of the influenza A virus **(D)** and the influenza B virus **(E)** NP are covered with a potential surface (positive and negative charges are colored blue and red with limits $\pm 20\ k_bT/e_c$). **(F)** Cryo-EM reconstruction of a native RNP by Arranz *et al.* [15] The viral polymerase complex is located at the bottom end of the RNP and is shown in green and orange. The two opposite-running NP–RNA strands are colored blue and pink, respectively. NP and RNA are fitted in the EM map in the lower panel. **(G)** Cryo-EM reconstruction of an authentic RNP by Moeller *et al.* [17] The viral polymerase is highlighted in red, whereas the NP–RNA is colored dark green. The structure of the NP protomer is fitted into the EM map in the lower panel; (F,G) were adapted with permission [15,17], respectively.

responses to double-stranded RNA (dsRNA) [3,5,39,40]. Moreover, based on the complex structure of LAFV NP with dTTP (a cap structure analog), the authors indicated that the N-terminal domain of LAFV NP acts as a cap-binding factor and that NP contributes directly to the cap-snatching process of viral replication. Brunotte *et al*. further analyzed this dTTP-binding site and showed that substitutions of the residues at this site do not specifically cause a defect in viral mRNA synthesis, and the nucleotide-binding pocket of LAFV NP could not be assigned a specific role in viral mRNA synthesis [25]. Furthermore, Hastie *et al*. presented the complex structure of the N-terminal domain of LAFV NP in complex with ssRNA, and this domain was assigned as an RNA-binding domain using an N-terminal domain construct (residues 1–340) [4]. Notably, their results revealed that the N-terminal domain of LAFV NP is unlikely to bind to the cap or its analogs *in vitro* [4].

The enzymatic activity of viral NP was subsequently reported in the CCHFV NP. Guo *et al*. showed that the head domain of CCHFV NP has high structural similarity to the LAFV NP N-terminal domain, although they have no primary sequence similarity [8]. However, the CCHFV NP (or the head domain alone) displays very low binding affinity for either a cap/cap analog or exogenous RNA, suggesting that at least its monomeric form is unlikely to have a cap-snatching or RNA-binding function similar to that of LAFV NP [8]. Most surprisingly, the CCHFV NP has a nuclease activity for both ss and ds DNA, which is stimulated to different extents by divalent cations [8].

Although the precise biological function of LAFV and CCHFV NPs remains elusive, these studies first revealed different enzymatic activities of virus-encoded NPs and extended our understanding of the biological roles of NP in the viral life cycle.

A new strategy for antiviral development

Because correct RNP formation and function is a key step in virus replication, transcription, and assembly, it is conceivable that blockade of RNP formation and its function would provide great potential for antiviral development. In the past few years great progress has been achieved based on this new strategy, particularly for influenza virus.

Through a chemical genetic method, Kao *et al*. first demonstrated that influenza virus NP is a druggable target [41]. They reported that a small-molecule compound, nucleozin, which triggers aggregation and inhibits the nuclear accumulation of NP, can inhibit the replication of influenza virus at a nanomolar median effective concentration (EC_{50}) [41]. In a parallel effort, Gerritz *et al*. discovered a series of influenza replication inhibitors and showed that they interfere with NP-dependent processes via the formation of higher-order NP oligomers with an EC_{50} up to 60 nM [42]. Notably, the structure of NP in

complex with a representative compound of these inhibitors revealed that two inhibitors in an antiparallel orientation lock two adjacent NP protomers. This quaternary complex explained viral inhibition via ligand-induced formation of stable NP oligomers [42]. In addition, disturbance of the polymerase complex of influenza virus was incorporated into antiviral strategies. Based on the complex structure of the PA C-terminal domain (PA_C) and the first 25 amino acids of PB1 [43], a subset of modifications of the N-terminal peptide of PB1 were shown to demolish the binding affinity of PA to PB1, as well as the polymerase activity of PA–PB1–PB2 polymerase complex, and attenuated the replication of influenza virus [44–46]. Given that both the polymerase complex and NP are highly conserved in different influenza viruses, these results cumulatively demonstrated that targeting the formation of viral RNP is a valid goal for the development of small-molecule therapies against viral resistance to currently available drugs, such as adamantanes or neuraminidase inhibitors.

