

Crystallization and preliminary crystallographic study of human coronavirus NL63 main protease in complex with an inhibitor

Fenghua Wang,^{a,b,c} ‡ Yusheng Tan,^a ‡ Huiyan Li,^d Xia Chen,^{b,c} Jinshan Wang,^{b,c} Shuang Li,^b Sheng Fu,^b Qi Zhao,^e Cheng Chen,^{a,b,*} Dan Su^{d,*} and Haitao Yang^{a,b}

^aSchool of Life Sciences, Tianjin University, Tianjin 300072, People's Republic of China,

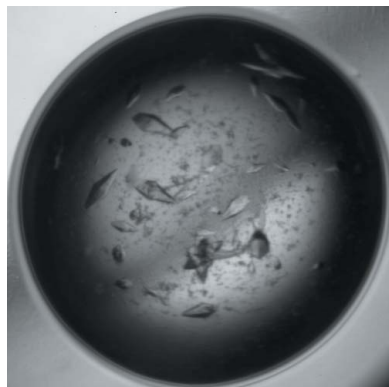
^bTianjin International Joint Academy of Biotechnology and Medicine, Tianjin 300457, People's Republic of China, ^cCollege of Life Sciences, Nankai University, Tianjin 300071, People's Republic of China, ^dState Key Laboratory of Biotherapy and Cancer Center, West China Medical School, West China Hospital, Sichuan University, Chengdu 610017, People's Republic of China, and ^eDepartment of Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, CT 06520, USA

‡ These authors contributed equally to this paper.

Correspondence e-mail: chengchen@tju.edu.cn, sudan@scu.edu.cn

Received 7 April 2014

Accepted 4 June 2014



© 2014 International Union of Crystallography
All rights reserved

Human coronavirus NL63 mainly infects younger children and causes cough, fever, rhinorrhoea, bronchiolitis and croup. It encodes two polyprotein precursors required for genome replication and transcription. Each polyprotein undergoes extensive proteolytic processing, resulting in functional subunits. This process is mainly mediated by its genome-encoded main protease, which is an attractive target for antiviral drug design. In this study, the main protease of human coronavirus NL63 was crystallized in complex with a Michael acceptor. The complex crystals diffracted to 2.85 Å resolution and belonged to space group $P4_12_12$, with unit-cell parameters $a = b = 87.2$, $c = 212.1$ Å. Two molecules were identified per asymmetric unit.

1. Introduction

Since the global outbreak of severe acute respiratory syndrome (SARS) in 2003, which was caused by SARS coronavirus (SARS-CoV), human CoVs have been in the spotlight. Shortly thereafter in 2004, another human CoV, named HCoV-NL63, was isolated from the nasopharyngeal aspirate of a seven-month-old child with coryza, conjunctivitis and fever hospitalized in Amsterdam (van der Hoek *et al.*, 2004). It has been shown that HCoV-NL63 mainly infects younger children and immunocompromised persons and causes mild upper respiratory symptoms such as cough, fever and rhinorrhoea or more serious lower respiratory tract involvement such as bronchiolitis and croup (Abdul-Rasool & Fielding, 2010). This CoV can be detected in as high as 1.0–9.3% of respiratory tract infections in children worldwide (Vabret *et al.*, 2005; van der Hoek *et al.*, 2006; Fielding, 2011).

Similar to other coronaviruses, HCoV-NL63 contains a single-stranded positive-sense polyadenylated RNA genome that encodes two large polyproteins (pp1a and pp1ab) which are processed into 16 nonstructural proteins (nsp1–16) by two virus-encoded proteinases. Nsp5, also named main protease (M^{pro}), is responsible for 11 out of 15 cleavage sites, thus playing a pivotal role in this digestion process, and is indispensable for viral replication (van der Hoek *et al.*, 2006; Pyrc *et al.*, 2004, 2007). The absence of its cellular homologue in humans makes it an ideal target for antiviral drug design (Yang *et al.*, 2005; Anand *et al.*, 2005). To date, several crystal structures of other coronavirus main proteases have been solved (Yang *et al.*, 2003, 2005). The sequence similarity of HCoV-NL63 protease to those of porcine transmissible gastroenteritis virus (TGEV) and SARS-CoV is 60 and 44%, respectively. Based on the structural analysis of these main proteases, the idea of designing wide-spectrum inhibitors against CoVs has been proposed. Here, we report the crystallization and preliminary crystallographic study of human coronavirus NL63 main protease in complex with a designed Michael acceptor inhibitor named N3.

