Acta Crystallographica Section F
Structural Biology Communications
ISSN 2053-230X

Crystallization and preliminary crystallographic study of *Porcine epidemic diarrhea virus* main protease in complex with an inhibitor

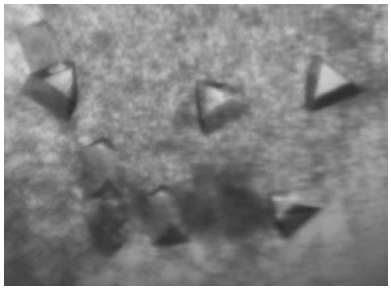
Yusheng Tan,[a]‡ Fenghua Wang,[a,b,c]‡ Xia Chen,[a,b,c] Jinshan Wang,[a,b,c] Qi Zhao,[d] Shuang Li,[b] Zefang Wang,[a] Sheng Fu,[b]* Cheng Chen[a,b]* and Haitao Yang[a,b]

[a]School of Life Sciences, Tianjin University, Tianjin 300072, People's Republic of China, [b]Tianjin International Joint Academy of Biotechnology and Medicine, Tianjin 300457, People's Republic of China, [c]College of Life Sciences, Nankai University, Tianjin 300071, People's Republic of China, and [d]Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, CT 06520, USA

‡ These authors contributed equally to this work.

Correspondence e-mail: fusheng@tjab.org, chengchen@tju.edu.cn

Received 16 August 2014
Accepted 3 October 2014



Porcine epidemic diarrhea virus (PEDV) mainly infects neonatal pigs, resulting in significant morbidity and mortality. Owing to problems such as long periods of virus shedding, existing vaccines cannot provide complete protection from PEDV infection. The PEDV genome 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 *Porcine epidemic diarrhea virus* in complex with a Michael acceptor was crystallized. The complex crystals diffracted to 2.5 Å resolution and belonged to space group *R*3, with unit-cell parameters a = 175.3, b = 175.3, c = 58.7 Å. Two molecules were identified per asymmetric unit.

1. Introduction

Porcine epidemic diarrhea virus (PEDV) mainly infects pigs and causes porcine epidemic diarrhoea (PED), a devastating enteric disease with clinical manifestations such as acute diarrhoea and dehydration (Wood, 1977). In particular, it results in significant morbidity and mortality for neonatal piglets (Hwang *et al.*, 1994). PEDV was first discovered in England in 1978 (Pensaert & de Bouck, 1978; Debouck & Pensaert, 1980). It then spread throughout Europe during the 1980s and 1990s. Currently, it is circulating in Asia, especially in those countries where the pork industry is prevalent, such as the Philippines, South Korea and China (Song & Park, 2012). Outbreaks in these countries have been found to be more acute and severe than those observed in Europe (Song & Park, 2012). To date, serious economic loss has been caused by PEDV. Although several vaccines have been developed against PEDV, they cannot provide complete protection from PEDV infection owing to problems such as long periods of virus shedding (Yuan *et al.*, 1998). Similar to other coronaviruses (CoVs), PEDV contains a single-stranded positive-sense polyadenylated RNA genome that encodes two large polyproteins (pp1a and pp1ab), which need to be processed into 16 nonstructural proteins (nsp1–16) for genome replication (Brian & Baric, 2005). This process is mediated by two virus-encoded proteinases. Nsp5, also named the main protease (M^{pro}), is responsible for 11 out of 15 cleavage sites, thus playing a pivotal role in this digestion process and being indispensable for viral replication (Pyrc *et al.*, 2004, 2007; Brian & Baric, 2005; van der Hoek *et al.*, 2006). The critical role of nsp5 in virus replication makes it an ideal target for anti-PEDV drug design (Yang *et al.*, 2005; Brian & Baric, 2005; Anand *et al.*, 2003). The crystal structures of the main proteases from *Human coronavirus 229E* (HCoV-229E; Anand *et al.*, 2003), *Transmissible gastroenteritis virus* (TGEV; Anand *et al.*, 2002), SARS-CoV (Yang *et al.*, 2003; Xue *et al.*, 2007) and *Infectious bronchitis virus* (IBV; Xue *et al.*, 2008) have been solved alone. Based on the structural analysis of these main proteases, the idea of designing inhibitors against CoVs has been proposed (Anand *et al.*, 2002, 2003; 2005; Yang *et al.*, 2003; Bacha *et al.*, 2004; Xue *et al.*, 2007). To date, the crystal structures of CoV main proteases from HCoV-229E (Lee *et al.*, 2009), TGEV (Kim *et al.*, 2012; Yang *et al.*, 2005), *Human coronavirus HKU1* (Zhao *et al.*, 2008), SARS-CoV (Verschueren *et al.*, 2008; Lu *et al.*, 2006; Lee *et al.*, 2005; Yang *et al.*, 2003, 2005; Bacha *et al.*, 2008), IBV (Xue *et al.*, 2008)

