Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Expression, crystallization and preliminary crystallographic study of mouse hepatitis virus (MHV) nucleocapsid protein C-terminal domain

Xiaohang Tong, Yanlin Ma and Xuemei Li*

National Laboratory of Macromolecules, Institute of Biophysics, Chinese Academy of Science, Beijing 100101, People's Republic of China

Correspondence e-mail: lixm@sun5.ibp.ac.cn

Received 17 March 2010
Accepted 2 April 2010

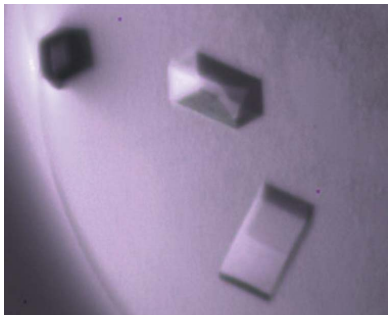
Mouse hepatitis virus (MHV) belongs to the group II coronaviruses. The virus produces nine genes encoding 11 proteins that could be recognized as structural proteins and nonstructural proteins and are crucial for viral RNA synthesis. The nucleocapsid (N) protein, one of the structural proteins, interacts with the 30.4 kb virus genomic RNA to form the helical nucleocapsid and associates with the membrane glycoprotein *via* its C-terminus to stabilize virion assembly. Here, the expression and crystallization of the MHV nucleocapsid protein C-terminal domain are reported. The crystals diffracted to 2.20 Å resolution and belonged to space group *P*422, with unit-cell parameters $a = 66.6$, $c = 50.8$ Å. Assuming the presence of two molecules in the asymmetric unit, the solvent content is 43.0% ($V_M = 2.16$ Å^3 Da^{-1}).

1. Introduction

Coronaviruses are large enveloped positive single-stranded RNA viruses which belong to the *Coronaviridae* family of the *Nidovirales* order. Coronaviruses have been identified as the causative agent of many animal and human diseases (Rota *et al.*, 2003). On the basis of antigenic cross-reactivity and sequence similarity, coronaviruses have been assigned into three groups, with human coronavirus 229E (HCoV-229E), mouse hepatitis virus (MHV) and avian infectious bronchitis virus (IBV) being representatives of groups I, II and III, respectively. MHV, which causes liver or neuron infections in mice, was the best-studied coronavirus (both *in vivo* and *in vitro*) prior to the 2003 SARS outbreak.

MHV (Bost *et al.*, 2000) contains a 31.4 kb positive-sense ssRNA genome (Lai & Stohlman, 1978; Sturman & Holmes, 1983). The virus produces nine genes encoding 11 proteins that could be recognized as structural proteins and non-structural proteins (nsps) and are crucial for viral RNA synthesis (Sturman & Holmes, 1983). The genomic RNA is encapsidated by the nucleocapsid (N) protein into the capsid core. The other four structural proteins, consisting of the spike (S), membrane (M), envelope (E) and haemagglutinin-esterase (HE) proteins, surround the capsid core to form the crown-like viral particles (Sturman & Holmes, 1983).

The MHV-A59 N protein (MHV-N) is well conserved across the various MHV strains. It interacts with the 30.4 kb genomic RNA to form the helical nucleocapsid (Almazán *et al.*, 2004; Baric *et al.*, 1988; Macneughton & Davies, 1978; Robbins *et al.*, 1986; Sawicki *et al.*, 2005) and associates with the membrane glycoprotein *via* its C-terminal part to stabilize virion assembly (Bednar *et al.*, 2006; Hurst *et al.*, 2005; Kuo & Masters, 2002; Verma *et al.*, 2006). It is also considered to be an RNA chaperone (Mir & Panganiban, 2006; Zúñiga *et al.*, 2007). The N protein of MHV-A59 has been determined to be a highly basic phosphoprotein with a molecular weight of 55 kDa that can be subdivided into three conserved domains: domains I (residues Met1–Ala139) and II (residues Asp163–Gln380) are basic, while the C-terminal domain III (residues Glu406–Val454) is acidic. A general RNA-binding region is located at residues



His136–Arg397 (Cologna *et al.*, 2000; Masters, 1992; Mir & Panganiban, 2006; You *et al.*, 2007), while the conserved negatively charged amino acids in domain III are believed to play an important role in N–M protein interactions during assembly (Bednar *et al.*, 2006; Hurst *et al.*, 2005; Kuo & Masters, 2002; Verma *et al.*, 2006). To date, no complete nucleocapsid protein structure has been reported, although the structure of the N-terminal domain (NTD; PDB code 3hd4; Grossoehme *et al.*, 2009) of MHV N protein has been determined. However, the structure of the C-terminal domain (CTD) remains elusive. This work will help to finally determine the structure of the CTD.