2. Materials and methods

2.1. Protein expression and purification

The coding sequence for HCoV-NL63 main protease was subcloned into pGEX-6P-1 plasmid and verified by sequencing. The recombinant plasmid was then transformed into *Escherichia coli* strain BL21 (DE3) for protein expression. Cultures were grown in LB medium containing 0.1 mg ml⁻¹ ampicillin at 310 K until the optical density at 600 nm reached 0.6. Isopropyl β -D-1-thiogalactopyranoside was added to a final concentration of 0.5 mM and the cultures were induced to express HCoV-NL63 main protease at 289 K for 16 h. Thereafter, centrifugation was used to harvest the cells and the bacterial pellets were resuspended in PBS (140 mM NaCl, 10 mM Na₂HPO₄, 2.7 mM KCl, 1.8 mM KH₂PO₄ pH 7.3) supplemented with 1 mM DTT, 10% glycerol. After sonication at 277 K, the lysate of the bacteria was centrifuged at 12 000g for 50 min at 277 K and the precipitate was discarded. The supernatant was loaded onto a disposable column containing glutathione Sepharose 4B affinity resin

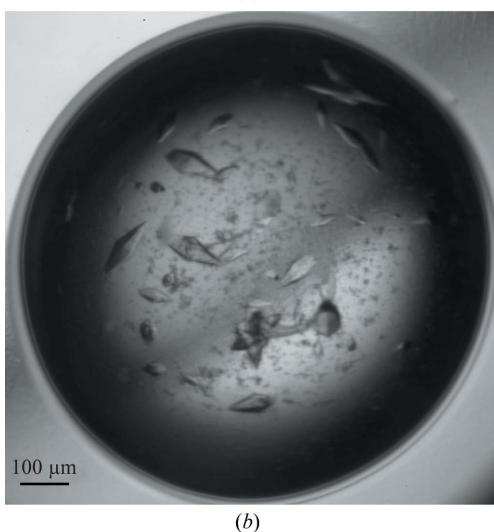
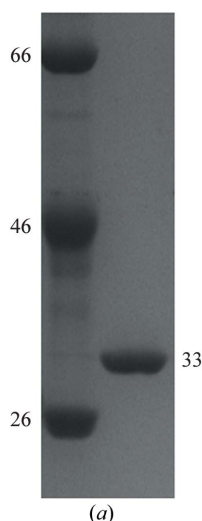


Figure 1

Purification and crystallization of HCoV-NL63 main protease. (a) SDS-PAGE analysis of purified HCoV-NL63 main protease. The molecular masses of the marker and HCoV-NL63 main protease are indicated in kDa. (b) Typical complex crystals of HCoV-NL63 main protease with inhibitor N3 grown by the hanging-drop method. These crystals with typical dimensions of 0.1 × 0.05 × 0.03 mm were used for subsequent diffraction and data collection.

Table 1

Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

Wavelength (Å)	1.0000
Resolution (Å)	50.0–2.85 (2.90–2.85)
Space group	<i>P</i> 4 ₁ 2 ₁ 2
Unit-cell parameters (Å, °)	<i>a</i> = 87.2, <i>b</i> = 87.2, <i>c</i> = 212.1, $\alpha = \beta = \gamma = 90$
No. of observed reflections	283898
No. of unique reflections	19967
Matthews coefficient (Å ³ Da ⁻¹)	3.06
Solvent content (%)	59.8
Average <i>I</i> / σ (<i>I</i>)	21.5 (6.1)
Completeness (%)	100 (100)
Multiplicity	14.2 (12.6)
<i>R</i> _{merge} [†] (%)	14.4 (46.6)
<i>R</i> _{meas} [‡] (%)	14.9 (48.6)
<i>R</i> _{p.i.m.} [§] (%)	4.0 (13.7)

[†] $R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$, where $I_i(hkl)$ is the intensity of the *i*th observation of reflection *hkl* and $\langle I(hkl) \rangle$ is the average intensity. [‡] $R_{\text{meas}} = \frac{\sum_{hkl} \{N(hkl)/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$. [§] $R_{\text{p.i.m.}} = \frac{\sum_{hkl} \{1/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$.

