Crystal structure of Junin virus nucleoprotein

Yinjie Zhang,1,2,3 Le Li,2,3 Xiang Liu,3 Shishang Dong,2,3 Wenming Wang,2,3 Tong Huo,1,2,3 Yu Guo,1 Zihe Rao1,2,3 and Cheng Yang1,2,3

Correspondence
Cheng Yang
Cheng.yang@htmdc.org

Received 8 May 2013
Accepted 17 July 2013

1State Key Laboratory of Medicinal Chemical Biology and College of Pharmacy, Nankai University, Tianjin 300071, PR China
2College of Life Sciences, Nankai University, Tianjin 300071, PR China
3High-Throughput Molecular Drug Discovery Center, Tianjin Joint Academy of Biomedicine and Technology, Tianjin 300457, PR China

Junin virus (JUNV) has been identified as the aetiological agent of Argentine haemorrhagic fever (AHF), which is a serious public health problem with approximately 5 million people at risk. It is treated as a potential bioterrorism agent because of its rapid transmission by aerosols. JUNV is a negative-sense ssRNA virus that belongs to the genus Arenavirus within the family Arenaviridae, and its genomic RNA contains two segments encoding four proteins. Among these, the nucleoprotein (NP) has essential roles in viral RNA synthesis and immune suppression, but the molecular mechanisms of its actions are only partially understood. Here, we determined a 2.2 Å crystal structure of the C-terminal domain of JUNV NP. This structure showed high similarity to the Lassa fever virus (LASV) NP C-terminal domain. However, both the structure and function of JUNV NP showed differences compared with LASV NP. This study extends our structural insight into the negative-sense ssRNA virus NPs.

INTRODUCTION

Junin virus (JUNV) has been identified as the aetiological agent of Argentine haemorrhagic fever (AHF), a disease endemic to the Pampas region of Argentina that represents a serious public health problem with approximately 5 million people at risk (Mills et al., 1994; Parodi et al., 1966). AHF occurs mainly during the harvesting season in Argentina, with a peak incident in May. The disease is four times as prevalent in males as it is in females, and is more prevalent among rural workers than in urban populations (Grant et al., 2012; Mills et al., 1994; Parodi et al., 1966). JUNV spreads through the aerosolization of host rodent Calomys musculinus and is treated as a potential bioterrorism agent because of its rapid transmission by aerosols. Whilst the nucleoside analogue ribavirin in combination with hyperimmune serum has been used for therapeutic treatment of JUNV infections, these agents are only effective if administered during a specific time period and have shown limited success (Enria et al., 2008). Thus, the development of novel antiviral drugs is significant for controlling AHF caused by JUNV.

JUNV belongs to the genus Arenavirus within the family Arenaviridae. The family Arenaviridae consists of this one unique genus containing more than 20 recognized virus species, which are classified into two distinct groups based on serological cross-reactivity and sequence-based phylogeny: Old World (OW) and New World (NW) arenaviruses (Martinez-Sobrido et al., 2007). At least five members of this family cause viral haemorrhagic fevers, including Lassa fever virus (LASV), an OW arenavirus, and JUNV, Machupo virus (MACV), Guanarito virus (GUAV) and Sabia virus (SABV) found in South America, which belong to the NW clade B arenaviruses (Cuevas et al., 2011). JUNV is an enveloped virus with a negative-sense ambisense ssRNA that contains two genomic segments encoding four proteins. The small (S) RNA segment (3.5 kb) encodes the viral glycoprotein precursor (GPC) and the nucleoprotein (NP), and the large (L) RNA segment (7.3 kb) encodes the viral RNA-dependent RNA polymerase and the small (11 kDa) RING finger protein Z found in many other negative-sense ssRNA viruses (Casabona et al., 2009; Levingston Macleod et al., 2011; Perez et al., 2003; Strecker et al., 2003; Urata et al., 2006, 2009).

Among these four proteins, NP is the most abundant viral polypeptide in both virions and infected cells. It encapsidates and packages genomic RNA to prevent it from degrading in the host cell, and also mediates the interaction between genomic RNA and viral RNA-dependent RNA polymerase to form the ribonucleoprotein complex (Kranzusch & Whelan, 2012; Ruigrok et al., 2011). Recent studies have shown that the virally encoded NPs have enzymic activity beyond RNA encapsidation activity, which extends our understandings of the multiple roles of...
virally encoded NPs (Kranzusch & Whelan, 2012; Ruigrok et al., 2011; Sun et al., 2012). NPs of LASV, JUNV, MACV and lymphocytic choriomeningitis virus (LCMV) have all been implicated in suppression of the innate interferon (IFN) response, resulting in the inhibition of nuclear translocation of IFN regulatory factor 3 (IRF-3) (Hastie et al., 2011a; Martinez-Sobrido et al., 2007). The arenavirus NP has distinct N- and C-terminal domains connected by a flexible linker (Brunotte et al., 2011; Hastie et al., 2011a; Levinston Macleod et al., 2011; Qi et al., 2010). The C-terminal domain functions as an exonuclease specific for dsRNA and is linked to antagonism of type I IFN (Hastie et al., 2011a, b; Qi et al., 2010).

