
Copyright © 2013 RNA Society

http://eprints.gla.ac.uk/82866

Deposited on: 12 September 2014
Crystal structure of Schmallenberg orthobunyavirus nucleoprotein–RNA complex reveals a novel RNA sequestration mechanism

Haohao Dong, Ping Li, Bettina Böttcher, et al.

*RNA* 2013 19: 1129-1136 originally published online June 24, 2013

Access the most recent version at doi:10.1261/rna.039057.113

---

### References

This article cites 25 articles, 11 of which can be accessed free at: [http://rnajournal.cshlp.org/content/19/8/1129.full.html#ref-list-1](http://rnajournal.cshlp.org/content/19/8/1129.full.html#ref-list-1)

### Creative Commons License

This article is distributed exclusively by the RNA Society for the first 12 months after the full-issue publication date (see [http://rnajournal.cshlp.org/site/misc/terms.xhtml](http://rnajournal.cshlp.org/site/misc/terms.xhtml)). After 12 months, it is available under a Creative Commons License (Attribution-NonCommercial 3.0 Unported), as described at [http://creativecommons.org/licenses/by-nc/3.0/](http://creativecommons.org/licenses/by-nc/3.0/).

### Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](http://rnajournal.cshlp.org/subscriptions).
Crystal structure of Schmallenberg orthobunyavirus nucleoprotein–RNA complex reveals a novel RNA sequestration mechanism

HAOHAO DONG,1,2,4 PING LI,2,4,5 BETTINA BÖTTCHER,3 RICHARD M. ELLIOTT,2,5,6 and CHANGJIANG DONG1,6

1Biomedical Research Centre, Norwich Medical School, University of East Anglia, Norwich Research Park NR4 7TJ, United Kingdom
2Biomedical Sciences Research Complex, School of Biology, University of St Andrews, St Andrews KY16 9ST, United Kingdom
3School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JN, United Kingdom

ABSTRACT

Schmallenberg virus (SBV) is a newly emerged orthobunyavirus (family Bunyaviridae) that has caused severe disease in the offspring of farm animals across Europe. Like all orthobunyaviruses, SBV contains a tripartite negative-sense RNA genome that is encapsidated by the viral nucleocapsid (N) protein in the form of a ribonucleoprotein complex (RNP). We recently reported the three-dimensional structure of SBV N that revealed a novel fold. Here we report the crystal structure of the SBV N protein in complex with a 42-nt-long RNA to 2.16 Å resolution. The complex comprises a tetramer of N that encapsidates the RNA as a cross-shape inside the protein ring structure, with each protomer bound to 11 ribonucleotides. Eight bases are bound in the positively charged cleft between the N- and C-terminal domains of N, and three bases are shielded by the extended N-terminal arm. SBV N appears to sequester RNA using a different mechanism compared with the nucleoproteins of other negative-sense RNA viruses. Furthermore, the structure suggests that RNA binding results in conformational changes of some residues in the RNA-binding cleft and the N- and C-terminal arms. Our results provide new insights into the novel mechanism of RNA encapsidation by orthobunyaviruses.

Keywords: Schmallenberg virus; orthobunyavirus; bunyavirus; nucleoprotein structure; RNA replication and transcription

INTRODUCTION

Schmallenberg virus (SBV) was first identified in November 2011 in northern Germany, and has since spread across Europe (Hoffmann et al. 2012; Beer et al. 2013), probably vectored by culicoid midges (De Regge et al. 2012; Rasmussen et al. 2012). SBV infects cattle, sheep, and goats, causing transient fever, diarrhea, reduction in milk yield, and abortion or malformations in offspring, leading to considerable economic losses (Garigliany et al. 2012; Tarlinton et al. 2012). No vaccines are yet available to control the disease.