(Pharmacia) for purifying the GST-tagged HCoV-NL63 main protease. The fusion protein was then subjected to on-column cleavage using commercial PreScission protease (Pharmacia) at 277 K for 18 h. The protease was added to a final concentration of 0.25 mg ml⁻¹ and the buffer used for proteolysis was PBS. The resulting protein of interest was further purified by anion-exchange chromatography using a HiTrap Q column (GE Healthcare) in a linear gradient from 25 to 250 mM NaCl with 20 mM Tris-HCl pH 8.0, 10% glycerol, 1 mM DTT and reached greater than 95% purity as determined by SDS-PAGE analysis (Fig. 1a).

2.2. Crystallization

The purified protein was immediately supplemented with 10% DMSO and concentrated to 1 mg ml⁻¹. The previously reported inhibitor N3 (Yang *et al.*, 2005), dissolved in 100% DMSO to a final concentration of 10 mM as a stock, was added to purified protein at a molar ratio between 3:1 and 5:1. After mixing at 277 K for 4 h, the protein complex was centrifuged at 12 000g for 10 min and exchanged into a buffer consisting of 10 mM HEPES pH 7.5, 150 mM NaCl, 1 mM DTT using Thermo iCON concentrators. The final protein was concentrated to 10 mg ml⁻¹ for crystallization. The method described by Birtley & Curry (2005) was used to screen for crystallization conditions for the protein complex. Generally speaking, the reservoir solution was prepared by mixing the original crystallization condition [0.1 M HEPES pH 5.5, 10% (w/v) polyethylene glycol 8000, 4% (v/v) ethylene glycol] for the apo form (PDB entry 3tlo; C. P. Chuck & K. B. Wong, unpublished work) and the commercial Crystal Screen and Crystal Screen 2 kits (Hampton Research, Laguna Niguel, California, USA) in a ratio of 80:20. The hanging-drop vapour-diffusion method was then used to screen for initial crystallization conditions. Crystallization drops were carefully set up by mixing 1.0 μ l protein solution with 1.0 μ l of the reservoir solution above and were then left in 16-well crystallization plates at 291 K to reach equilibrium. Crystals with typical dimensions of 0.1 × 0.05 × 0.03 mm appeared after about 48 h using condition No. 11 of Crystal Screen, which consisted of 0.1 M sodium citrate tribasic dihydrate pH 5.6, 1.0 M ammonium phosphate monobasic. These crystals were used for subsequent diffraction and data collection (Fig. 1b).