Table 1
Macromolecule-production information.

Source organism	PEDV strain CV777 (GenBank accession No. AAK38661)
DNA source	Synthetic DNA
Forward primer†	CGGGATCCGCTGGCTTGCGTAAGAT
Reverse primer‡	CCGCTCGAGTTACTGAAGATTAACGCCAT
Cloning vector	pGEX-6P-1
Expression vector	pGEX-6P-1
Expression host	*E. coli* BL21 (DE3)
Complete amino-acid sequence of the construct produced	GPLGSAGLRKMAQPSGVVEKCIVRVCYGNMALNGLWLGDIVM-CPRHVIASSTTSTIDYDYALSVLRLHNFSISSGNVFLGVV-SATMRGALLQIKVNQNNVHTPKYTYRTVRPGESFNILACY-DGAAAGVYGVNMRSNYTIRGSFINGACGSPGYNINNGTVE-FCYLHQLELGSGCHVGSDLDGVMYGGYEDQPTLQVEGASS-LFTENVLAFLYAALINGSTWWLSSSRIAVDRFNEWAVHNG-MTTVGNTDCFSILAAKTGVDVQRLLASIQSLHKNFGGKQI-LGHTSLTDEFTTGEVVRQMYGVNLQ

† Underlined sequence: *Bam*HI site. ‡ Underlined sequence: *Xho*I site.

and *Middle east respiratory syndrome coronavirus* (Ren *et al.*, 2013) have been solved in complex with inhibitors. In this study, we report the crystallization and preliminary crystallographic study of PEDV main protease in complex with a designed Michael acceptor inhibitor named N3 (Yang *et al.*, 2005; Supplementary Fig. S1[1]). Michael acceptors are a class of conjugated α,β-unsaturated carbonyl compounds that have been successfully introduced to devise irreversible covalent inhibitors of cysteine proteases from viruses such as *Human enterovirus 68* (Tan *et al.*, 2013), rhinoviruses (Matthews *et al.*, 1999) and coronaviruses (Yang *et al.*, 2005).

2. Materials and methods

2.1. Macromolecule production

The coding sequence for PEDV main protease was synthesised and cloned into the vector pGEX-6P-1 (GE Healthcare) using the *Bam*HI and *Xho*I restriction sites (Table 1). The recombinant plasmid was verified by sequencing and then transformed into *Escherichia coli* strain BL21 (DE3) cells 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 m*M* and the cultures were induced to express PEDV 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 m*M* NaCl, 10 m*M* Na_2HPO_4, 2.7 m*M* KCl, 1.8 m*M* KH_2PO_4 pH 7.3) supplemented with 1 m*M* dithiothreitol (DTT) and 10% glycerol. After sonication at 277 K, the bacterial lysate was centrifuged at 12 000*g* 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 (Pharmacia) for purification of the GST-tagged PEDV 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} for proteolysis in PBS. Five additional residues (GPLGS) were left at the N-terminus of the PEDV main protease. The resulting protein of interest was further purified by anion-exchange chromatography using a HiTrap Q column (GE Healthcare) with a linear gradient from 25 to 250 m*M* NaCl in 20 m*M* Tris–HCl pH 8.0, 10% glycerol, 1 m*M* DTT and reached greater than 90% purity by SDS–PAGE analysis (Fig. 1*a*). Typical yields were 5 mg of purified protein per litre of bacterial culture.