To date, structural information on several NPs throughout a wide range of negative-sense ssRNA virus families has been reported. The negative-sense ssRNA viruses are classified depending on whether their viral genomes are non-segmented or segmented. The non-segmented negative-sense ssRNA viruses comprise the families Rhabdoviridae, Paramyxoviridae, Bornaviridae and Filoviridae, whereas the segmented negative-sense ssRNA viruses comprise the families Arenaviridae, Bunyaviridae and Orthomyxoviridae. The crystal structure of the NP of Borna disease virus (BDV) from the family Bornaviridae, in the absence of RNA, is the first non-segmented negative-sense ssRNA virus NP structure that crystallized as a tetramer (Rudolph et al., 2003). The crystal structures of the NP–RNA complexes of vesicular stomatitis virus (VSV) and rabies virus from the family Rhabdoviridae and respiratory syncytial virus (RSV) from the family Paramyxoviridae first illustrated the structure and molecular mechanism by which non-segmented negative-sense ssRNA virus NP encapsidates RNA (Albertini et al., 2008; Ge et al., 2010; Tawar et al., 2009). For segmented negative-sense ssRNA viruses, two independent groups reported the crystal structure of RNA-free NP from influenza A virus (IAV) from the family Orthomyxoviridae (Ng et al., 2008; Ye et al., 2006), whereas another work reported the crystal structure of NP from influenza B virus (IBV) (Ng et al., 2012) showing totally different oligomerization from IAV NP. Recently, four groups reported NP–RNA complexes of four kinds of viruses from the genus Orthobunyavirus, in the family Bunyaviridae, that all showed a tetrameric organization with the bound RNA (Ariza et al., 2013; Li et al., 2013; Niu et al., 2013; Reguera et al., 2013). Besides the function of RNA encapsidation, recent studies have revealed unexpected enzymic activities of NPs. For example, the NP structure of Crimean–Congo hemorrhagic fever virus (CCHFV) from the family Bunyaviridae has been solved and shows intrinsic nuclease activity on both ssDNA and dsDNA (Guo et al., 2012). Until now, in the family Arenaviridae, only the structures of LASV NP and Tacaribe arenavirus (TCRV) NP C-terminal domain have been solved (Jiang et al., 2013; Qi et al., 2010). The structures revealed that the C-terminal domain possesses 3′→5′ RNA-specific exoribonuclease activity that can suppress IFN production by degrading immune-stimulatory RNAs (Hastie et al., 2011a; Jiang et al., 2013; Qi et al., 2010). However, two research groups hold different opinions on the function of the LASV NP N-terminal domain. Qi et al. (2010) explained it as a cap-binding factor based on the complex structure with dTTP, whereas Hastie et al. (2011b) reported it as an RNA-binding domain that controls RNA binding through a gating mechanism.

To further address the functional mechanisms of NP in viral RNA synthesis and host immune suppression, we set out to determine the crystal structure for JUNV NP, knowledge of which can be extended to other negative-sense ssRNA virus NP proteins.

RESULTS

Overall structure of JUNV NPΔ340

The recombinant C-terminal domain of JUNV NP (NPΔ340) existed mainly in two forms during purification, but only the form comprising the front peak during the heparin column purification was able to crystallize well (Fig. 1a). We used the C-terminal domain structure of LASV NP as the model for phasing. The JUNV NPΔ340 structure formed a typical α/β/α sandwich architecture. It consisted of six α-helices connected by a series of loops and five-stranded β-sheets with one antiparallel strand. The β-sheets were surrounded by three groups of α-helices (x1, x3 and x4; x2; and x5 and x6) that were assigned by their geographical positions. Two loops corresponding to residues 407–431 and 499–521 were missing in the structure that would connect β3–z3 and z5–z6 (Fig. 2). The mass spectrometry result of JUNV NP crystals showed that these two loops existed during the crystallization process, suggesting that they were invisible due to high flexibility (Fig. 1b).

