Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of the X (ADRP) domain of nsp3 from feline coronavirus

Justyna A. Wojdyla,[a] Ioannis Manolaridis,[a] Eric J. Snijder,[b] Alexander E. Gorbalenya,[b] Bruno Coutard,[c] Yvonne Piotrowski,[d] Rolf Hilgenfeld[d] and Paul A. Tucker[a]*

[a]EMBL Hamburg Outstation, c/o DESY, Notkestrasse 85, D-22603 Hamburg, Germany, [b]Molecular Virology Laboratory, Department of Medical Microbiology, Center of Infectious Diseases, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands, [c]Laboratoire Architecture et Fonction des Macromolécules Biologiques, UMR 6098, AFMB–CNRS–ESIL, Case 925, 163 Avenue de Luminy, 13288 Marseille, France, and [d]Institute of Biochemistry, Center for Structural and Cell Biology in Medicine, University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany

Correspondence e-mail: tucker@embl-hamburg.de

Received 24 June 2009
Accepted 1 October 2009

PDB References: FCoV X domain, 3ew5, r3ew5sf; 3eti, r3etisf; ADP-ribose-bound, 3jzt, r3jztsf.

The structure of the X (or ADRP) domain of a pathogenic variant of feline coronavirus (FCoV) has been determined in tetragonal and cubic crystal forms to 3.1 and 2.2 Å resolution, respectively. In the tetragonal crystal form, glycerol-3-phosphate was observed in the ADP-ribose-binding site. Both crystal forms contained large solvent channels and had a solvent content of higher than 70%. Only very weak binding of this domain to ADP-ribose was detected *in vitro*. However, the structure with ADP-ribose bound was determined in the cubic crystal form at 3.9 Å resolution. The structure of the FCoV X domain had the expected macro-domain fold and is the first structure of this domain from a coronavirus belonging to subgroup 1a.

1. Introduction

Coronaviruses are positive-stranded RNA viruses that belong to the order *Nidovirales* (Gorbalenya *et al.*, 2006). Their virion ranges from 80 to 120 nm in diameter and contains one molecule of RNA of approximately 30 kb (Weiss & Navas-Martin, 2005). Transcription and replication of the coronaviral RNA genome takes place in the cytoplasm of the host cell at replication sites that are associated with a characteristic reticulovesicular network of modified endoplasmic reticulum (Stertz *et al.*, 2007; Snijder *et al.*, 2006; Prentice *et al.*, 2004; Knoops *et al.*, 2008). The first two open reading frames (ORF1a and ORF1b) of the viral genome overlap and encode two replicase precursor polyproteins, pp1a and pp1ab. These long polyproteins (∼4000 and ∼7000 amino acids, respectively) are the substrates for the autocatalytic production of 15–16 mature nonstructural proteins (nsps), which is driven by viral proteases: a 3C-like main protease (M^{pro}; nsp5) and one or two (depending on the virus) accessory papain-like proteases ($PL1^{pro}$ and $PL2^{pro}$ in nsp3; Weiss & Navas-Martin, 2005; Ziebuhr *et al.*, 2000). The largest cleavage product derived from pp1a/pp1ab is nsp3. This subunit includes diverse enzymatic and poorly characterized domains that are involved in the replication and expression of the virus genome and virus–host interactions (Thiel *et al.*, 2003; Snijder *et al.*, 2003; Neuman *et al.*, 2008; Kanjanahaluethai *et al.*, 2007).

One of the ubiquitous domains present in nsp3 of all coronaviruses is the X domain [also known as the adenosine diphosphate-ribose-1″-phosphatase (ADRP) domain], which was originally identified as a domain that is conserved in vertebrate RNA viruses of the alphavirus-like supergroup and coronaviruses (Koonin *et al.*, 1992; Gorbalenya *et al.*, 1991). Its