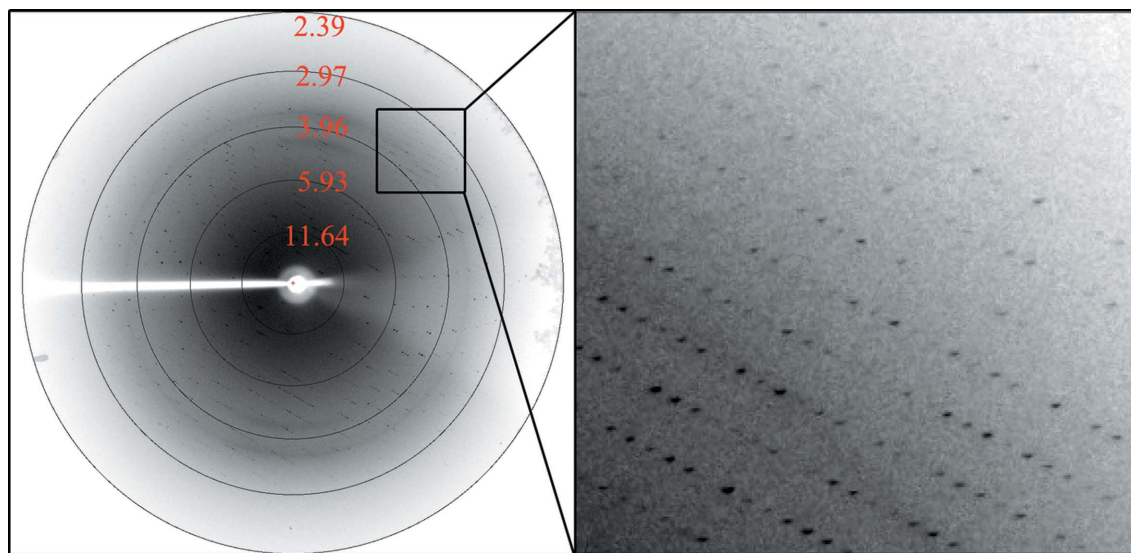


Figure 2

A typical diffraction pattern of the HCoV-NL63 main protease complex crystal collected on a MAR165 charge-coupled device detector on beamline 1W2B at the BSRF. The black circles and numbers correspond to the resolution shells (in Å). The rectangular box shows a diffraction spot in the outer resolution shell.

2.3. X-ray data collection and processing

The crystal was cryoprotected in a solution consisting of 0.1 M HEPES pH 5.5, 10% (w/v) polyethylene glycol 8000, 4% (v/v) ethylene glycol, 20% glycerol and then mounted in a nylon loop and flash-cooled in a nitrogen stream at 100 K using a MAR165 charge-coupled device detector on beamline 1W2B of the Beijing Synchrotron Radiation Facility (BSRF) with a wavelength of 1.0000 Å. The crystals diffracted to 2.85 Å resolution (Fig. 2). All intensity data were indexed, integrated and scaled with the *HKL-2000* package (Otwinowski & Minor, 1997). The R_{merge} and completeness in the outer resolution shell were less than 50% and greater than 90%, respectively. The diffraction data set was firstly indexed in space group *P4*. Subsequent calculation of the self-rotation function with *MOLREP* (Vagin & Teplyakov, 2010) using diffraction data between 10 and 3 Å resolution indicated an additional twofold symmetry axis (Fig. 3). The data set was then processed as *P422* and evaluated using *phenix.xtriage* (Adams *et al.*, 2010); no twinning or pseudo-translational symmetry were detected. Further aided by the systematic absences, a twofold screw axis along the *a* axis and a fourfold screw axis along the *c* axis were identified. The following molecular-replacement search over all alternative subgroups of *P422* helped to determine the space group as *P4₁2₁2* with a TFZ score of 30.5. Based on the molecular weight of the protomer, the Matthews coefficient (Matthews, 1968) was calculated to be 3.06 Å³ Da⁻¹ and the solvent content was 59.8%, with two molecules per asymmetric unit. A final diffraction data set was collected to 2.85 Å resolution and the related data-collection and processing statistics are summarized in Table 1.

3. Results and discussion

Initially, we crystallized apo-form HCoV-NL63 main protease based on the reported crystallization condition (PDB entry 3tlo) and tried soaking it with the inhibitor N3. However, most of the crystals cracked. Although the remaining crystals, which seemed to be intact, were able to produce high-resolution diffraction patterns, the collected data images could not be indexed. We then turned to an alternative method to co-crystallize the protease with the inhibitor

N3 (Birtley & Curry, 2005). The crystals of the protease complex belonged to space group *P4₁2₁2*, with unit-cell parameters $a = b = 87.2$, $c = 212.1$ Å and two molecules per asymmetric unit, corresponding to a Matthews coefficient and solvent content of 3.06 Å³ Da⁻¹ and 59.8%, respectively. Further structural and functional analysis of the HCoV-NL63 main protease in complex with the Michael acceptor N3 will lead to better design and optimization of this lead drug against HCoV-NL63-associated diseases.

