STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Crystallographic analysis of the N-terminal domain of *Middle East respiratory syndrome coronavirus* nucleocapsid protein

Yong-Sheng Wang,[a,b] Chung-ke Chang[c] and Ming-Hon Hou[a,b]*

[a]Department of Life Sciences, National Chung Hsing University, Taichung 402, Taiwan, [b]Institute of Genomics and Bioinformatics, National Chung Hsing University, Taichung 402, Taiwan, and [c]Institute of Biomedical Sciences, Academia Sinica, Taipei 11529, Taiwan. *Correspondence e-mail: mhho@nchu.edu.tw

Received 18 March 2015
Accepted 26 May 2015

Edited by M. S. Weiss, Helmholtz-Zentrum Berlin für Materialien und Energie, Germany

Keywords: *Middle East respiratory syndrome coronavirus*; nucleocapsid protein; N-terminal domain; RNA binding.

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

The N-terminal domain of the nucleocapsid protein from *Middle East respiratory syndrome coronavirus* (MERS-CoV NP-NTD) contains many positively charged residues and has been identified to be responsible for RNA binding during ribonucleocapsid formation by the virus. In this study, the crystallization and crystallographic analysis of MERS-CoV NP-NTD (amino acids 39–165), with a molecular weight of 14.7 kDa, are reported. MERS-CoV NP-NTD was crystallized at 293 K using PEG 3350 as a precipitant and a 94.5% complete native data set was collected from a cooled crystal at 77 K to 2.63 Å resolution with an overall $R_{\rm merge}$ of 9.6%. The crystals were monoclinic and belonged to space group $P2_1$, with unit-cell parameters a = 35.60, b = 109.64, c = 91.99 Å, β = 101.22°. The asymmetric unit contained four MERS-CoV NP-NTD molecules.

1. Introduction

Between 2003 and 2004, *Severe acute respiratory syndrome coronavirus* (SARS-CoV) caused a worldwide epidemic and had a significant economic impact in countries affected by the outbreak (Lai, 2003). In 2004, another alphacoronavirus, *Human coronavirus NL63* (HCoV-NL63), was isolated from a seven-month-old child in the Netherlands suffering from bronchiolitis and conjunctivitis (Pyrc *et al.*, 2004). In 2005, Woo and coworkers discovered the novel betacoronavirus *Human coronavirus HKU1* in patients with respiratory-tract infections (Woo *et al.*, 2005). Recently, *Middle East respiratory syndrome coronavirus* (MERS-CoV) has been found in patients with severe acute respiratory-tract infections in the Middle East (Woo *et al.*, 2014). As is the case for all coronaviral infections, there are no efficacious therapies currently available against coronaviral diseases, making the development of anticoronaviral compounds a priority. The virion envelope surrounding the nucleocapsid contains the following structural proteins: S (spike), M (matrix), E (envelope) and N (nucleocapsid). Some of them have a third glycoprotein, HE (haemagglutinin-esterase), which is present in most alphacoronaviruses. The primary function of the HCoV N protein is to recognize a stretch of RNA that serves as a packaging signal and leads to the formation of the ribonucleoprotein (RNP) complex during assembly (Chang *et al.*, 2014). The RNP may be important in keeping the RNA in an ordered conformation suitable for replication and transcription of the viral genome (Lai, 2003; Nelson *et al.*, 2000; Huang *et al.*, 2004; Navas-Martín & Weiss, 2004). In addition, the N protein has been identified as an important diagnostic marker and the most

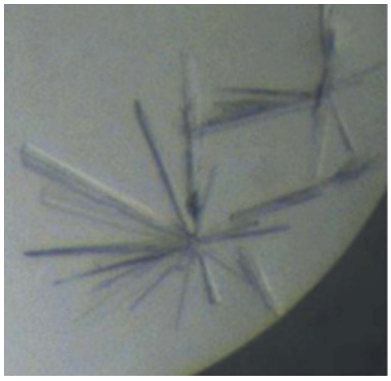

Table 1
Macromolecule-production information.