[1] Supporting information has been deposited in the IUCr electronic archive (Reference: DP5077).

2.2. Crystallization

The purified protein was immediately supplemented with 10% dimethyl sulfoxide (DMSO) and concentrated to 1 mg ml^{-1} using Thermo iCON concentrators. The protein concentration was determined by ultraviolet–visible spectrophotometry using a NanoDrop 2000 spectrophotometer (Thermo Scientific, USA). The previously reported inhibitor N3 (Yang *et al.*, 2005), dissolved in 100% DMSO to a final concentration of 10 m*M* as a stock, was then added to the purified protein at a molar ratio from 3:1 to 5:1. After mixing at 277 K for 4 h, the protein complex was centrifuged at 12 000*g* for 10 min and exchanged into a buffer consisting of 10 m*M* HEPES pH 7.5, 150 m*M* NaCl, 1 m*M* DTT using concentrators. The final protein was concentrated to 10 mg ml^{-1} for crystallization. In the initial stage, commercial screening kits including Crystal Screen, Crystal Screen 2, Index, PEGRx and PEGRx 2 (Hampton Research, Laguna Niguel, California, USA) were used to screen for preliminary crystallization conditions for PEDV main protease in complex with N3. Crystallization trials were set up in 16-well crystallization plates (Xiangyushun, People's Republic of China) at 291 K using the hanging-drop vapour-diffusion method. Crystallization drops were carefully set up by mixing 1.0 µl protein solution with 1.0 µl reservoir solution and were then left to equilibrate against 200 µl reservoir solution. Initial

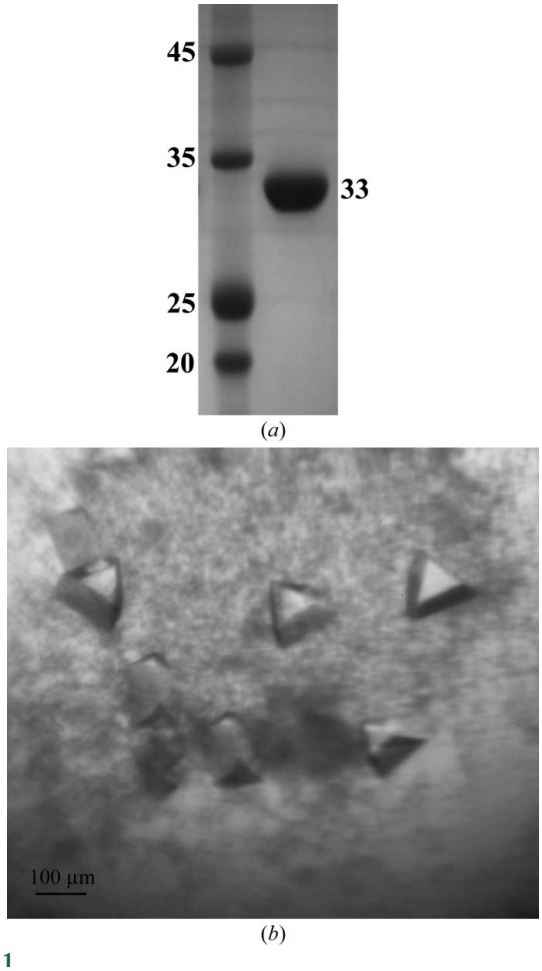


Figure 1
Purification and crystallization of PEDV main protease. (*a*) SDS–PAGE analysis of purified PEDV main protease. The molecular masses of the marker and PEDV main protease are indicated in kDa. (*b*) Typical crystals of PEDV main protease complexed with the inhibitor N3 grown by the hanging-drop method. Crystals with typical dimensions of 0.08 × 0.08 × 0.05 mm were used for subsequent X-ray diffraction tests and data collection.