proposed enzymatic activity was based on its distant similarity to a cellular enzyme (Snijder *et al.*, 2003), the activity of which was first identified in yeast proteins (Shull *et al.*, 2005; Martzen *et al.*, 1999; Kumaran *et al.*, 2005) and subsquently in the *Archeoglobus fulgidus* protein AF1521 (Karras *et al.*, 2005). ADRP activity has also been shown for the severe acute respiratory syndrome virus (SARS-CoV), human coronavirus HCoV-229E and porcine transmissible gastroenteritis virus (TGEV) X domains (Putics *et al.*, 2005, 2006). The RNA viral X domains and their cellular homologues, which can be found in eukaryotes, bacteria and archaea, belong to a large structural family prototyped by a macro domain (Pehrson & Fuji, 1998). The human macroH2A1 macro domain has been shown to bind *O*-acetyl-ADP-ribose (Kustatscher *et al.*, 2005) and both the yeast protein and the *A. fulgidus* protein AF1521 can bind ADP-ribose (Karras *et al.*, 2005; Kumaran *et al.*, 2005). It is believed that this activity or the binding of related ligands [for example poly(ADP)-ribose] might determine the function of the macro domain. The X domains of several coronaviruses [SARS-CoV, HCoV-229E and infectious bronchitis virus (IBV) strain M41; Egloff *et al.*, 2006; Xu *et al.*, 2009] and alphaviruses (Malet *et al.*, 2009) have also been shown to bind ADP-ribose. Catalytic phosphatase activity and ADP-ribose binding are two related but distinct properties of the X domain. Despite considerable progress in the *in vitro* characterization of coronaviral X domains, their activity and function *in vivo*, which are apparently linked to virus pathogenesis, remain poorly understood (Eriksson *et al.*, 2008).

Coronaviruses may be divided into several genetic subgroups, the first of which were originally established using serological cross-reactivity (Gorbalenya *et al.*, 2004; Lai & Holmes, 2001). Subgroup 1a includes feline coronavirus (FCoV) and TGEV. HCoV-229E belongs to subgroup 1b, together with HCoV-NL63 and porcine epidemic diarrhoea virus (PEDV). Subgroup 2a contains the murine hepatitis virus MHV and the porcine haemagglutinating encephalomyelitis virus BCoV. In 2003 the new SARS-CoV was classified as member of a new subgroup 2b. Group 3 originally included only avian viruses such as IBV. Very recently, a number of new coronaviruses have been identified that could be prototypes of separate subgroups in groups 2 and 3 (Woo *et al.*, 2009; Mihindukulasuriya *et al.*, 2008).

The X domains of nsp3 from coronaviral subgroups 1b (HCoV-229E; Piotrowski *et al.*, 2009; Xu *et al.*, 2009), 2b (SARS-CoV; Egloff *et al.*, 2006; Saikatendu *et al.*, 2005) and 3 (IBV; Piotrowski *et al.*, 2009; Xu *et al.*, 2009) have been structurally characterized. Here, we report the first crystal structure of the X domain from feline infectious peritonitis virus (FIPV), which belongs to coronaviral subgroup 1a. Feline infectious peritonitis virus (FIPV) is a pathogenic FCoV variant that emerged by mutation of the relatively benign enteric FCoV (Vennema *et al.*, 1998; Poland *et al.*, 1996) and causes a fatal immune-mediated disease in cats (Pedersen, 1995). Since the variations are minor and despite the fact that we are working with a FIPV construct, we will henceforth use FCoV as an abbreviation for the virus. The FCoV X-domain structure was determined in tetragonal and cubic space groups to 3.1 and 2.2 Å resolution, respectively. In addition, the structure with bound ADP-ribose was determined in the cubic crystal form to 3.9 Å resolution. We analyzed the similarity of the FCoV X-domain structure to those of other coronaviral X domains and macro domains from other organisms.

2. Materials and methods

2.1. Cloning

Vector pDEST14 (Invitrogen) containing the X domain of nsp3 (residues 1254–1421 of the polyprotein pp1a) from feline coronavirus (FCoV; strain FIPV WSU-79/1146; GenBank/RefSeq accession No. NC_007025.1) was originally created using Gateway cloning technology (with a C-terminal His_6 tag). The X domain was recloned into the pETM-11 vector (EMBL Hamburg), which allows the removal of an N-terminal His_6 tag with tobacco etch virus (TEV) 3C-like protease. The point mutant Leu122Met was created using the Stratagene QuikChange mutagenesis kit. The nucleotide substitution was confirmed by DNA sequencing (MWG Biotech).

2.2. Expression and purification

Protein expression was performed in *Escherichia coli* strain Rosetta (DE3) pLysS (Novagen). Cultures were grown in LB medium at 310 K until the OD_{600} reached 0.8, induced with 1 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) and left shaking overnight at 288 K. Cells were collected by centrifugation at 4000*g* (30 min, 277 K) and frozen at 253 K.