2. Expression and purification

The gene encoding the MHV N protein (MHV-N) was amplified by polymerase chain reaction (PCR) from strain MHV-A59 (located at nucleotides 29 669–31 033 in the genome). Following this, the gene for the CTD (N282–397) of MHV-N, comprised of nucleotides 30 514–30 859, was subcloned for protein expression and crystallization. The PCR for CTD was performed with the primer pairs 5′-CGC GGA TCC CCA GTG CAG CAG TGT TTT GGA AAG-3′ and 5′-CG CTC GAG TTA ACG CCC TTT TCT TTG GGG CTT TG-3′. The PCR strategy introduced a *Bam*HI site *via* the forward primer and an *Xho*I site in the reverse primer such that the PCR products could be inserted into the pGEX-6p-1 vector (Amersham Biosciences) using T4 ligase.

The resulting plasmids were subsequently transformed into *Escherichia coli* strain BL21 (DE3). For each plasmid, a well isolated colony was transferred into 5 ml LB medium containing 0.1 mg ml^{-1} ampicillin and incubated at 310 K overnight. The cell culture was further grown at 310 K in 2×YT medium supplemented with ampicillin (0.1 mg ml^{-1}) until the cells reached an OD_{600} of 0.8. Protein expression was induced by the addition of 0.4 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) for another 16 h at 289 K.

Cells were harvested and then lysed by mild sonication in 1×PBS (phosphate-buffered saline; 140 m*M* NaCl, 2.7 m*M* KCl, 10 m*M* Na_2HPO_4 and 1.8 m*M* KH_2PO_4 pH 7.3). The supernatant containing the recombinant glutathione *S*-transferase (GST) fusion proteins, GST-CTD, was applied onto a Glutathione Sepharose 4B (Pharmacia) column, which was followed by on-bead cleavage with PreScission protease (Amersham Biosciences). After cleavage, five foreign residues (GPLGS) remained at the N-terminus of the target protein. Following cleavage, the eluate, which contains the protein of interest, was purified using two columns: a Resource S ion-exchange column and a Superdex 75 10/30 gel-filtration column (Pharmacia). SDS–PAGE analysis showed that the protein purity was approximately greater than 90% by visual estimation and that its apparent molecular mass was consistent with the theoretical molecular mass of CTD (12.6 kDa). The purified CTD was concentrated to 5 mg ml^{-1} using a spin filter for crystallization.

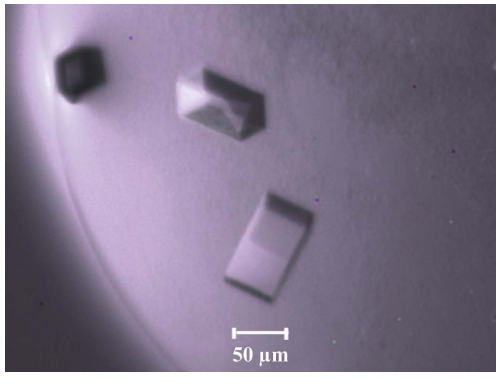


Figure 1
Typical crystals of MHV-N CTD. Maximum crystal dimensions are about 0.1 × 0.1 × 0.05 mm.

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

Space group	*P*422
Unit-cell parameters (Å, °)	a = 66.6, c = 50.8
Wavelength (Å)	1.0000
Resolution (Å)	50.00–2.20 (2.28–2.20)
Completeness (%)	98.1 (100)
Total No. of reflections	76746 (8069)
No. of unique reflections	6102 (606)
Average $I/\sigma(I)$	32.5 (14.8)
$R_{\rm merge}$† (%)	8.30 (31.4)
Matthews coefficient (Å^3 Da^{-1})	2.16

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean of the observations $I_i(hkl)$ of reflection hkl.