Table 2
Crystallization.

Method	Hanging-drop vapour diffusion
Plate type	16-well crystallization plates
Temperature (K)	291
Protein concentration (mg ml^{-1})	10
Buffer composition of protein solution	10 m*M* HEPES pH 7.5, 150 m*M* NaCl, 1 m*M* DTT
Composition of reservoir solution	25–29%(*v*/*v*) 2-propanol, 0.1 *M* Tris pH 7.5–7.9, 5%(*w*/*v*) polyethylene glycol 8000
Volume and ratio of drop	1.0 µl + 1.0 µl
Volume of reservoir (µl)	200

crystals of PEDV main protease in complex with N3 were obtained after 24 h using condition No. 38 of PEGRx 2, which consisted of 20%(*v*/*v*) 2-propanol, 0.1 *M* Tris pH 8.0, 5%(*w*/*v*) polyethylene glycol 8000. The optimized condition consisted of 25–29%(*v*/*v*) 2-propanol, 0.1 *M* Tris pH 7.5–7.9, 5%(*w*/*v*) polyethylene glycol 8000 (Fig. 1*b*). The crystallization information is summarized in Table 2.

2.3. X-ray data collection and processing

The crystals were cryoprotected in a solution consisting of 25%(*v*/*v*) 2-propanol, 0.1 *M* Tris pH 7.7, 5%(*w*/*v*) polyethylene glycol 8000, 20% glycerol and were then mounted in a nylon loop and flash-cooled in a nitrogen stream at 100 K. The X-ray diffraction data sets were collected using an ADSC Q315r detector on beamline BL-5A of the Photon Factory, KEK, Japan at a wavelength of 1.0000 Å. The crystals showed high-quality diffraction patterns after annealing for about 30 s (Fig. 2). The annealing procedure was performed following a previously described protocol (Yeh & Hol, 1998; Kriminski *et al.*, 2002). Briefly, the cold nitrogen was first diverted. The crystal was left at room temperature to anneal for about 30 s and subjected to data collection. All intensity data were indexed, integrated and scaled with the *HKL*-2000 package (Otwinowski & Minor, 1997). A complete data set was collected to 2.5 Å resolution and the data-collection and processing statistics are summarized in Table 3.

Table 3
Data-collection and processing statistics.

Values in parentheses are for the outer shell.

Diffraction source	Beamline BL-5A, Photon Factory
Wavelength (Å)	1.0000
Temperature (K)	100
Detector	ADSC Q315r
Crystal-to-detector distance (mm)	369
Rotation range per image (°)	0.5
Total rotation range (°)	180
Space group	*R*3
Unit-cell parameters (Å, °)	$a = b = 175.3$, $c = 58.7$, $\alpha = \beta = 90$, $\gamma = 120$
Mosaicity (°)	0.74
Resolution range (Å)	50.0–2.44 (2.48–2.44)
Total No. of reflections	144327 (5570)
No. of unique reflections	24884 (1211)
Completeness (%)	100.0 (99.2)
Multiplicity	5.8 (4.6)
$\langle I/\sigma(I)\rangle$	32.6 (5.3)
$R_{\rm merge}$† (%)	5.2 (29.9)
$R_{\rm meas}$‡ (%)	5.7 (33.7)

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

3. Results and discussion

The crystals were initially directly subjected to X-ray diffraction validation. However, only weak diffraction could be obtained. We then used the annealing method to improve the diffraction quality of the complex crystals (Yeh & Hol, 1998; Kriminski *et al.*, 2002). The crystals finally diffracted to a highest resolution of of 2.5 Å using 25%(*v*/*v*) 2-propanol, 0.1 *M* Tris pH 7.7, 5%(*w*/*v*) polyethylene glycol 8000, 20% glycerol as a cryoprotectant. The crystals belonged to space group *R*3, with unit-cell parameters $a = 175.3$, $b = 175.3$, $c = 58.7$ Å. The data set was evaluated by *phenix.xtriage* (Adams *et al.*, 2010) and no twinning or pseudo-translational symmetry was detected. Subsequent calculation of the self-rotation function with *MOLREP* (Vagin & Teplyakov, 2010) using diffraction data between