The bacterial pellet from 1 l cell culture was resuspended in 20 ml lysis buffer [50 m*M* Tris pH 8.0, 150 m*M* NaCl, 5%(*v*/*v*) glycerol]. Cells were lysed by sonication and centrifuged at 38 000*g* in SS-34 centrifuge tubes (50 min, 277 K). The supernatant was filtered through a 0.45 µm pore-size membrane (Sartorius Stedim Biotech), loaded onto 5 ml Ni–NTA beads (Qiagen) pre-equilibrated with lysis buffer and incubated for 30 min at 277 K. The beads were then washed with high-salt buffer (50 m*M* Tris pH 8.0, 500 m*M* NaCl) and the protein was eluted with 50 m*M* Tris pH 8.0, 150 m*M* NaCl and 500 m*M* imidazole. The fractions collected were analyzed by SDS–PAGE and the protein sample was buffer-exchanged to lysis buffer using PD-10 columns (GE Healthcare).

The protein obtained using the pETM-11 vector (wild type/Leu122Met) was incubated overnight with His-tagged TEV protease (EMBL Hamburg) at 277 K (in a 50:1 ratio). The sample was then loaded onto Ni–NTA beads equilibrated with lysis buffer. The flowthrough was collected and the beads were washed with 20 ml lysis buffer. The fractions collected were analyzed by SDS–PAGE.

The protein sample was loaded onto a Superdex 75 (16/60) gel-filtration column (GE Healthcare) equilibrated with lysis buffer. The elution volume of the protein corresponded to a monomer according to the column calibration (Gel Filtration Standards, Bio-Rad). The purity of the sample was checked by SDS–PAGE. The protein solution was concentrated to 11 mg ml^{-1} using a 10 kDa molecular-weight cutoff centrifuge

Table 1
Data-collection and refinement statistics.

	Wild type	Leu122Met mutant	Wild type + ADP-ribose
Data collection			
Crystallization condition	Ammonium sulfate	Ammonium citrate	Ammonium citrate
X-ray source	ESRF ID29	ESRF ID14-2	EMBL X13
Space group	$P4_12_12$	$P2_13$	$P2_13$
Unit-cell parameters (Å)	$a = b = 161.1$, $c = 98.3$	$a = b = c = 218.9$	$a = b = c = 220.2$
Wavelength (Å)	0.979	0.993	0.812
Resolution range (Å)	50.0–3.1	50.0–2.2	20.0–3.9
Mosaicity (°)	0.17	0.14	0.16
Mean $I/\sigma(I)$	19.2 (3.5)	20.3 (5.6)	4.5 (2.6)
R_{fac} (linear) (%)	7.1 (46.0)	15.4 (60.7)	55.9 (90.6)
Redundancy	4.0	21.5	7.5
R_{meas} (%)	8.0 (51.6)	15.8 (62.1)	60.1 (97.5)
No. of observations	94994	3791311	240574
No. of unique reflections	23603	176075	32246
Completeness (%)	99.1 (98.1)	99.8 (98.9)	98.9 (98.5)
Refinement			
Resolution range (Å)	24.7–3.1	24.8–2.2	19.9–3.9
No. of reflections (working/free)	22625/1225	167093/8797	31727/465
No. of protein residues	*A*, 168; *B*, 166; *C*, 168	*A*–*H*, 168	*A*–*H*, 168
No. of waters	79	2343	—
No. of ADP-ribose molecules	—	—	3
No. of G3P molecules	3	—	—
No. of sulfate molecules	6	—	—
No. of chloride ions	4	—	2
No. of sodium ions	—	—	1
R_{work}/R_{free} (%)	20.01/24.20	14.70/18.13	19.88/23.91
Average B (Å^2)	60.0	39.9	31.0
Geometry bonds (Å)/angles (°)	0.014/1.6	0.020/1.6	0.029/2.6
Diffraction data precision indicator (Å)	0.31	0.19	0.53

concentrator (Vivaspin, Vivascience). The protein concentration was determined using a NanoDrop spectrophotometer (Thermo Scientific). The protein sample was stored at 277 K (for up to one week) or flash-frozen in the presence of 20%(*v*/*v*) glycerol and kept at 193 K (for up to six months).

2.3. Crystallization

Initial crystallization trials were carried out at 292 K using the sitting-drop vapour-diffusion method in 96-well Greiner plates at the EMBL Hamburg High-throughput Crystallization Facility (Mueller-Dieckmann, 2006). Crystals were obtained from 2.4 *M* ammonium sulfate, 0.1 *M* MES pH 6.0 (for protein expressed from pDEST14 and pETM-11) and 1.8 *M* triammonium citrate pH 7.0 (for protein expressed from pETM-11 vector). Further optimization of the crystallization conditions was performed manually in 24-well plates (Qiagen) using the hanging-drop vapour-diffusion method at 292 K. Crystals were obtained from a 2 µl drop of 6 mg ml^{-1} protein in 2.6–2.8 *M* ammonium sulfate and 0.1 *M* MES pH 5.0–6.0 or 11 mg ml^{-1} protein in 1.4–1.6 *M* diammonium citrate pH 6.0–7.0.