3. Crystallization

Crystals of MHV-N CTD were grown at 289 K using the hanging-drop vapour-diffusion method. 1 µl protein solution at a concentration of 5 mg ml^{-1} in stabilizing buffer (20 m*M* MES pH 6.5, 150 m*M* NaCl) was mixed with 1 µl well solution and equilibrated against 200 µl well solution. The Crystal Screen reagent kits (Hampton Research) were applied for initial crystallization screening. Protein crystallization was performed at 291 K in 16-well plates using the hanging-drop vapour-diffusion method.

Initial crystals of CTD were observed using condition No. 28 from Hampton Research Crystal Screen 2, which consists of 1.6 *M* sodium citrate pH 6.5. The condition was optimized by low-gradient variation of the precipitant concentration. A condition consisting of 1.3 *M* sodium citrate pH 6.5 was chosen as the optimum condition for seeding trials as it led to a small improvement in crystal quality. The seeding method involved microseeding combined with streak-seeding. In brief, the protocol used was as follows. We picked up some initial crystals and suspended them as a slurry in mother liquor (1.6 *M* sodium citrate pH 6.5). We then crushed them by vibration into invisible seeds. The seed slurry was diluted to 0.5×, 0.1×, 0.05× and 0.01× concentrations in order to form a suitable concentration-screening matrix. We then dipped a cat whisker into the diluted seed slurry and streaked the whisker across a pre-equilibrated drop under the optimum condition (1.3 *M* sodium citrate pH 6.5). This procedure was repeated for all the different concentrations of seed slurry. Finally, high-quality well shaped crystals were obtained (Fig. 1).

4. Data collection and processing

Prior to data collection, the crystals were plunged into liquid nitrogen and transported to the cold nitrogen stream of beamline 17A at the Photon Factory synchrotron facility (Tsukuba, Japan). Diffraction data were collected on an ADSC Q270 detector at 100 K at a wavelength of 1.0000 Å and were processed to 2.20 Å resolution using the *HKL*-2000 program suite (Otwinowski & Minor, 1997); the crystal belonged to space group *P*422, with unit-cell parameters a = 66.6, c = 50.8 Å. No potential rotational symmetry was revealed on checking the diffraction extinctions. Based on the molecular weight of the monomer, the Matthews coefficient (Matthews, 1968)