The C-terminal domain of JUNV NP is similar to the homologous LASV NP

Multiple sequence alignment and comparison of the structure with reported structures in the RCSB Protein Data Bank (PDB) using the Dali structure comparison service (Holm & Rosenström, 2010) revealed that JUNV NPΔ340 had the motif DEDD and was similar to the C-terminal domain of LASV NP. The DEDD superfamily contains strictly conserved Asp–Glu–Asp–Asp catalytic residues in the active site, which is a conserved motif in this kind of exonuclease (Cisneros et al., 2009; de Silva et al., 2007; Zuo et al., 2007) (Fig. 3). In this structure, the active site consisted of the amino acids D380, E382, D457 and D529. In addition, H524 was also near to the active site of the DEDD superfamily (Fig. 4a). These amino acids formed an acidic cavity that is shared by some exonucleases, such as LASV NP and human TREX2, which could protect human immunodeficiency virus from the host innate immune response (Hastie et al., 2011a; Qi et al., 2010; Yan et al.,
Fig. 1. Ion-exchange chromatography and mass spectrometry data of the JUNV NPΔ340 protein solution. (a) JUNV NPΔ340 samples were injected onto a heparin column. The upper panel shows SDS-PAGE analysis of the ion-exchange chromatography elution fractions corresponding to the two peaks. (b) JUNV NP (JNP) coverage map. The residues of missing loops are boxed. The existing residues determined by mass spectrometry are coloured. Residues matching peptides are coloured blue, and residues matching modified peptides are coloured green. The colouring is transparent, so that regions where peptides overlap are visible.

Fig. 2. Structure of JUNV NPΔ340. (a) Topology diagram of JUNV NPΔ340. α-Helices are shown in blue, β-sheets are yellow and loops are green lines. (b) Schematic representation of JUNV NPΔ340 from the front (left) and the top (right). N-ter, N-terminal end; C-ter, C-terminal end.
However, no divalent ions could be found in the acidic cavity. To investigate further the similarity between JUNV NP and the C-terminal domain of LASV NP, the structure of LASV NP was superposed on to our structure and the root mean square deviation (RMSD) value was 1.51 Å for all Cα atoms of the 147 aligned residues (Fig. 5a).

In LASV NP, a Mn2+ is present in the active site coordinated with D389, E391 and D533, whilst no electric densities were found in the corresponding location in our structure. Thus, our structure remained in an ion-free state. Therefore, without the Mn2+, which can coordinate the surrounding amino acids, the amino acids near the active site in our structure were shifted slightly compared to those in LASV NP, especially residues E382 and D457 (Fig. 5b).

In LASV NP, a Zn2+ was also found that coordinated to the residues E399, C506, H509 and C529. The corresponding amino acids from our structure were E390, C497, H500 and C525, which were also slightly different from those in LASV NP. Due to no ion here leading to flexible conformations and poor electric densities, the side chain of H500 could not be properly placed; the residue C497 partially participated in forming a disulfide bond with C525 and was partially directed to the possible Zn2+ binding position (Fig. 5c). All the residues coordinated around the Mn2+ and Zn2+ active sites were strictly conserved in arenaviruses, indicating their important roles in these viruses (Fig. 3).

The C-terminal domain of JUNV NP has no exonuclease activity in vitro

As the homologue LASV NP has been shown to be a 3′→5′ exoribonuclease (Hastie et al., 2011a, b; Qi et al., 2010), we investigated the exonuclease activity of JUNV NPA340. We conducted an analysis by atomic absorption spectrometry and in vitro 3′→5′ exonuclease activity assays to...
determine the JUNV NPΔ340 cooperative capability with divalent ions and its exonuclease activity. According to the results of atomic absorption spectrometry, there was no Mn$^{2+}$ or Zn$^{2+}$ atom in the NP$^\Delta$340 protein solution before crystallization (Table 1). Therefore, NP$\Delta$340 could not incorporate divalent ions during the crystallization process, and we could not find Mn$^{2+}$ or Zn$^{2+}$ in the structure. The exonuclease activity assays demonstrated that the JUNV NP$\Delta$340 had no exonuclease activity (Fig. 6a). However, the LASV NP C-terminal domain could digest dsRNA well under the same experimental conditions (Fig. 6b), which indicated JUNV NP$\Delta$340 and the LASV NP C-terminal domain have different functions.

**DISCUSSION**

In this work, we determined the structure of the C-terminal domain of JUNV NP at a resolution of 2.2 Å, and found a similarity with the LASV NP C-terminal domain. However, both the structure and function also showed differences compared with LASV NP.