The crystals obtained using the ammonium sulfate condition initially diffracted to 4.5 Å resolution. In order to improve their diffraction quality, a dehydration method was implemented (Heras & Martin, 2005). Crystallization drops containing crystals were equilibrated for 12 h at 292 K in reservoirs containing crystallization-condition components plus increasing concentrations of one of the following: glycerol [4.25–17%(*v*/*v*)], MPD [5–20%(*v*/*v*)], ethylene glycol [5–20%(*v*/*v*)] or PEG 400 [2.5–10%(*v*/*v*)]. In each case after 12 h a few crystals were flash-cooled in liquid nitrogen. Crystals incubated in the 4.25% glycerol condition diffracted to 3.1 Å resolution.

2.4. Data collection and processing

Crystals grown with ammonium sulfate (from protein expressed from pDEST14 vector) were dehydrated using 4.25% glycerol and flash-cooled in liquid nitrogen. Single-wavelength X-ray diffraction data were collected from single crystals at 100 K on European Synchrotron Radiation Facility (ESRF) beamline ID29 using an ADSC Quantum 315R detector. The crystal-to-detector distance was maintained at 425.1 mm with an oscillation range of 0.6°. 100 images were collected to a maximum resolution of 3.0 Å.

Crystals grown in ammonium citrate (Leu122Met protein) were cryoprotected in 4.5 *M* sodium formate. A single-wavelength data set containing 360 images was collected from single crystals at 100 K on ESRF beamline ID14-2 using an ADSC Quantum 4 CCD detector. The crystal-to-detector distance was kept at 203 mm with an oscillation range of 0.5°. The crystal diffracted to 2.2 Å resolution.

A crystal grown in ammonium citrate (wild-type protein) was transferred to a 1 µl drop containing 4.5 *M* sodium formate and 2 m*M* ADP-ribose. After 1.5 h incubation the crystal was flash-cooled in liquid nitrogen. A single-wavelength data set containing 64 images was collected from a single crystal fragment at 100 K on the EMBL Hamburg X13 beamline using a MAR CCD 165 mm detector. The crystal-to-detector distance was kept at 300 mm with an oscillation range of 1°. The crystal diffracted to 3.9 Å resolution.

The recorded images were processed with *XDS* (Kabsch, 1988) and the reflection intensities were processed with *COMBAT* and scaled with *SCALA* (Evans, 1993) from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994). Data-collection statistics are shown in Table 1.

2.5. Structure determination

The structure of the FCoV X domain in space group $P4_12_12$ was determined by the molecular-replacement method using the program *MOLREP* (Vagin & Teplyakov, 1997). The coordinates of the HCoV-NL63 X domain (Piotrowski *et al.*, to be published) served as a search model. The solution showed three molecules in the asymmetric unit. Refinement was carried out using the program *REFMAC*5 (Murshudov *et al.*,

1997). The structure was visualized and rebuilt into the electron density using the program *Coot* (Emsley & Cowtan, 2004). The stereochemistry of the model was evaluated using the program *MolProbity* (Davis *et al.*, 2007).

Molecule *A* from the tetragonal space-group solution was used as a search model for molecular replacement in space group $P2_13$. The solution found by the program *Phaser* (McCoy *et al.*, 2007) consisted of six molecules. Appropriately weighted $2F_o - F_c$ and $F_o - F_c$ maps were calculated at this stage. These maps showed additional electron density indicating the presence of two additional protein molecules in the asymmetric unit. The molecular-replacement solution was used as a preliminary model for *ARP*/*wARP* (Perrakis *et al.*, 1999), which built most of the eight molecules into the electron density. The final refinement was carried out with *REFMAC*5. Where necessary, the structure was manually modified using the program *Coot*. The stereochemistry of the model was evaluated with the program *MolProbity*.

The structure of the ADP-ribose-bound FCoV X domain in space group $P2_13$ was solved using the molecular-replacement protocol of the automated crystal structure-determination platform *Auto-Rickshaw* (Panjikar *et al.*, 2005). The structure solution from the cubic space group was used as a search model for molecular replacement. Appropriately weighted $2F_o - F_c$ and $F_o - F_c$ maps were calculated for the coordinate file obtained from *Auto-Rickshaw*. These maps showed additional electron density in the binding pockets of three molecules (*B*, *C* and *D*), indicating the presence of ADP-ribose. The refinement was carried out using the program *REFMAC*5. The structure was visualized and the fit to the electron density was verified using the program *Coot*. The stereochemistry of the model was evaluated with the program *MolProbity*.