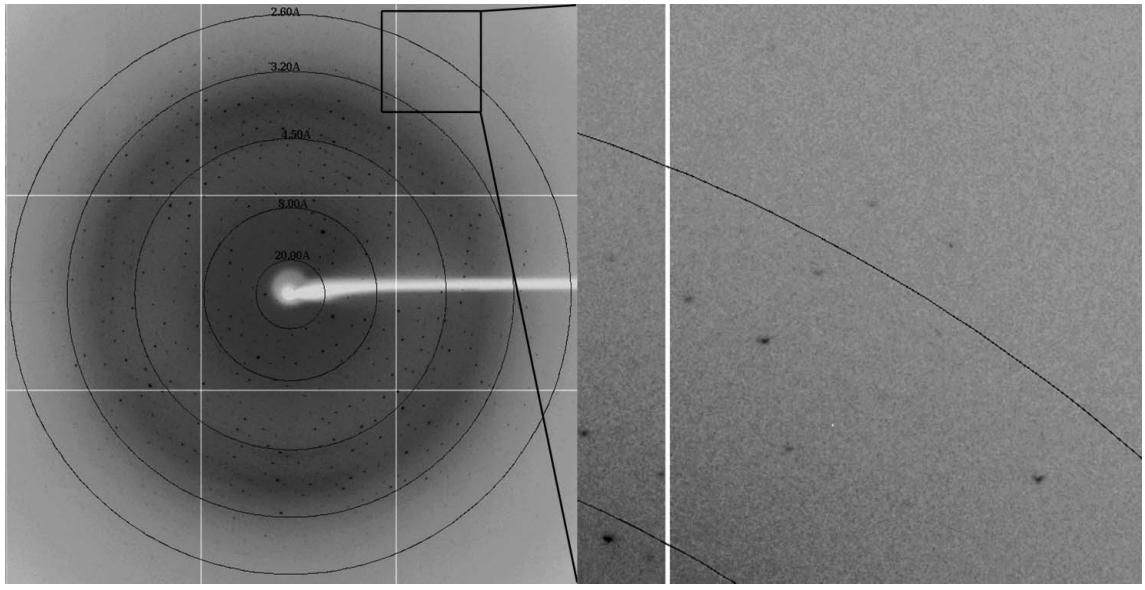


Figure 2
A typical diffraction pattern from a crystal of the PEDV main protease complex collected on beamline BL-5A of the Photon Factory, KEK, Japan. The edge of the frame is at 2.44 Å resolution. Diffraction spots in the outer resolution shell are indicated by the box.