SBV is most closely related to Sathuperi virus (Goller et al. 2012) in the Simbu serogroup of the genus Orthobunyavirus in the family Bunyaviridae. There are more than 350 recognized bunyaviruses that are divided into five genera, Orthobunyavirus, Hantavirus, Nairovirus, Phlebovirus, and Tospovirus (Plyusnin et al. 2012). The orthobunyaviruses make up the largest genus in Bunyaviridae, comprising about 170 members. Some orthobunyaviruses, such as La Crosse virus (LACV) and Oropouche virus, are significant human pathogens, whereas others like Akabane and Cache Valley viruses cause economically important disease in ruminants. Like all bunyaviruses, SBV is an enveloped negative-stranded RNA virus with a tripartite genome. The large (L) segment encodes a RNA-dependent RNA polymerase (L protein), and the middle segment (M) encodes two surface glycoproteins (Gc and Gn) and a nonstructural protein (NSm), while the small segment encodes a nucleoprotein (N) and, for most orthobunyaviruses, a second nonstructural protein called NSs (Elliott et al. 2011). The nucleoprotein interacts with viral genomic and antigenomic (replicative intermediate) RNA species to form ribonucleoprotein complexes (RNP) that are the functional templates for RNA replication and transcription (Elliott et al. 2011). We recently determined the structure of SBV N protein, which showed a novel fold, and proposed a new mechanism for genome RNA encapsidation and replication (Dong et al. 2013). Although the potential RNA-binding cleft was identified, details of the interaction between N and the viral RNA to form the ribonucleoprotein complex, essential for understanding viral RNA encapsidation, replication, and transcription, were not elucidated. We now report...
the structure of the SBV nucleoprotein in complex with a 42-base-long RNA.

RESULTS

SBV N forms a tetramer when complexed with 42-nt RNA

The SBV N protein was expressed in bacteria, and then either purified in the native form or the protein was denatured and refolded as previously described (Dong et al. 2013). As natively purified N binds host RNA nonspecifically and heterogeneously, we attempted to reconstitute a homogenous N–RNA complex by incubating RNA-free N with chemically synthesized RNAs (polyU) of 21, 28, or 42 bases in length. Only crystals of SBV N in complex with the 42-base-long RNA diffracted at high resolution (see below). The oligomeric state of the N–42-nt RNA complex was examined by chemical cross-linking and SDS-PAGE and compared with that of the natively purified N or RNA-free refolded N, suggesting that the RNA-free refolded N protein and the N–42-nt RNA complex exist as trimer and tetramer, respectively. This correlates with the tetrameric state of natively purified SBV N that contains host RNA (Dong et al. 2013).

Structure of tetrameric N in complex with 42-nt RNA

Crystals of SBV N complexed with 42-nt RNA were obtained in space group P2₁ with unit cell dimensions \(a = 76.50 \text{ Å}, b = 86.05 \text{ Å}, c = 77.46 \text{ Å}, \) and \(\beta = 101.98^\circ\). The complex structure was determined by molecular replacement using Phaser (McCoy et al. 2007) and the native SBV N structure (PDB access code 4IDU) as the searching model. The N–RNA complex formed a tetrameric ring structure (Fig. 1A,B), in which the C-terminal arm of each protomer stretches out to bind to a hydrophobic region in the C-terminal domain of the neighboring N molecule (Fig. 1C), as previously described (Dong et al. 2013). Meanwhile, the N-terminal arms of three N protomers move away from their own or neighboring RNA-binding clefts and interact with their neighboring N molecules to further stabilize the oligomeric structure, while the fourth N-terminal arm is disordered (Fig. 1D). The RNA molecule is bent extensively between the protomers to form a cross-shape inside the tetrameric ring (Fig. 1A,B).

RNA is mainly bound at the central RNA-binding cleft

The electron density of the 42-nt RNA is clearly visible alongside the inner face of the tetramer (Fig. 2A). We previously suggested that the central positively charged cleft between the N- and C-domains of N is used for RNA binding (Dong et al. 2013). Indeed, our structure of the N–RNA complex clearly shows that the cleft can host eight RNA bases and