Atomic coordinates have been deposited in the Protein Data Bank (the PDB code for the tetragonal space group is 3ew5, that for the cubic space group is 3eti and that for the ADP-ribose bound structure is 3jzt).

Interfaces between molecules in both crystal forms were analyzed with the *PISA* server (Krissinel & Henrick, 2007). Interface areas larger than 300 Å^2 were treated as being significant in describing the crystal packing, although the domain was clearly monomeric in dilute solution. Interactions between molecules were calculated with the *CCP*4 program *CONTACT*. The maximum contact distance considered was 3.6 Å.

2.6. Binding assay

The ADP-ribose-binding assay is based upon the pull-down experiment described by Karras *et al.* (2005). The Ni–NTA slurry (~100 µl) was equilibrated with lysis buffer [50 m*M* Tris pH 8.0, 150 m*M* NaCl, 5%(*v*/*v*) glycerol]. 1 ml 27 µ*M* protein sample in lysis buffer (from protein expressed using the pETM-11 vector) was loaded onto the column and incubated for 10 min. In order to check the effectiveness of the immobilization of the protein on the resin, the absorbance of the collected supernatant was measured at 280 nm. 1 ml 27 µ*M* ADP-ribose solution was loaded onto the column and incubated for 30 min. The absorbance of the collected supernatant was measured at 259 nm. The percentage of ADP-ribose bound to the protein was calculated as the ratio between the supernatant absorbance and the absorbance of 27 µ*M* ADP-ribose solution at 259 nm. The ADP-ribose concentration was calculated using an extinction coefficient of 15 400 M^{-1} cm^{-1} at 260 nm.

3. Results and discussion

3.1. Structure determination

The wild-type FCoV X domain (residues 1254–1421 of pp1a, here renumbered as 34–201 in both the pDEST14 and pETM-11 vectors) cloned into vector pDEST14 was expressed in *E. coli* cells at high yield (~60 mg per litre of cell culture). The protein crystallized overnight from a solution containing ammonium sulfate and MES. The crystals belonged to space group $P4_12_12$, with unit-cell parameters $a = b = 161.1$, $c = 98.3$ Å. In this crystal form there were three molecules (chains *A*, *B* and *C*) in the asymmetric unit, corresponding to a solvent content of 78% (Matthews, 1968; see Supplementary Fig. S6[1]). The structure was refined at 3.1 Å resolution to a final *R* value of 20.0% (R_{free} = 24.2%). All residues, except for the C-terminal His_6 tag, two N-terminal residues in chains *A* and *C* and four in chain *B*, could be placed into the electron density. The Ramachandran plot showed 91.7% of the residues in preferred regions and 8.3% in allowed regions. The refined structure contained three glycerol-3-phosphate molecules, six sulfate ions, four chloride ions and 79 solvent molecules.

The mutated Leu122Met X domain crystallized in the presence of ammonium citrate. These crystals belonged to space group $P2_13$, with unit-cell parameters $a = b = c = 218.9$ Å. There were eight molecules (chains *A*–*H*) in the asymmetric unit, corresponding to 74% solvent content. The crystals contained large solvent channels with a diameter of approximately 80 Å (see Supplementary Fig. S1[1]). The structure was refined at 2.2 Å resolution to a final *R* value of 14.7% (R_{free} = 18.1%). In all chains only the first three N-terminal residues showed no electron density. The Ramachandran plot showed 97.8% of the residues to be in preferred regions and 2.2% to be in allowed regions. The refined structure contained 2343 solvent molecules.

The structure with ADP-ribose bound was determined in the cubic crystal form at 3.9 Å resolution. The crystal belonged to space group $P2_13$, with unit-cell parameters $a = b = c = 220.2$ Å. The structure was refined to a final *R* value of 19.9% (R_{free} = 23.9%). Data-collection and structure-refinement statistics are shown in Table 1.

3.2. Overall structure

The FCoV X domain is a single domain with a mixed α/β structure (Fig. 1*a*). The core of the structure is a single mixed

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: DZ5170). Services for accessing this material are described at the back of the journal.

β-sheet. The order of the strands in the sheet is β1, β2, β7, β6, β3, β5, with the last strand in two stretches, β4a and β4b. The central five strands are parallel. The β-sheet is sandwiched between six helices, with α1, α2 and α3 packing onto one face, and α4, α5 and α6 onto the other. The SARS-CoV and HCoV-229E X domains have the same arrangement of β-strands, while the IBV X domain lacks the N-terminal strand. The unique feature of the FCoV X domain β-sheet is the last broken strand. The topology is similar to other coronaviral X domains (Fig. 1*b*).