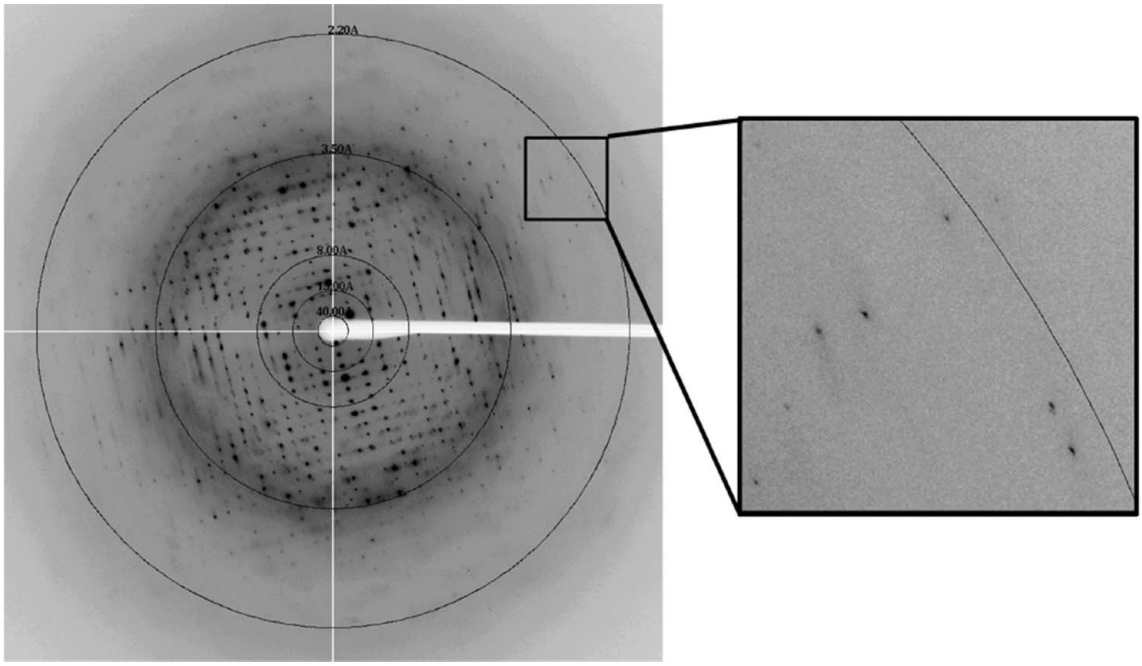


Figure 2
A typical diffraction pattern of an MHV-N CTD crystal. The exposure time was 2 s, the crystal-to-detector distance was 180.0 mm and the oscillation range per frame was 1°. The diffraction image was collected on an ADSC Q270 detector. An enlarged image is shown on the right.

was calculated to be 2.16 Å^3 Da^{-1} and the solvent content was 43.0%, assuming the presence of two molecules per asymmetric unit. Data-collection statistics are shown in Table 1. A typical diffraction image of the MHV-N CTD crystal is shown in Fig. 2.

This work was supported by the National Natural Science Foundation of China (NSFC; grant No. 30730022), the Ministry of Science and Technology (MOST) 973 Project (grant No. 2006CB806503) and 863 Project (grant Nos. 2006AA02A322 and 2006AA020502) and CAS (China) grant KSCX2-YW-R-05 to ZR.

References

Almazán, F., Galán, C. & Enjuanes, L. (2004). *J. Virol.* **78**, 12683–12688.
Baric, R. S., Nelson, G. W., Fleming, J., Deans, R. J., Keck, J. G., Casteel, N. & Stohlman, S. A. (1988). *J. Virol.* **62**, 4280–4287.
Bednar, V., Verma, S., Blount, A. & Hogue, B. G. (2006). *Adv. Exp. Med. Biol.* **581**, 127–132.
Bost, A., Carnahan, R., Lu, X. & Denison, M. (2000). *J. Virol.* **74**, 3379–3387.
Cologna, R., Spagnolo, J. & Hogue, B. (2000). *Virology*, **277**, 235–249.
Grossoehme, N., Li, L., Keane, S., Liu, P., Dann, C. III, Leibowitz, J. & Giedroc, D. (2009). *J. Mol. Biol.* **394**, 544–557.
Hurst, K., Kuo, L., Koetzner, C., Ye, R., Hsue, B. & Masters, P. (2005). *J. Virol.* **79**, 13285–13297.
Kuo, L. & Masters, P. (2002). *J. Virol.* **76**, 4987–4999.
Lai, M. M. & Stohlman, S. A. (1978). *J. Virol.* **26**, 236–242.
Macneughton, M. & Davies, H. (1978). *J. Gen. Virol.* **39**, 545–549.
Masters, P. S. (1992). *Arch. Virol.* **125**, 141–160.
Matthews, B. (1968). *J. Mol. Biol.* **33**, 491–497.
Mir, M. & Panganiban, A. (2006). *J. Virol.* **80**, 6276–6285.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Robbins, S., Frana, M., McGowan, J., Boyle, J. & Holmes, K. (1986). *Virology*, **150**, 402–410.
Rota, P. *et al.* (2003). *Science*, **300**, 1394–1399.
Sawicki, S., Sawicki, D., Younker, D., Meyer, Y., Thiel, V., Stokes, H. & Siddell, S. (2005). *PLoS Pathog.* **1**, e39.
Sturman, L. S. & Holmes, K. V. (1983). *Adv. Virus Res.* **28**, 35–112.
Verma, S., Bednar, V., Blount, A. & Hogue, B. G. (2006). *J. Virol.* **80**, 4344–4355.
You, J.-H., Reed, M. L. & Hiscox, J. A. (2007). *Biochem. Biophys. Res. Commun.* **358**, 1015–1020.
Zúñiga, S., Sola, I., Moreno, J., Sabella, P., Plana-Durán, J. & Enjuanes, L. (2007). *Virology*, **357**, 215–227.