Source organism	*Middle East respiratory syndrome coronavirus*
DNA source	The templates for the MERS-CoV N protein were provided by AllBio Science Incorporation, Taichung, Taiwan. The primers for the MERS-CoV N protein were provided by Genomics BioSci & Tech Ltd, New Taipei City, Taiwan.
Forward primer	CTTATCGCATATGAACACCGTGAGCTGGTATACCGGC
Reverse primer	CTTACGGCTCGAGGGTGCCTTCAATATGAAAGTTTTTCG
Cloning vector	pET-28a (Novagen)
Expression vector	pET-28a (Novagen)
Expression host	*Escherichia coli* BL21 (DE3) pLysS
Complete amino-acid sequence of the construct produced	HHHHHHSSGLVPRGSHMNTVSWYTGLTQHGKVPLTFPPGQGVPLNANSTPAQNAGYWRRQDRKINTGNGIKQLAPRWYFYYTGTGPEAALPFRAVKDGIVWVHEDGATDAPSTFGTRNPNNDSAIVTQFAPGTKLPKNFHIEGT

'immunodominant' antigen in infected hosts (Chan *et al.*, 2005; Woo *et al.*, 2004; Liang *et al.*, 2013).

The N protein of MERS-CoV, with a molecular weight of 45.6 kDa and a pI of 10.05, is highly basic and hydrophilic (Woo *et al.*, 2014). Previous studies revealed that the N-terminal domain of CoV (NP-NTD) contains mostly positively charged residues, which are responsible for RNA binding, while the C-terminal domain (NP-CTD) mainly acts as an oligomerization module to form a capsid (Huang *et al.*, 2009; Saikatendu *et al.*, 2007; Lo *et al.*, 2013; Chen *et al.*, 2013). The central disordered region of the N protein has also been shown to contain an RNA-binding region (Chang *et al.*, 2009). We have shown that compounds that bind to the NP-NTD and interfere with NP–RNA interactions provide valuable leads for the development of anti-coronaviral therapeutics (Lin *et al.*, 2014). The crystal structures of several NP-NTDs, including those from SARS-CoV, *Infectious bronchitis virus* (IBV), *Human coronavirus OC43* (HCoV-OC43) and *Mouse hepatitis virus* (MHV), have been described (Chen *et al.*, 2013; Ma *et al.*, 2010; Saikatendu *et al.*, 2007; Jayaram *et al.*, 2006; Yu *et al.*, 2006; Fan *et al.*, 2005). In order to clarify the mechanism by which the N protein of MERS-CoV bind to nucleic acids, we have undertaken the determination of the crystal structure of the N-terminal domain of MERS-CoV spanning residues 39–165, which shares 58% identity to NP-NTD of SARS-CoV. The results presented in this paper mainly concern the crystallization and preliminary X-ray structural analysis of MERS-CoV NP-NTD.

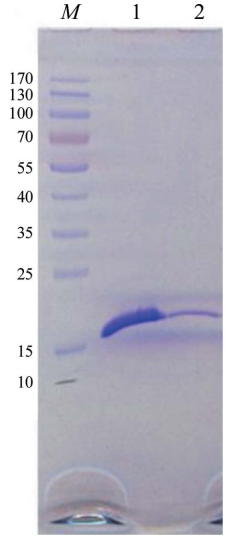

Figure 1
SDS–PAGE analysis of MERS-CoV NP-NTD stained with Coomassie Brilliant Blue. Lane *M*, protein molecular-mass marker (labelled in kDa); lane 1, concentrated MERS-CoV NP-NTD after dialysis; lane 2, purified MERS-CoV NP-NTD.

2. Materials and methods

2.1. Macromolecule production

The templates for the MERS-CoV N protein were purchased from AllBio Science Incorporated, Taichung, Taiwan. Truncated forms of recombinant MERS-CoV NP-NTD were generated by polymerase chain reaction (PCR) of a pGENT plasmid encoding the N-protein gene using different primers. The PCR products were digested with NdeI and XhoI, and the DNA fragments were cloned into pET-28a (Novagen) using T4 ligase (NEB). Expression of the protein was initiated by adding IPTG to a final concentration of 1 m*M* followed by incubation at 10°C for 24 h. After harvesting the bacteria by centrifugation (8000 rev min^{-1}, 12 min, 4°C), the bacterial pellets were resuspended in lysis buffer (50 m*M* Tris–HCl, 150 m*M* NaCl, 15 m*M* imidazole, 1 m*M* PMSF pH 7.5) and lysed by sonication on ice using 3 s pulses with 6 s pauses for a total of 10 min. Soluble proteins were obtained from the supernatant after centrifugation (13 000 rev min^{-1}, 40 min, 4°C). NP-NTD carrying an N-terminal His_6 tag fused to a SSGLVPRGSHM linker sequence was purified using an Ni–NTA column (Novagen) and eluted with a buffered imidazole gradient of 15–250 m*M*. Fractions containing pure protein