![FIGURE 1. Tetrameric structure of SBV N–RNA complex. (A) Cartoon representation of the tetrameric structure of the SBV N–RNA complex. Four protomers are shown in blue, green, yellow, and cyan. The 42-nt-long RNA is shown in stick form in orange bound inside the tetrameric ring, forming a cross shape. The black dotted circle shows the gap in the RNA. (B) Same as in A but rotated along the \(\gamma\)-axis at 180°. (C) Surface representation of the tetrameric structure. The red arrows point to the C-terminal arms, which bind to the neighboring protomer’s C-terminal binding sites, thus playing an essential role for oligomerization. (D) Same as in C but rotated along the \(\gamma\)-axis at 180°. The light blue arrows point to the N-terminal arms, which bind RNAs and to neighboring protomers. One of the N-terminal arms, indicated by a star, is disordered.](image-url)
that residues K48, K51, R166, H77, R95, K178, K179, R182, and R184 play a crucial role in binding the RNA (Fig. 2B–D). This is consistent with our previous observations that mutations of residues K48 and K51 impaired RNA-binding activity (Dong et al. 2013). It is also noticeable that the protrusion of the hydrophobic residues F176 and F18 changes the orientation of the RNA bases and results in an “S” shape for the RNA chain in the cleft (Fig. 2B, C). In detail, bases U1, U2, and U3 face toward the inside of the RNA-binding cleft, interacting with positively charged residues K48, K51, and R166, and form hydrophobic interactions with residues L126 and F176 (Fig. 2B,D); F176 lies

**FIGURE 2.** SBV N–RNA interactions. (A) Unbiased Fo–Fc electron density map of the 42-nt polyU contoured at 3 σ. The electron density is shown in blue and the RNA is shown in stick form in orange. The four protomers are shown in ribbon form. The black dotted circle shows the RNA gap. (B) Electrostatic potential surface map showing RNA binding in a SBV N protomer. The RNA is shown as stick form in orange. One protomer binds 11 nt of RNA with 8 nt in the RNA-binding cleft and 3 nt to the N-terminal arm. Bases U1 to U3 face into protein inside the RNA-binding cleft, bases U4 to U7 point outside the RNA-binding cleft and are exposed to solvent, while base U8 is bent into the bottom of the binding cleft and faces the protein. (C) Interaction between the RNA and residues of the RNA-binding cleft. The RNA forms an “S” shape in the RNA-binding cleft. The RNA is shown as stick form in orange, while amino acids of the RNA-binding cleft are shown in cyan. (D) Schematic diagram showing interactions between RNA and amino acids in SBV N. Bases oriented to the top indicate those that face into protein in the RNA-binding cleft, while the bases oriented toward the bottom indicate the bases exposed to solvent. The dark dotted lines show the interactions. (E) Interaction of the N-terminal arm and the RNA. Bases U9–U11 interact with atoms of the main chain of the N-terminal arm. Base U9 is stacked above base U7. The RNA is shown in orange and the protein in cyan. (F) The N-terminal arm binds to another hydrophobic area of a neighboring protomer. In particular, residues F7 and F66 form π–π interaction. The N-terminal arm is in cyan and the hydrophobic area of the neighboring protomer is in blue.
between bases U3 and U4. Bases U4, U5, U6, and U7, which are facing outward to solvent, stack together at the surface of the cleft, where these bases interact with positively charged residues R95, K178, K179, and R182 (Fig. 2B–D). Meanwhile, base U8 sits deeply in the cleft and is sandwiched by residues R184 and N19 (Fig. 2B–D), whereas residue F18 sits between bases U7 and U8.

The N-terminal arm is also involved in RNA binding

Beside the RNA-binding cleft, the N-terminal arm of N also contributes to RNA binding. The arm, which is located between two neighboring N protomers, can host the remaining three bases (U9–11) (Fig. 2B–D). In particular, residues A15, A16, and Q12 on the N-terminal arm interact with bases U9, U10, and U11 to shield the RNA from exposure to the outside of the tetramer. This is consistent with our previous observation that truncation of the N-terminal arm completely abrogated RNA-binding activity and the N-terminal arm was disordered in the refolded RNA-free N structure (Dong et al. 2013). The N-terminal arm binds to a hydrophobic area consisting of residues V42, V53, V62, L64, and F66 in a neighboring protomer (Fig. 2F) and, in particular, residues F7 and F66 form π–π hydrophobic interactions. As each N protomer can bind 11 bases, the tetramer is able to bind 44-nt RNA. However, because we used a 42-nt RNA to generate the complex for crystallization, this results in a gap that is visible in the tetrameric complex structure (Fig. 1A,B), and more clearly seen in the electron density map of the RNA (Fig. 2A).