Numerous 3′→5′ exoribonucleases are members of the DEDDh family. Two independent research groups have demonstrated the 3′→5′ exoribonuclease activity of the LASV NP C-terminal domain by *in vitro* exonuclease activity assays compared with wild-type LASV NP, as well as using NP mutants at putative catalytic sites, and the LASV NP C-terminal domain belongs to the DEDDh family (Hastie *et al.*, 2011a; Qi *et al.*, 2010). As well as LASV, another RNA virus, severe acute respiratory syndrome coronavirus also encodes a non-structural protein, nsp14, with exonuclease activity. The nsp14 protein possesses activity against both ssRNA and dsRNA for controlling genome fidelity and is also a member of the DEDDh family (Eckerle *et al.*, 2010; Minskaia *et al.*, 2006). Another DEDDh family member, human TREX2, can digest ssDNA produced by human immunodeficiency virus instead of RNA in order to protect the virus from the host innate immune response (Yan *et al.*, 2010).

Exoribonucleases in the DEDDh family share a common catalytic mechanism that depends on two metal ions (Steitz & Steitz, 1993), whereas JUNV NPΔ340 could not incorporate divalent ions during the crystallization process.
or in the structure. In addition, exonuclease activity assays demonstrated that NPΔ340 has no exonuclease activity.

Focusing on the possible Mn$^{2+}$-binding position in the JUNV NPΔ340 structure, we found that the residue D457 maintained a distance away from the possible active site so that the motif DEDD may not exist in the JUNV NPΔ340 structure, and JUNV NPΔ340 may not belong to the DEDD superfamily of exonucleases (Fig. 4a).

Multimerization of NPs has been observed in several negative-sense ssRNA viruses (Albertini et al., 2006; Green et al., 2006; Tawar et al., 2009; Ye et al., 2006). However, JUNV NPΔ340 mainly existed in a monomeric form in our study (data not shown). Two factors possibly account for this monomeric form of JUNV NPΔ340. One is that the structure we determined is not the full-length JUNV NP, as the N-terminal domain may play an essential role in protein oligomerization. The other one is that JUNV NP might execute its function by the monomeric form, as the NP of CCHFV reported recently was found to be able to act as a divalent cation-dependent endonuclease in its monomeric form in vitro (Guo et al., 2012).

As the JUNV genomic RNA only encodes four proteins during its life cycle, the NP is likely to play essential roles in multiple stages of virus replication and transcription. Thus, we suggest that JUNV NPΔ340 might possess other important functions instead of the exonuclease activity.

Here, we have presented a crystal structure of the C-terminal domain of NP in JUNV. It showed similarities to

<table>
<thead>
<tr>
<th>Ion</th>
<th>JUNV NPΔ340 #1 (3.8 mg ml$^{-1}$)</th>
<th>NPA340: ion molar ratio</th>
<th>JUNV NPΔ340 #2 (19 mg ml$^{-1}$)</th>
<th>NPA340: ion molar ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mn$^{2+}$</td>
<td>0.031</td>
<td>253</td>
<td>0.019</td>
<td>2046</td>
</tr>
<tr>
<td>Zn$^{2+}$</td>
<td>0.275</td>
<td>34</td>
<td>0.207</td>
<td>222</td>
</tr>
<tr>
<td>Mg$^{2+}$</td>
<td>0.168</td>
<td>20</td>
<td>0.107</td>
<td>158</td>
</tr>
<tr>
<td>Ca$^{2+}$</td>
<td>1.871</td>
<td>3</td>
<td>0.950</td>
<td>30</td>
</tr>
</tbody>
</table>

**Table 1.** Atomic absorption spectrometry analysis of JUNV NPΔ340 protein solution

Four divalent ions, Mn$^{2+}$, Zn$^{2+}$, Mg$^{2+}$ and Ca$^{2+}$ were detected. The protein concentrations of samples #1 and #2 were different as indicated; results are shown in mg l$^{-1}$. Analysis was carried out using a Hitachi 180-80 Polarized Zeeman Atomic Absorption Spectrometer.
the C-terminal domain of NP in LASV. However, there were also differences between LASV and JUNV NP. As arenavirus NPs within the NW group share high similarity, the knowledge obtained from the structure of JUNV NP could be extended to other pathogenic arenaviruses that can cause severe hemorrhagic fever diseases in humans.