Chain *A* from the cubic crystal form was compared with all structures in the PDB using the *DALI* server (http://ekhidna.biocenter.helsinki.fi/dali_server/). The structure of the FCoV X domain is very similar to those of X domains from other members of the *Coronaviridae* (*Z* scores from 20.1 to 27.6), as well as to the mammalian nonhistone domain of the histone variant macroH2.A (*Z* scores between 19.7 and 20.0), confirming that the FCoV X domain has a macro-domain-like fold.

The r.m.s.d. between the C^{α} atoms of the molecules in the tetragonal space group was less than 0.4 Å (for 136 C^{α} atoms). In the cubic crystal form the r.m.s.d. between molecules was less than 0.15 Å (for 142 C^{α} atoms). The r.m.s.d. between C^{α} atoms from the tetragonal and cubic space groups was between 0.47 and 0.72 Å. The smaller r.m.s.d. differences in the cubic crystal form are most likely to be a result of the higher resolution structure determination (2.2 Å).

3.3. Crystal packing

The FCoV X domain has been crystallized in three different space groups: tetragonal, cubic and orthorhombic (diffraction to a maximum of 3 Å, data not shown). Three molecules were found in the asymmetric unit cell of the tetragonal crystal form and eight in that of the cubic crystal form. Orthorhombic crystals were obtained in buffer containing sodium formate and sodium acetate. Interestingly, soaking cubic crystals with heavy-atom solutions led to a change of the space group to the orthorhombic form. The implied similarity between the packing in the two crystal forms would suggest the presence of 24 molecules in the asymmetric unit of the orthorhombic crystal form. The tendency of the FCoV X domain to oligomerize is striking, taking into consideration the fact that no buffer molecules were interpretable in the electron density of either the tetragonal or cubic space groups. Interestingly, *PISA*-mediated analysis of X-domain interfaces in both the tetragonal and cubic space groups revealed no specific interactions that could result in the formation of stable quaternary structures, which is in agreement with the observation that the X domain is monomeric in dilute solution. Analytical size-exclusion chromatography demonstrated that under the experimental conditions used the protein eluted as a monomer (results not shown). However, solutions of the protein do yield a heavy precipitate overnight. The observed tendency to form multimers might be biologically relevant since the X domain is one of many domains of nsp3, which is likely to be part of a large replication/transcription complex. Indeed, interactions between several nonstructural proteins have been identified and are of importance in the structure and function of the viral enzyme complexes that have been analyzed (Imbert *et al.*, 2008; von Brunn *et al.*, 2007). The crystal structure of the IBV X domain contains a crystallographic dimer with a relatively large interface area of 2600 $Å^2$ (Xu *et al.*, 2009). In contrast, the interface between FCoV X-domain monomers is small but well defined, with some interactions occurring in both crystals forms. The interactions result in rigid lattices with large solvent channels, which might make the crystals capable of acting of a host lattice for other biological molecules. The