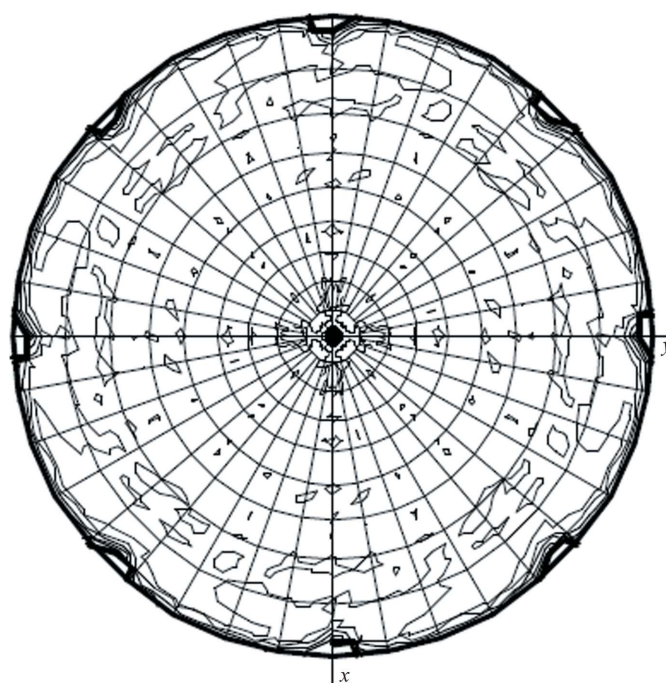


Figure 3

Stereographic projection of the self-rotation function of the *P4* diffraction data set for $\kappa = 180^\circ$ calculated for data between 10 and 3 Å resolution with a radius of integration of 30 Å using *MOLREP* from the *CCP4* suite (Winn *et al.*, 2011). Eight coplanar peaks, 45° apart from each other, are indicated.

We would like to thank Zengqiang Gao and Tianyi Zhang for their help with data collection on beamline 1W2B at the Beijing Synchrotron Radiation Facility (BSRF). This work was supported by the National Natural Science Foundation of China (31300150 and 31370735), the Specialized Research Fund for the Doctoral Program of Higher Education of China (20130032120090) and Tianjin Municipal Natural Science Foundation (General Program: 13JCYBJC42500).

References

- Abdul-Rasool, S. & Fielding, B. C. (2010). *Open Virol. J.* **4**, 76–84.
- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Anand, K., Yang, H., Bartlam, M., Rao, Z. & Hilgenfeld, R. (2005). *Coronaviruses with Special Emphasis on First Insights Concerning SARS*, edited by A. Schmidt, M. H. Wolff & O. Weber, pp. 173–199. Basel: Birkhäuser Verlag.
- Birtley, J. R. & Curry, S. (2005). *Acta Cryst.* **D61**, 646–650.
- Fielding, B. C. (2011). *Future Microbiol.* **6**, 153–159.
- Hoek, L. van der, Pirc, K. & Berkhout, B. (2006). *FEMS Microbiol. Rev.* **30**, 760–773.
- Hoek, L. van der, Pirc, K., Jebbink, M. F., Vermeulen-Oost, W., Berkhout, R. J., Wolthers, K. C., Wertheim-van Dillen, P. M., Kaandorp, J., Spaargaren, J. & Berkhout, B. (2004). *Nature Med.* **10**, 368–373.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Pirc, K., Berkhout, B. & van der Hoek, L. (2007). *J. Virol.* **81**, 3051–3057.
- Pirc, K., Jebbink, M. F., Berkhout, B. & van der Hoek, L. (2004). *Virol. J.* **1**, 7.
- Vabret, A., Mourez, T., Dina, J., van der Hoek, L., Gouarin, S., Petitjean, J., Brouard, J. & Freymuth, F. (2005). *Emerg. Infect. Dis.* **11**, 1225–1229.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Yang, H. *et al.* (2005). *PLoS Biol.* **3**, e324.
- Yang, H., Yang, M., Ding, Y., Liu, Y., Lou, Z., Zhou, Z., Sun, L., Mo, L., Ye, S., Pang, H., Gao, G. F., Anand, K., Bartlam, M., Hilgenfeld, R. & Rao, Z. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 13190–13195.