**METHODS**

**Cloning, protein expression and purification.** Codon-optimized cDNAs for full-length JUNV NP were synthesized by GENEWIZ. The C-terminal domain (NPΔ340, residues 341–564) of NP was PCR amplified (forward primer, 5′-CGGGATCCATGCCGCGCCAACCGC-3′; reverse primer, 5′-GGCCTCGAGTTACAGCGCGTATGCGGCTTTT-3′). The PCR product was digested with BamHI and XhoI restriction enzymes (Takara) and inserted into a linearized pGEX-6p-1 plasmid (GE Healthcare). Eluted NP340 protein was further purified using a heparin column (GE Healthcare). SDS-PAGE analysis revealed over 95 % purity of the final purified recombinant protein. Protein concentration was adjusted to 10 mg ml⁻¹ in buffer containing 20 mM Tris/HCl (pH 8.5), 500 mM NaCl, 1 µg RNase ml⁻¹ and 1 µg DNase I ml⁻¹ and homogenized with a low-temperature ultrahigh pressure cell disrupter (JNBIO). The lysate was centrifuged at 25 000 g for 30 min at 277 K to remove cell debris. Affinity purification was achieved by two passes of the supernatant through 2 ml glutathione S-transferase affinity medium. In-column digestion was carried out for 16 h at 277 K with PreCission Protease (GE Healthcare). Eluted NP340 protein was further purified using a heparin column (GE Healthcare). SDS-PAGE analysis revealed over 95 % purity of the final purified recombinant protein. Protein concentration was adjusted to 10 mg ml⁻¹ in buffer containing 20 mM Tris/HCl (pH 8.0), 50 mM NaCl for crystallization trials.

**Crystallization.** Initial crystallization conditions were screened by the hanging-drop vapour-diffusion method using commercial crystal screening kits at 289 K including the Index, Crystal Screen, PEG/Ion, Salt/RX and Crystal Screen cryo from Hampton Research and Wizard kits (Emerald BioSystems). Each crystallization drop contained 1 µl protein solution and 1 µl reservoir solution and was equilibrated against 200 µl reservoir solution. Small crystals first appeared after 2 weeks in 0.1 M Tris/HCl (pH 8.5) and 2 M ammonium sulfate. Further optimization was performed, and the final optimized crystals were grown in buffer containing 0.1 M Tris/HCl (pH 8.5), 2 M ammonium sulfate. All chemicals used for crystallization were purchased from Sigma. The crystals grew to a final size of 5 × 10 × 200 µm within 2 weeks.

**Data collection and processing.** X-ray diffraction data were collected on a Beamline BL-17A (100 K, λ=1.000 Å) of the Photon

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>NP</th>
<th>Lasv</th>
<th>JUNV</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>20</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>40</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>60</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>120</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>120</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>120</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>120</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
Factory (Tsukuba, Japan) and processed to 2.2 Å resolution, using HEL-2000 (Otwinowski & Minor, 1997) for data indexing and scaling. Molecular replacement was performed by Phaser using the LAVS NP C-terminal domain (PDB 3Q7B) as a searching template (Terwilliger, 2000; Terwilliger & Berendzen, 1999). Further model building was performed automatically by Phenix AutoBuild (Adams et al., 2002), and manually by Coot (Emsley & Cowtan, 2004). The structure was refined using Phenix (Adams et al., 2002). Data were indexed to space group P2_1_2_1, with cell parameters a=40.4 Å, b=78.5 Å and c=136.0 Å. Refinement converged to Rwork/Rfree=0.19/0.22 with good stereochemistry judged from a Ramachandran plot. Most residues appeared in the favoured region (96.6%) and no residues were disallowed in the Ramachandran plot. Model geometry was verified using Procheck (Laskowski et al., 1993). All structure figures were drawn with PyMOL (DeLano, 2002). Coordinates and structure factors have been deposited in the RCSB PDB under accession code 4K7E and the statistics are summarized in Table 2.

Exonuclease activity assays. An exonuclease activity assay of JUNV NP with dsRNA sequence 5'-UUCCCGAAGUGUCUGACGUTT-3' (sense) as substrate was carried out in 10 µl reaction solution containing 0.2 M NaCl, 20 mM HEPES (pH 7.5), 5 mM MnCl₂ and 4 µg NP, in the presence of dsRNA substrate, at 310 K for 10–120 min. The control reactions included 10 mM EDTA. All structure figures were drawn with PyMOL (DeLano, 2002). Coordinates and structure factors have been deposited in the RCSB PDB under accession code 4K7E and the statistics are summarized in Table 2.

Acknowledgements

We thank the staff of Beamline BL-17A at the Photon Factory (KEK, Japan), for their help with data collection. This work was supported by the National Natural Science Foundation of China (nos 31200641, 31200586), National Biomedical Special Project of International Innovation Park (11ZCKFSY06300 and 11ZCKFSY06900) and Fundamental Research Funds for the Central University.

References