EM structure of SBV N

The oligomeric states of the natively purified N, RNA-free refolded N, and the refolded N complexed with 42-nt RNA were also examined by electron microscopy. The natively purified N and the N–RNA complex were mainly seen as tetramers but minor amounts of other oligomeric states, such as trimers or pentamers, were observable (Fig. 3A,C). The oligomeric states of the RNA-free refolded N were heterogeneous and irregular compared with the natively purified N (Fig. 3B). However, the predominant population of the refolded protein was identified as trimers, although tetramers and pentamers were also present (Fig. 3B). The EM projection maps also match well with the crystal structures, as shown in overlays of the crystal structures to selected class averages (Fig. 3D). Thus, the EM observations were largely in agreement with our chemical cross-linking analysis and crystal structure data.

RNA-free SBV N undergoes conformational changes upon RNA binding

We noticed that the RNA-free refolded N was unstable and liable to aggregation during the denaturation and refolding process, suggesting that N requires RNA binding to stabilize its structure. In solution the refolded SBV N is heterogeneous with a trimer as the predominant species (Figs. 3B, 4A), and the hexameric form may be required to form crystals based on the observed crystal packing. The hexameric structure was generated according to symmetry-related molecules (Dong et al. 2013).

In the structure of trimeric refolded N without bound RNA (Fig. 4A) (PDB code 4IDX) all of the positively charged residues in the RNA-binding clefts face toward the solution phase and, therefore, would be accessible to RNA in solution. When incubated with the 42-nt polyU RNA, the refolded SBV N converted from a trimer to a tetramer, where the RNA-binding clefts now face toward the inside of the ring, thus shielding the RNA from the outside (Fig. 1A,B). The
overall structure of the SBV N in the N–RNA complex is similar to the refolded RNA-free SBV N with an RMSD of 1.05 over 201 residues (Ca backbones) (Fig. 4B). The most notable changes are that three of the N-terminal arms in the tetramer become ordered in the SBV N–RNA complex structure (Fig. 1D) and that there are conformational changes in the C-terminal arms. Furthermore, the side chains of the RNA-binding residues R41, K48, L45, K51, H77, V82, F18, N19, P20, R184, R182, K178, K179, F176, and R166 undergo conformational changes in order to bind the RNA (Fig. 4C).

**DISCUSSION**

The genomic RNA of negative-stranded viruses is encapsidated by nucleoprotein in the form of RNP. The RNP not only serves as a template for viral RNA replication and transcription by the cognate viral RNA-dependent RNA polymerase (RdRp), but also protects the viral genome from ribonucleases and helps to avoid alerting the host innate immune response (Qi et al. 2010; Jiang et al. 2013). Although the nucleoproteins have similar functions in different families of negative-sense RNA viruses, their structures seem to be very variable, as exemplified by our previous publication reporting the crystal structure of the SBV N protein (Dong et al. 2013). Here, we have extended this work by determining the structure of the SBV N protein in complex with RNA at high resolution, which sheds new light on details of the N–RNA interaction for orthobunyaviruses in general.

Rift Valley fever virus (RVFV) and SBV are classified in different genera in the family Bunyaviridae and have similarly sized N proteins, but their N protein folds are different (Raymond et al. 2010; Ferron et al. 2011). The nucleoprotein of RVFV uses its flexible N-terminal arm for N multimerization (Fig. 5A), and the RNPs form irregular structures due to this flexibility. Both the N- and C-terminal arms of SBV N are highly flexible and our data show that the SBV N–RNA complex forms a ring-like tetrameric structure mediated by these arms. Intriguingly, this is reminiscent of the nucleoproteins of rabies and vesicular stomatitis viruses (VSV), two nonsegmented genome viruses that form ring structures by multimerization mediated by the N-terminal arm and a loop in the C-terminal domain (Fig. 5B,C). This arrangement allows the RNP to form a super helix. Whether the same pertains for SBV RNPs is currently unknown. Although the majority of the native SBV N forms tetramers as observed by EM and chemical cross-linking, further study is needed to determine whether the building block for RNP assembly is the tetramer or some other form, such as the monomer.