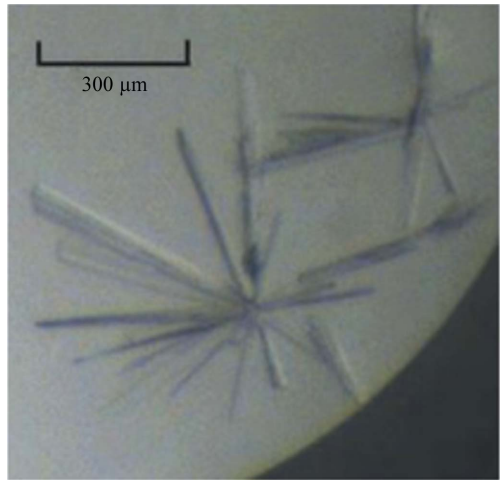

Figure 2
Crystals of MERS-CoV NP-NTD obtained by the sitting-drop vapour-diffusion method. The largest crystals in the needle clusters are approximately 300 × 20 × 10 µm in size.

were collected at ∼150 m*M* imidazole and were dialyzed against 50 m*M* Tris-buffered solution at pH 7.5 containing 150 m*M* NaCl for 3 h at 4°C (Fig. 1). The purified NP-NTD was further concentrated using Amicon Ultra centrifugal filter units and centrifuged at 3500*g* for 10 min at 4°C several times until the concentration of NP-NTD reached 10 mg ml^{-1} as determined by the Bradford method (BioShop Canada Inc.). Macromolecule-production information is summarized in Table 1.

2.2. Crystallization

Initial crystallization conditions were identified by using the sitting-drop vapour-diffusion method with crystal screening kits from Molecular Dimensions as described previously (Till *et al.*, 2013; Chen *et al.*, 2014). Each of the solutions (2 µl) from the crystal screening kits was mixed with 2 µl purified protein solution (10 mg ml^{-1}) and allowed to equilibrate against 300 µl solution in the well at room temperature (∼25°C). The conditions were refined and crystals were grown from a well solution using the sitting-drop vapour-diffusion method by equilibrating a mixture of 2 µl protein solution (10 mg ml^{-1}) and 2 µl reservoir solution against 300 µl reservoir solution consisting of 2 m*M* NaBr, 75 m*M* ammonium sulfate, 29% PEG 3350 (Sigma). Crystals appeared within two weeks and the largest crystal in the needle clusters grew to dimensions of approximately 300 × 20 × 10 µm (Fig. 2). Crystallization information is summarized in Table 2.

2.3. Data collection and processing

X-ray data were collected using synchrotron radiation with a crystal-to-detector distance of 350 mm. The oscillation width and exposure time for each frame were 1° and 20 s, respectively. Crystallographic data integration and reduction were performed with the *HKL*-2000 program package (Otwinowski & Minor, 1997). The crystallographic data-collection statistics for NP-NTD are listed in Table 3.

Table 2
Crystallization.

Method	Vapour diffusion
Plate type	24-well sitting-drop plate (Hampton Research)
Temperature (K)	293
Protein concentration (mg ml^{-1})	10
Buffer composition of protein solution	50 m*M* Tris–HCl pH 7.5, 75 m*M* NaCl
Composition of reservoir solution	2 m*M* NaBr, 75 m*M* ammonium sulfate, 29% PEG 3350
Volume and ratio of drop	1:1; 2 µl reservoir solution was mixed with 2 µl purified protein solution
Volume of reservoir (µl)	300

3. Results and discussion

The MERS-CoV NP-NTD crystal chosen for this study diffracted to 2.63 Å resolution (Fig. 3) and belonged to space group $P2_1$, with unit-cell parameters a = 35.60, b = 109.64, c = 91.99 Å, β = 101.22°. The Matthews coefficient of 2.63 Å^3 Da^{-1} calculated using *MATTHEWS_COEF* from *CCP*4 (Winn *et al.*, 2011; Matthews, 1968) suggested that there were four molecules in an asymmetric unit with a solvent content of 59.2%. A homology search for the MERS-CoV NP-NTD structure was performed using the *BLAST* server (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The model-selection criterion was based on the *E*-value and the estimated precision value. The sequence-alignment search indicated that MERS-CoV NP-NTD shares high identity with other NP-NTDs from coronaviruses (Supplementary Fig. S1). For example, the MERS-CoV NP-NTD (residues 39–165) has 58% sequence