Moreover, the structure of SFTSV NP in complex with Suramin, an antiviral inhibitor, revealed that the blocker binds to the RNA-binding cavity and attenuates SFTSV replication, indicating a new therapeutic antiviral approach to impact RNP formation [9].

Concluding remarks

Structural studies on NSRV RNPs from crystallography and EM provide insights into the dynamic processes of RNP formation. Both nsNSRV and sNSRV RNPs are formed through the conformational shifts of their NPs. During this process, specific position(s) of viral NPs, for example the N-arm of phleboviral NP, the N- and C-terminal extensions of orthobunyaviral NPs, and the domain exchange loop of influenza virus NP, play crucial roles in RNP formation.

Although the structures of NSRV-encoded NPs vary distinctly, nsNSRV NPs share homologous folding, whereas sNSRV NPs have multiple variations: some, for example influenza virus and orthobunyaviruses, have similarities with nsNSRVs, whereas others have distinct structural features. Interestingly, previous results revealed that most RNPs of sNSRV have a relaxed architecture, but form flexible supercoils in influenza virus and orthobunyavirus [31,33], which are similar to the flexible helical RNPs of nsNSRVs [47]. It is likely that the similar folding of NPs of influenza virus, orthobunyavirus, and nsNSRVs may easily lead the viral NPs to form higher-order helical structures. Most importantly, sNSRV NPs can have enzymatic activities beyond genome encapsidation. These data significantly extend our knowledge on viral RNP formation and function, and the structural information regarding the similarities and variations among viral NPs should be useful for viral taxonomy.

Although the structural and functional studies summarized here suggest an overall picture of RNP formation in NSRV, many important questions remain unanswered. For example, the genomes of nsNSRVs encode an accessory phosphoprotein (P) that participates in RNP formation and function [37]. Several studies have shown that sNSRVs also encode an accessory protein that regulates the formation and function of viral RNPs, in other words Z protein in arenavirus [48], NSm/NSs in orthobunyavirus [49], and NS1 in influenza virus [50]. However, structural details on the participation of these viruses in RNP formation have yet to be defined (Box 2). We anticipate that future developments and discussions will continue to explore structural aspects of viral RNPs, and antiviral reagents that impact upon RNP formation are likely to have clinical applications in the future.

Box 2. Outstanding questions

- What is the 'real' architecture of native RNP that is packaged into virions and what is their function in host cells? Are there any differences between the 'real' structures of native RNPs in variable environments?
- What is the dynamic mechanism to form the highly ordered structure of native RNP? What is the structure of the NP protein immediately after it is expressed? What is the key factor to induce the conformational change of the NP protein for RNA binding and oligomerization?
- What is the mechanism of RdRp accessibility to the RNA bases in the various NSRV RNPs during template copying? How do the 3′ and 5′ termini of the bunyaviral genome interact with RdRp? Although there is some biological evidence [54], the structural details are still unknown.
- How can we use structural findings in the development of antiviral therapeutics that target RNP formation and function?

Acknowledgments

This work was supported by the Ministry of Science and Technology of China '973' Project (Grants 2013CB911103 and 2014CB542800) and the National Natural Science Foundation of China (Grants 31000332 and 31170678).