A common feature of negative-sense RNA virus nucleoproteins is the presence of an RNA-binding cleft or groove that is formed by mainly positively charged residues, but to date few structures of RNA bound to nucleoprotein have been determined. The available structures indicate that the mechanisms by which the nucleoproteins sequester RNAs are different. For both rabies virus and VSV, each N subunit binds nine ribonucleotides (six bases in RNA-binding cleft and three bases between N subunits) with six bases exposed to solvent and three bases facing into the protein (Fig. 5B, C) (Albertini et al. 2006; Green et al. 2006). The Lassa fever virus nucleoprotein binds six RNA bases that face the outside of the RNA-binding cleft and are exposed to solvent (Fig. 5D; Hastie et al. 2011), while in RVFV nucleoprotein, four RNA bases locate to the bottom of the RNA-binding cleft and face into the protein (Fig. 5A); an additional two or three bases are bound between N–N subunits (Raymond et al. 2012). In a further variation, we show that each SBV N subunit sequesters eight bases in the RNA-binding cleft with bases U1, U2, U3, and U8 located to the bottom of the RNA-binding cleft to face into the protein, bases U4–U7 exposed to solvent, and the N-terminal arm binding another three bases in the N–N interface (Fig. 2D).

For genomic RNA replication and transcription, the bases of the RNA have to be "read" by the viral RNA-dependent RNA polymerase. As the genomic RNA is closely encapsidated in the N oligomeric ring, the question arises as to how the RNA becomes accessible to the polymerase. In the SBV N–RNA complex there are four bases facing into the protein in the RNA-binding cleft, which suggests that
the encapsidated RNA has to be partially released from the RNA-binding cleft or that the N undergoes a conformational change (or rotation) to permit polymerase access. Comparing the RNA-free N and RNA–N complex structures, we observe conformational change in some of the RNA-binding residues in the RNA-binding cleft and N-terminal and C-terminal arms. To achieve this, the flexible N- and C-terminal arms may play an essential role. Indeed, truncation of the N-terminal arm of SBV N completely abrogates its RNA-binding activity, truncation of the C-terminal arm decreases RNA-binding ability, while truncations of either or both arms results in loss of RNA replication activity (Dong et al. 2013).

In summary, we have determined the SBV N–RNA complex structure to high resolution that suggests a new RNA sequestration mechanism in orthobunyavirus RNP formation. These data will help to further our understanding of viral RNA encapsidation, replication, and transcription, and may aid in the development of antiviral compounds against diseases caused by medically important orthobunyaviruses.

Note added in Revision: While responding to the reviewers’ comments on this manuscript, lower resolution structures of N protein–ssRNA complexes from other viruses in the Orthobunyavirus genus, including La Crosse (3.4 Å resolution), Leanyer (2.78 Å resolution), and Bunyamwera viruses (3.2 Å resolution) were reported (Ariza et al. 2013; Li et al. 2013; Niu et al. 2013; Reguera et al. 2013). They all share similar structural features and proposed mode of RNA binding. For example, all complexes form tetrameric structures that encapsidate the RNA inside the ring. Each N protomer binds 10 or 11 ribonucleotides, and the C-terminal and N-terminal arms are involved in N multimerization.