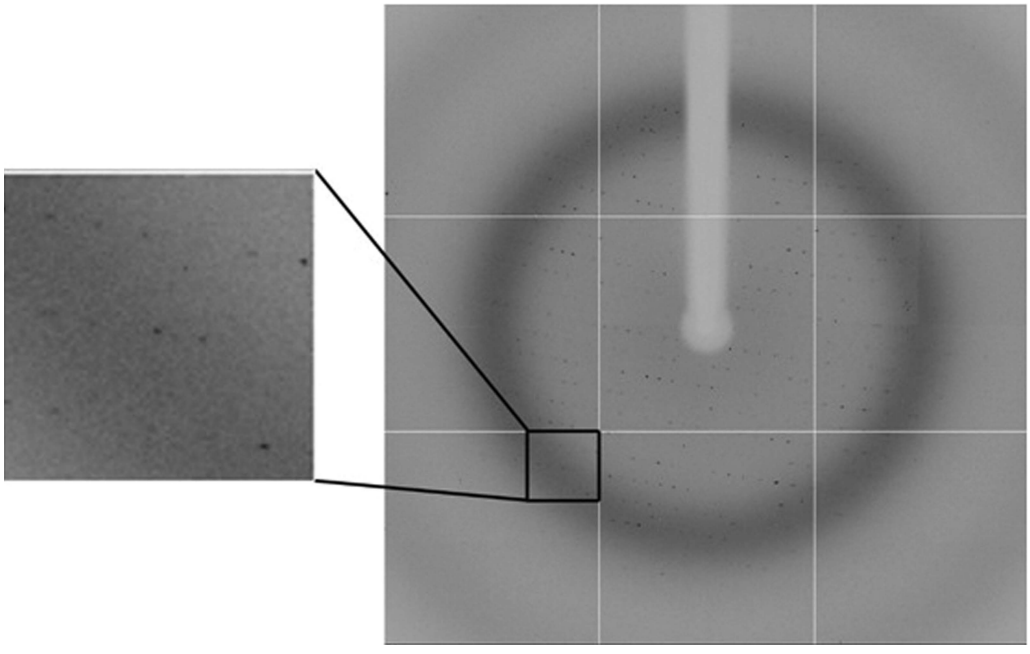

Figure 3
Typical X-ray diffraction pattern of MERS-CoV NP-NTD.

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	BL13B1 beamline, National Synchrotron Radiation Research Center (NSRRC), Hsinchu, Taiwan
Wavelength (Å)	1.000
Temperature (K)	77
Detector	ADSC Q315r
Crystal-to-detector distance (mm)	350
Rotation range per image (°)	1
Total rotation range (°)	360
Exposure time per image (s)	20
Space group	$P2_1$
a, b, c (Å)	35.60, 109.64, 91.99
α, β, γ (°)	90, 101.22, 90
Mosaicity (°)	1.053
Resolution range (Å)	30–2.63 (2.73–2.63)
Total No. of reflections	138439
No. of unique reflections	13974
Completeness (%)	94.5 (91.4)
Multiplicity	7.1 (5.9)
$\langle I/\sigma(I)\rangle$	18.89 (3.03)
$R_{\rm r.i.m.}$	0.01 (0.08)
Overall B factor from Wilson plot (Å^2)	43.1

identity to SARS-CoV NP-NTD. The NTD from SARS-CoV (PDB entry 2ofz; Saikatendu *et al.*, 2007) was chosen as the initial model as its *E*-value was 1×10^{-23}, and a total of 116 residues were modelled. The first molecular-replacement trial was performed using the *PERON* automated interface at the Protein Tectonics Platform (PTP), RIKEN SPring-8 Center, Japan (Sugahara *et al.*, 2008). The core of the model consisted of a tightly packed β-sheet surrounded by large loops. The molecular-replacement method was then applied to the model using *MOLREP* (Vagin & Teplyakov, 2010) using reflections in the resolution range 8.5–3.0 Å. Single and unambiguous solutions for four NP-NTD molecules in one asymmetric unit were obtained in the rotation and translation functions, yielding a final correlation coefficient of 0.81 and an *R* factor of 0.32. Structural refinement of MERS-CoV NP-NTD is currently in progress.

Acknowledgements

This work was supported by the MOST 103-2113-M-005-007-MY3. We thank the beamline staff at the BL13B1 beamline, National Synchrotron Radiation Research Center (NSRRC), Hsinchu, Taiwan for assistance. We are also grateful for access to the Protein Tectonics Platform (PTP), which is a structural biology facility collaboration between RIKEN SPring-8 Center, Japan and NSRRC, Taiwan.