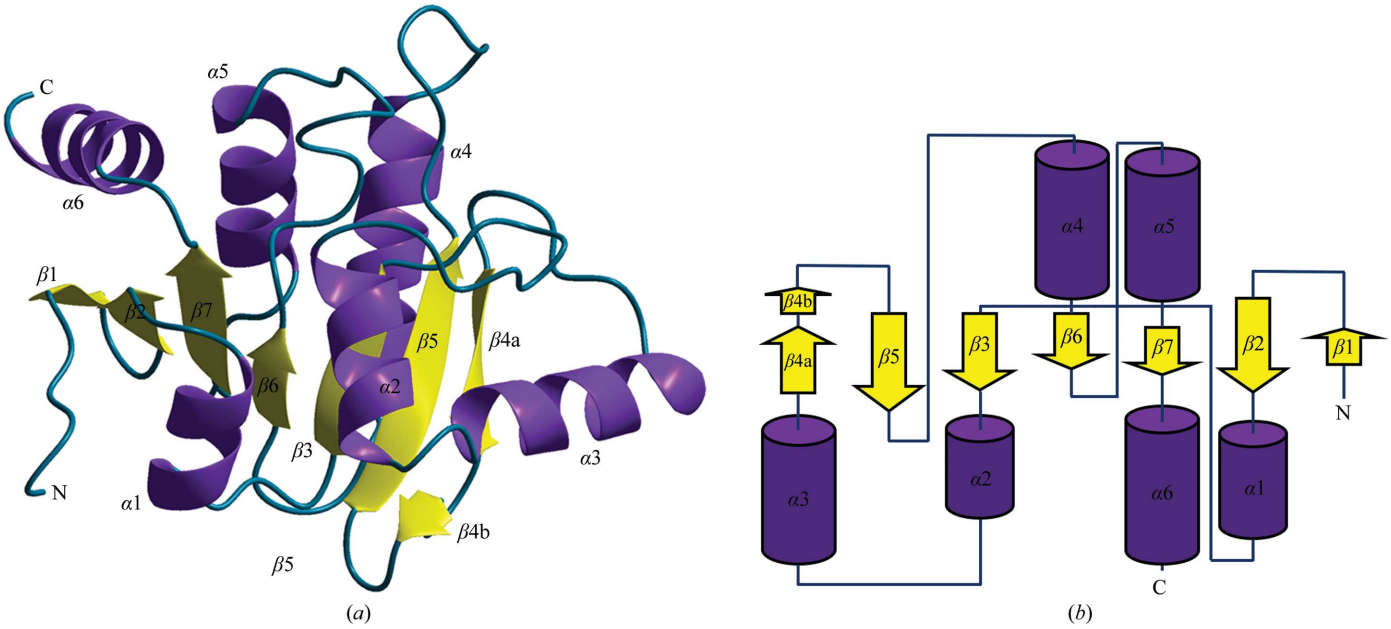


Figure 1
(*a*) Ribbon representation of the FCoV X domain. Loops are shown in blue, α-helices in purple and β-strands in yellow. (*b*) Topology diagram of the FCoV X domain coloured as in (*a*).

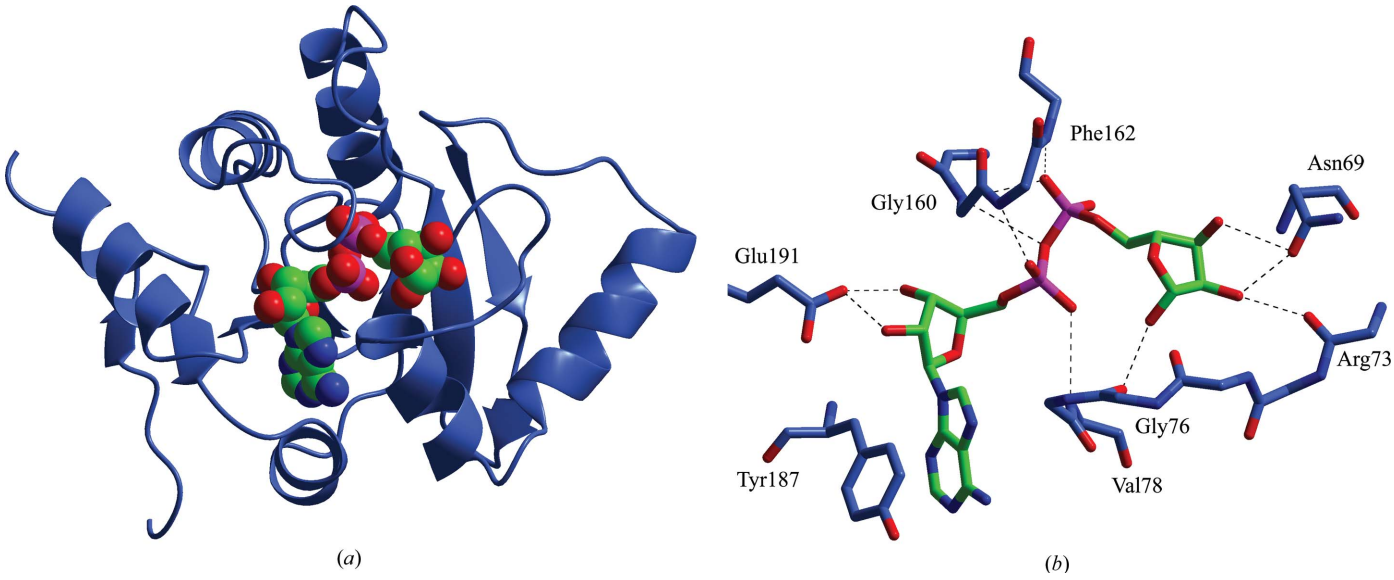


Figure 2
(*a*) Ribbon representation of the FCoV X domain with ADP-ribose bound. ADP-ribose is shown as a space-filling model coloured by atom type. (*b*) Interactions between residues of the FCoV X domain and ADP-ribose. Protein residues (in light blue) and ADP-ribose are shown in stick representation. For protein regions 73–78 and 159–162 only the main chain is shown. Hydrogen bonds are shown as dashed lines. N atoms are shown in blue, C atoms in green, O atoms in red and P atoms in magenta.