References

1 King, A.M.Q. *et al.*, eds (2012) *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses*, Academic Press
2 Kranzusch, P.J. and Whelan, S.P. (2012) Architecture and regulation of negative-strand viral enzymatic machinery. *RNA Biol.* 9, 941–948
3 Hastie, K.M. *et al.* (2011) Structure of the Lassa virus nucleoprotein reveals a dsRNA-specific 3′ to 5′ exonuclease activity essential for immune suppression. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2396–2401
4 Hastie, K.M. *et al.* (2011) Crystal structure of the Lassa virus nucleoprotein–RNA complex reveals a gating mechanism for RNA binding. *Proc. Natl. Acad. Sci. U.S.A.* 108, 19365–19370
5 Qi, X. *et al.* (2011) Cap binding and immune evasion revealed by Lassa nucleoprotein structure. *Nature* 468, 779–783
6 Dong, H. *et al.* (2013) Structure of Schmallenberg orthobunyavirus nucleoprotein suggests a novel mechanism of genome encapsidation. *J. Virol.* 87, 5593–5601
7 Ferron, F. *et al.* (2011) The hexamer structure of Rift Valley fever virus nucleoprotein suggests a mechanism for its assembly into ribonucleoprotein complexes. *PLoS Pathog.* 7, e1002030
8 Guo, Y. *et al.* (2012) Crimean-Congo hemorrhagic fever virus nucleoprotein reveals endonuclease activity in bunyaviruses. *Proc. Natl. Acad. Sci. U.S.A.* 109, 5046–5051
9 Jiao, L. *et al.* (2013) Structure of severe Fever with thrombocytopenia syndrome virus nucleocapsid protein in complex with suramin reveals therapeutic potential. *J. Virol.* 87, 6829–6839
10 Li, B. *et al.* (2013) Bunyamwera virus possesses a distinct nucleocapsid protein to facilitate genome encapsidation. *Proc. Natl. Acad. Sci. U.S.A.* 110, 9048–9053
11 Niu, F. *et al.* (2013) Structure of the Leanyer orthobunyavirus nucleoprotein–RNA complex reveals unique architecture for RNA encapsidation. *Proc. Natl. Acad. Sci. U.S.A.* 110, 9054–9059
12 Raymond, D.D. *et al.* (2010) Structure of the Rift Valley fever virus nucleocapsid protein reveals another architecture for RNA encapsidation. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11769–11774