References

Chan, K. H., Cheng, V. C. C., Woo, P. C. Y., Lau, S. K. M., Poon, L. L. M., Guan, Y., Seto, W. H., Yuen, K. Y. & Peiris, J. S. M. (2005). *Clin. Diagn. Lab. Immunol.* **12**, 1317–1321.
Chang, C.-K., Hou, M.-H., Chang, C.-F., Hsiao, C.-D. & Huang, T.-H. (2014). *Antiviral Res.* **103**, 39–50.
Chang, C.-K., Hsu, Y.-L., Chang, Y.-H., Chao, F.-A., Wu, M.-C., Huang, Y.-S., Hu, C.-K. & Huang, T.-H. (2009). *J. Virol.* **83**, 2255–2264.
Chen, I.-J., Yuann, J.-M. P., Chang, Y.-M., Lin, S.-Y., Zhao, J., Perlman, S., Shen, Y.-Y., Huang, T.-H. & Hou, M.-H. (2013). *Biochim. Biophys. Acta*, **1834**, 1054–1062.
Chen, Y.-W., Jhan, C.-R., Neidle, S. & Hou, M.-H. (2014). *Angew. Chem. Int. Ed.* **53**, 10682–10686.
Fan, H., Ooi, A., Tan, Y. W., Wang, S., Fang, S., Liu, D. X. & Lescar, J. (2005). *Structure*, **13**, 1859–1868.
Huang, C.-Y., Hsu, Y.-L., Chiang, W.-L. & Hou, M.-H. (2009). *Protein Sci.* **18**, 2209–2218.
Huang, Q., Yu, L., Petros, A. M., Gunasekera, A., Liu, Z., Xu, N., Hajduk, P., Mack, J., Fesik, S. W. & Olejniczak, E. T. (2004). *Biochemistry*, **43**, 6059–6063.
Jayaram, H., Fan, H., Bowman, B. R., Ooi, A., Jayaram, J., Collisson, E. W., Lescar, J. & Prasad, B. V. (2006). *J. Virol.* **80**, 6612–6620.
Lai, M. M. C. (2003). *J. Biomed. Sci.* **10**, 664–675.
Liang, F.-Y., Lin, L.-C., Ying, T.-H., Yao, C.-W., Tang, T.-K., Chen, Y.-W. & Hou, M. H. (2013). *J. Virol. Methods*, **187**, 413–420.
Lin, S.-Y., Liu, C.-L., Chang, Y.-M., Zhao, J., Perlman, S. & Hou, M.-H. (2014). *J. Med. Chem.* **57**, 2247–2257.
Lo, Y.-S., Lin, S.-Y., Wang, S.-M., Wang, C.-T., Chiu, Y.-L., Huang, T.-H. & Hou, M.-H. (2013). *FEBS Lett.* **587**, 120–127.
Ma, Y., Tong, X., Xu, X., Li, X., Lou, Z. & Rao, Z. (2010). *Protein Cell*, **1**, 688–697.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Navas-Martín, S. R. & Weiss, S. (2004). *J. Neurovirol.* **10**, 75–85.
Nelson, G. W., Stohlman, S. A. & Tahara, S. M. (2000). *J. Gen. Virol.* **81**, 181–188.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pyrc, K., Jebbink, M. F., Berkhout, B. & van der Hoek, L. (2004). *Virol. J.* **1**, 7.
Saikatendu, K. S., Joseph, J. S., Subramanian, V., Neuman, B. W., Buchmeier, M. J., Stevens, R. C. & Kuhn, P. (2007). *J. Virol.* **81**, 3913–3921.
Sugahara, M. *et al.* (2008). *J. Struct. Funct. Genomics*, **9**, 21–28.
Till, M., Robson, A., Byrne, M. J., Nair, A. V., Kolek, S. A., Shaw Stewart, P. D. & Race, P. R. (2013). *J. Vis. Exp.* (78), e50548.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Woo, P. C. Y. *et al.* (2005). *J. Infect. Dis.* **192**, 1898–1907.
Woo, P. C. Y., Lau, S. K. P., Wernery, U., Wong, E. Y. M., Tsang, A. K. L., Johnson, B., Yip, C. C. Y., Lau, C. C. Y., Sivakumar, S., Cai, J.-P., Fan, R. Y. Y., Chan, K.-H., Mareena, R. & Yuen, K.-Y. (2014). *Emerg. Infect. Dis.* **20**, 560–572.
Woo, P. C. Y., Lau, S. K. P., Wong, B. H. L., Chan, K.-H., Hui, W.-T., Kwan, G. S. W., Peiris, J. S. M., Couch, R. B. & Yuen, K.-Y. (2004). *J. Clin. Microbiol.* **42**, 5885–5888.
Yu, I.-M., Oldham, M. L., Zhang, J. & Chen, J. (2006). *J. Biol. Chem.* **281**, 17134–17139.