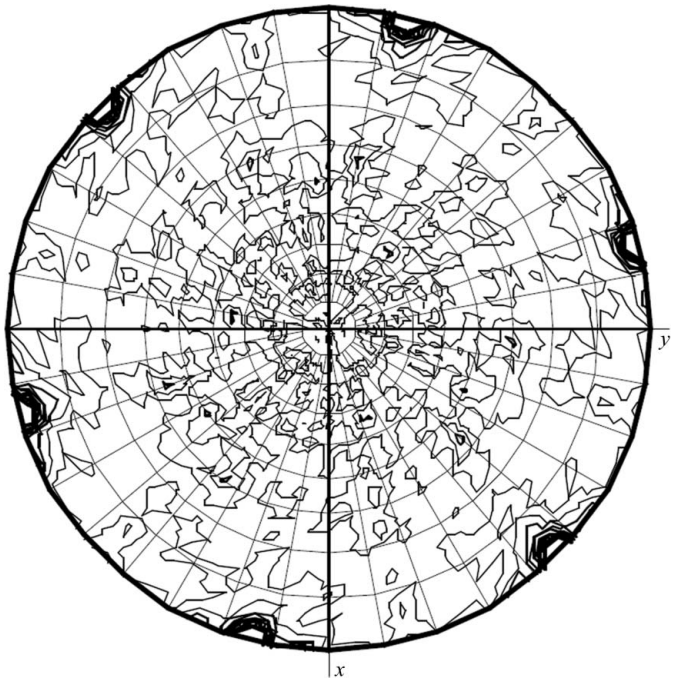


Figure 3
Stereographic projection of the self-rotation function of the diffraction data set for $\kappa = 180^\circ$ calculated for data between 10 and 3 Å resolution at a radius of integration of 30 Å using *MOLREP* from the *CCP*4 suite. The *x* and *z* axes of the plot align with the *a* and *c** crystallographic axes, respectively, and the *xy* plane aligns with the *ab* crystallographic plane.

10 and 3 Å resolution indicated that there was a twofold noncrystallographic symmetry axis (Fig. 3). This axis is approximately parallel to the crystal axis. The angle between these two axes is about 15°. Among the available CoV main protease structures (Anand *et al.*, 2002, 2003; Chuck *et al.*, 2013; Kim *et al.*, 2012; Lee *et al.*, 2005, 2009; Lu *et al.*, 2006; Ren *et al.*, 2013; Bacha *et al.*, 2008; Verschueren *et al.*, 2008; Xue *et al.*, 2007, 2008; Yang *et al.*, 2003, 2005; Zhao *et al.*, 2008; Zhu *et al.*, 2011), the primary sequence of HCoV-229E main protease (PDB entries 2zu2 and 1p9s; Lee *et al.*, 2009; Anand *et al.*, 2003) shows 70% identity to the PEDV main protease, indicating that the structure of HCoV-229E main protease may serve as a suitable search model to solve the structure of the PEDV main protease by molecular replacement. Based on the molecular weight of the monomer, the Matthews coefficient (Matthews, 1968) was calculated to be 2.62 $Å^3\ Da^{-1}$ and the solvent content to be 53.1%, assuming the presence of two molecules per asymmetric unit. Further structural and biochemical analysis of the PEDV main protease in complex with the Michael acceptor N3 will lead to better design and optimization of an antiviral against PED.

We would like to thank Dunquan Jiang for his help with data collection on beamline BL-5A at the Photon Factory, KEK, Japan. This work was supported by the National Natural Science Foundation of China (31300150), the Specialized Research Fund for the Doctoral Program of Higher Education of China (20130032120090) and the Tianjin Municipal Natural Science Foundation (General Program: 13JCYBJC42500).