13 Ng, A.K. *et al.* (2008) Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. *FASEB J.* 22, 3638–3647
14 Ye, Q. *et al.* (2006) The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature* 444, 1078–1082
15 Arranz, R. *et al.* (2013) The structure of native influenza virion ribonucleoproteins. *Science* 338, 1634–1637
16 Chenavas, S. *et al.* (2013) Monomeric nucleoprotein of influenza a virus. *PLoS Pathog.* 9, e1003275
17 Moeller, A. *et al.* (2013) Organization of the influenza virus replication machinery. *Science* 338, 1631–1634
18 Green, T.J. *et al.* (2006) Structure of the vesicular stomatitis virus nucleoprotein–RNA complex. *Science* 313, 357–360
19 Albertini, A.A. *et al.* (2006) Crystal structure of the rabies virus nucleoprotein–RNA complex. *Science* 313, 360–363
20 Tawar, R.G. *et al.* (2009) Crystal structure of a nucleocapsid-like nucleoprotein–RNA complex of respiratory syncytial virus. *Science* 326, 1279–1283
21 Ge, P. *et al.* (2010) Cryo-EM model of the bullet-shaped vesicular stomatitis virus. *Science* 327, 689–693
22 Desfosses, A. *et al.* (2013) Self-organization of the vesicular stomatitis virus nucleocapsid into a bullet shape. *Nat. Commun.* 4, 1429
23 Buchmeier, M.J. *et al.* (2007) Arenaviridae: the viruses and their replication. In *Fields Virology* (5th edn) (Knipe, D.M. and Howley, P.M., eds), pp. 1791–1827, Lippincott Williams & Wilkins
24 Martinez-Sobrido, L. *et al.* (2007) Differential inhibition of type I interferon induction by arenavirus nucleoproteins. *J. Virol.* 81, 12696–12703
25 Brunotte, L. *et al.* (2011) Structure of the Lassa virus nucleoprotein revealed by X-ray crystallography, small-angle X-ray scattering, and electron microscopy. *J. Biol. Chem.* 286, 38748–38756
26 Raymond, D.D. *et al.* (2012) Phleboviruses encapsidate their genomes by sequestering RNA bases. *Proc. Natl. Acad. Sci. U.S.A.* 109, 19208–19213
27 Yu, X.J. *et al.* (2011) Fever with thrombocytopenia associated with a novel bunyavirus in China. *N. Engl. J. Med.* 364, 1523–1532
28 Zhou, H. *et al.* (2013) The nucleoprotein of severe fever with thrombocytopenia syndrome virus processes a stable hexameric ring to facilitate RNA encapsidation. *Protein Cell* 4, 445–455
29 Carter, S.D. *et al.* (2012) Structure, function, and evolution of the Crimean-Congo hemorrhagic Fever virus nucleocapsid protein. *J. Virol.* 86, 10914–10923
30 Wang, Y. *et al.* (2012) Structure of Crimean-Congo hemorrhagic fever virus nucleoprotein: superhelical homo-oligomers and the role of caspase-3 cleavage. *J. Virol.* 86, 12294–12303
31 Reguera, J. *et al.* (2013) Structural basis for encapsidation of genomic RNA by La Crosse Orthobunyavirus nucleoprotein. *Proc. Natl. Acad. Sci. U.S.A.* 110, 7246–7251
32 Dong, H. *et al.* (2013) Crystal structure of Schmallenberg orthobunyavirus nucleoprotein-RNA complex reveals a novel RNA sequestration mechanism. *RNA* 19, 1129–1136
33 Ariza, A. *et al.* (2013) Nucleocapsid protein structures from orthobunyaviruses reveal insight into ribonucleoprotein architecture and RNA polymerization. *Nucleic Acids Res.* 41, 5912–5926
34 Yuan, P. *et al.* (2009) Crystal structure of an avian influenza polymerase PA(N) reveals an endonuclease active site. *Nature* 458, 909–913
35 Zhao, C. *et al.* (2009) Nucleoside monophosphate complex structures of the endonuclease domain from the influenza virus polymerase PA subunit reveal the substrate binding site inside the catalytic center. *J. Virol.* 83, 9024–9030
36 Liu, Y. *et al.* (2009) Structure–function studies of the influenza virus RNA polymerase PA subunit. *Sci. China C: Life Sci.* 52, 450–458
37 Morin, B. *et al.* (2013) The polymerase of negative-stranded RNA viruses. *Curr. Opin. Virol.* 3, 103–110
38 Ng, A.K. *et al.* (2012) Structural basis for RNA binding and homo-oligomer formation by influenza B virus nucleoprotein. *J. Virol.* 86, 6758–6767
39 Jiang, X. *et al.* (2013) Structures of arenaviral nucleoproteins with triphosphate dsRNA reveal a unique mechanism of immune suppression. *J. Biol. Chem.* 288, 16949–16959
40 Hastie, K.M. *et al.* (2012) Structural basis for the dsRNA specificity of the Lassa virus NP exonuclease. *PLoS ONE* 7, e44211
41 Kao, R.Y. *et al.* (2010) Identification of influenza A nucleoprotein as an antiviral target. *Nat. Biotechnol.* 28, 600–605
42 Gerritz, S.W. *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
43 He, X. *et al.* (2008) Crystal structure of the polymerase PA(C)–PB1(N) complex from an avian influenza H5N1 virus. *Nature* 454, 1123–1126
44 Wunderlich, K. *et al.* (2011) Identification of high-affinity PB1-derived peptides with enhanced affinity to the PA protein of influenza A virus polymerase. *Antimicrob. Agents Chemother.* 55, 696–702
45 Wunderlich, K. *et al.* (2009) Identification of a PA-binding peptide with inhibitory activity against influenza A and B virus replication. *PLoS ONE* 4, e7517
46 Manz, B. *et al.* (2011) Disruption of the viral polymerase complex assembly as a novel approach to attenuate influenza A virus. *J. Biol. Chem.* 286, 8414–8424
47 Ruigrok, R.W. *et al.* (2011) Nucleoproteins and nucleocapsids of negative-strand RNA viruses. *Curr. Opin. Microbiol.* 14, 504–510
48 Kranzusch, P.J. and Whelan, S.P. (2011) Arenavirus Z protein controls viral RNA synthesis by locking a polymerase-promoter complex. *Proc. Natl. Acad. Sci. U.S.A.* 108, 19743–19748
49 Guu, T.S. *et al.* (2012) Bunyavirus: structure and replication. *Adv. Exp. Med. Biol.* 726, 245–266
50 Robb, N.C. *et al.* (2011) The influenza A virus NS1 protein interacts with the nucleoprotein of viral ribonucleoprotein complexes. *J. Virol.* 85, 5228–5231
51 Rizzetto, M. (2009) Hepatitis D: thirty years after. *J. Hepatol.* 50, 1043–1050
52 Mielke-Ehret, N. and Muhlbach, H.P. (2012) Emaravirus: a novel genus of multipartite, negative strand RNA plant viruses. *Viruses* 4, 1515–1536
53 Kormelink, R. *et al.* (2011) Negative-strand RNA viruses: the plant-infecting counterparts. *Virus Res.* 162, 184–202
54 Mir, M.A. and Panganiban, A.T. (2004) Trimeric hantavirus nucleocapsid protein binds specifically to the viral RNA panhandle. *J. Virol.* 78, 8281–8288