**MATERIALS AND METHODS**

**Protein expression, purification, and N–RNA complex (RNP) reconstitution**

Construction of bacterial expression plasmid p14SBVN and purification of natively purified SBV N (with host RNA bound) and refolded RNA-free N have been described previously (Dong et al. 2013). The natively purified and refolded RNA-free SBV N was concentrated to 9 mg/mL in 20 mM Tris (pH 7.5), 300 mM NaCl, and 10% glycerol. To reconstitute RNP, the refolded N was mixed in a 1:1 molar ratio with RNA (poly U of 21, 28, or 42 bases in length; Eurogentec) and incubated on ice for 90 min. The preparation was screened for crystallization conditions.
Chemical cross-linking of SBV N protein

Cross-linking of natively purified SBV N, refolded RNA-free N, and the refolded N–RNA complex was performed using dithiobis (succinimidyl propionate; DSP), as previously reported (Dong et al. 2013). Briefly, samples containing 10–20 µg of protein were treated with 1 mM DSP for 30 min at room temperature in PBS buffer and then analyzed by SDS-PAGE under reducing or nonreducing conditions, followed by Coomassie blue staining.

Negatively stained EM structures of SBV N

For electron microscopy the refolded SBV N, native SBV N, or refolded SBV N–42-nt RNA complex were diluted to concentrations between 2 and 20 µg/mL. The protein was incubated for 1–2 min on carbon-coated grids, and then excess sample was removed. The grids were washed twice with water and twice with 2% uranyl acetate before staining for 5 min with 2% uranyl acetate. Excess staining solution was blotted off and the grids were air dried.

For imaging, grids were transferred to a FEI F20 electron microscope equipped with a 4k × 4k CMOS camera (TVIPS F416). Micrographs were recorded with EM-Tools (TVIPS GmbH) under low-dose conditions at 200 kV, a primary magnification of 66,000 (calibrated pixel size of 1.73 Å), and an approximate dose of 20 e/Å².

TABLE 1. Statistics of data collection and structure refinement

<table>
<thead>
<tr>
<th>Data collection</th>
<th>Schmallenberg virus N–42-nt RNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (Å)</td>
<td>0.92</td>
</tr>
<tr>
<td>Resolution (highest shell) (Å)</td>
<td>37.82–2.12 (2.23–2.16)²</td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁</td>
</tr>
<tr>
<td>Cell constants (Å)</td>
<td>a = 76.35, b = 85.96, c = 77.33</td>
</tr>
<tr>
<td></td>
<td>β = 102.02</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>54303 (7984)</td>
</tr>
<tr>
<td>Average redundancy</td>
<td>4.1 (4.3)</td>
</tr>
<tr>
<td>I/σ</td>
<td>11.8 (2.6)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>97.9 (99.2)</td>
</tr>
<tr>
<td>Rmerge (%)</td>
<td>0.061 (0.572)</td>
</tr>
<tr>
<td>Wilson B-factor</td>
<td>33.14</td>
</tr>
</tbody>
</table>

Refinement

Number of atoms

Protein: 7026
RNA: 840
Solvent: 152

Mean B-factors (Å²)

Protein: 45.0
RNA: 48.0
Solvent: 52.0

R_factor

25.94

R_mol

27.66

RMSD bonds (Å)/angles (°) 0.010/1.500

Residues in Ramaanchandran core (%) 95.6

PDBe accession code 4JNG

*Values in parentheses are for the highest-resolution shell.

DATA DEPOSITION

The atomic coordinates and the structure factors have been deposited in the Protein Data Bank with PDBe access code 4JNG for Schmallenberg virus N–42-nt RNA complex.

ACKNOWLEDGMENTS

We thank the staff of IO4 and I24 beam stations at Diamond Light Resources UK and Dr. Xiaohong Shi for critical reading of the manuscript. This work was supported by the Medical Research Council (grant No G1100110/1 to C.D.) and Wellcome Trust (Career Development Fellowship WT083501MA to C.D., and Programme Grant 079810 and Project Grant 091785 to R.M.E.). R.M.E. is a Welcome Trust Senior Investigator. B.B. was supported by the Darwin Trust of Edinburgh. The use of the electron microscopy facility was supported by the Wellcome Trust Equipment grant WT087658 and by the Scottish University Life Sciences Alliance (SULSA).

Received March 15, 2013; accepted May 31, 2013.

REFERENCES

Dong et al.