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Anand, K., Palm, G. J., Mesters, J. R., Siddell, S. G., Ziebuhr, J. & Hilgenfeld, R. (2002). *EMBO J.* **21**, 3213–3224.
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.
Anand, K., Ziebuhr, J., Wadhwani, P., Mesters, J. R. & Hilgenfeld, R. (2003). *Science*, **300**, 1763–1767.
Bacha, U., Barrila, J., Gabelli, S. B., Kiso, Y., Amzel, L. M. & Freire, E. (2008). *Chem. Biol. Drug Des.* **72**, 34–49.
Bacha, U., Barrila, J., Velazquez-Campoy, A., Leavitt, S. A. & Freire, E. (2004). *Biochemistry*, **43**, 4906–4912.
Brian, D. A. & Baric, R. S. (2005). *Curr. Top. Microbiol. Immunol.* **287**, 1–30.
Chuck, C.-P., Chen, C., Ke, Z., Wan, D. C.-C., Chow, H.-F. & Wong, K.-B. (2013). *Eur. J. Med. Chem.* **59**, 1–6.
Debouck, P. & Pensaert, M. (1980). *Am. J. Vet. Res.* **41**, 219–223.
Hoek, L. van der, Pyrc, K. & Berkhout, B. (2006). *FEMS Microbiol. Rev.* **30**, 760–773.
Hwang, E. K., Kim, J. H., Jean, Y. H., Bae, Y. C., Yoon, S. S., Park, C. K., Kweon, C. H., Yoon, Y. D. & Ackermann, M. (1994). *RDA J. Agric. Soc.* **36**, 587–596.
Kim, Y., Lovell, S., Tiew, K.-C., Mandadapu, S. R., Alliston, K. R., Battaile, K. P., Groutas, W. C. & Chang, K.-O. (2012). *J. Virol.* **86**, 11754–11762.
Kriminski, S., Caylor, C. L., Nonato, M. C., Finkelstein, K. D. & Thorne, R. E. (2002). *Acta Cryst.* D**58**, 459–471.
Lee, C.-C., Kuo, C.-J., Ko, T.-P., Hsu, M.-F., Tsui, Y.-C., Chang, S.-C., Yang, S., Chen, S.-J., Chen, H.-C., Hsu, M.-C., Shih, S.-R., Liang, P.-H. & Wang, A. H.-J. (2009). *J. Biol. Chem.* **284**, 7646–7655.
Lee, T.-W., Cherney, M. M., Huitema, C., Liu, J., James, K. E., Powers, J. C., Eltis, L. D. & James, M. N. G. (2005). *J. Mol. Biol.* **353**, 1137–1151.
Lu, I.-L., Mahindroo, N., Liang, P.-H., Peng, Y.-H., Kuo, C.-J., Tsai, K.-C., Hsieh, H.-P., Chao, Y.-S. & Wu, S.-Y. (2006). *J. Med. Chem.* **49**, 5154–5161.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Matthews, D. A. *et al.* (1999). *Proc. Natl Acad. Sci. USA*, **96**, 11000–11007.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pensaert, M. B. & de Bouck, P. (1978). *Arch. Virol.* **58**, 243–247.
Pyrc, K., Berkhout, B. & van der Hoek, L. (2007). *J. Virol.* **81**, 3051–3057.
Pyrc, K., Jebbink, M. F., Berkhout, B. & van der Hoek, L. (2004). *Virol. J.* **1**, 7.
Ren, Z., Yan, L., Zhang, N., Guo, Y., Yang, C., Lou, Z. & Rao, Z. (2013). *Protein Cell.*, **4**, 248–250.
Song, D. & Park, B. (2012). *Virus Genes*, **44**, 167–175.
Tan, J., George, S., Kusov, Y., Perbandt, M., Anemüller, S., Mesters, J. R., Norder, H., Coutard, B., Lacroix, C., Leyssen, P., Neyts, J. & Hilgenfeld, R. (2013). *J. Virol.* **87**, 4339–4351.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Verschueren, K. H., Pumpor, K., Anemüller, S., Chen, S., Mesters, J. R. & Hilgenfeld, R. (2008). *Chem. Biol.* **15**, 597–606.
Wood, E. N. (1977). *Vet. Rec.* **100**, 243–244.
Xue, X., Yang, H., Shen, W., Zhao, Q., Li, J., Yang, K., Chen, C., Jin, Y., Bartlam, M. & Rao, Z. (2007). *J. Mol. Biol.* **366**, 965–975.
Xue, X., Yu, H., Yang, H., Xue, F., Wu, Z., Shen, W., Li, J., Zhou, Z., Ding, Y., Zhao, Q., Zhang, X. C., Liao, M., Bartlam, M. & Rao, Z. (2008). *J. Virol.* **82**, 2515–2527.
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.
Yeh, J. I. & Hol, W. G. J. (1998). *Acta Cryst.* D**54**, 479–480.
Yuan, L., Kang, S.-Y., Ward, L. A., To, T. L. & Saif, L. J. (1998). *J. Virol.* **72**, 330–338.
Zhao, Q., Li, S., Xue, F., Zou, Y., Chen, C., Bartlam, M. & Rao, Z. (2008). *J. Virol.* **82**, 8647–8655.
Zhu, L., George, S., Schmidt, M. F., Al-Gharabli, S. I., Rademann, J. & Hilgenfeld, R. (2011). *Antiviral Res.* **92**, 204–212.