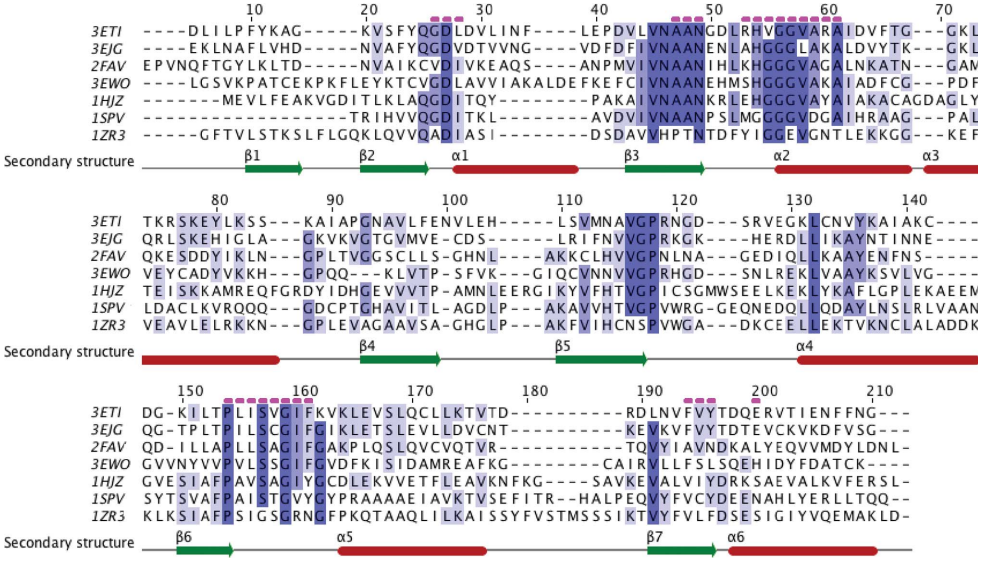


Figure 3
Structure-based sequence alignment of macro domains. The alignment is based on the PDB structures of the FCoV X domain (PDB code 3eti; UniProtKB reference Q98VG9), the HCoV-229E X domain (3ejg; UniProtKB reference P0C6X1), the SARS-CoV X domain (2fav; UniProtKB reference P0C6U8), the IBV X domain (3ewo; UniProtKB reference P0C6V5), *A. fulgidus* AF1521 protein (1hjz; UniProtKB reference O28751), *E. coli* ER58 protein (1spv; UniProtKB reference P0A8D6) and human macroH2.2A domain (1zr3; UniProtKB reference O75367). The secondary structure of the FCoV X domain is shown. Regions of high sequence identity are shaded blue. Residues forming the ADP-ribose-binding site are indicated by dashed magenta lines. The alignment was generated using the EBI *SSM* tool (http://www.ebi.ac.uk/msd-srv/ssm/) and modified with *Jalview*.

packing is described in more detail in the supplementary material.

3.4. ADP-ribose-binding site

By soaking a native FCoV X-domain crystal in 2 m*M* ADP-ribose for 1.5 h, we were able to observe bound ADP-ribose in the cubic crystal form. ADP-ribose molecules could be clearly identified from the electron density in molecules *B*, *C* and *D*. In all three molecules the ADP-ribose molecule is located in a binding pocket formed mainly by the N-terminal part of α1, the C-terminal part of β3, the long loop $L_{\beta3-\alpha2}$, the N-terminal residues of α2, the loop $L_{\beta7-\alpha5}$ and the region between the C-terminus of β8 and the N-terminus of α7 (Fig. 2). The ADP-ribose-binding cavity is open and solvent-accessible, with a positively charged floor. Upon binding to the FCoV X domain the ADP-ribose molecule adopts a slightly bent conformation. The adenine moiety lies in the hydrophobic cavity formed by Leu52, Val78, Ala81, Pro155, Phe185 and Tyr187. The side chain of Tyr187 stacks against the adenine ring. This type of interaction has also been observed in the structures of HCoV-229E (Tyr152) and in the AF1521 protein (Tyr176), but the residue is not conserved in macro domains. It is Asn in SARS-CoV, Leu in IBV and Phe in human macroH.2A (Fig. 3). Interestingly, the highly conserved Asp51 that has been postulated to be critical for binding specificity does not interact with the adenine (at a distance of 4.7 Å) in this structure. The adenosine ribose forms strong hydrogen bonds

to Glu191. The distal ribose interacts with the main-chain atoms of Arg73, Val75 and Gly76 and also forms a hydrogen bond to the side chain of Asn69. Two phosphate groups are hydrogen bonded to the main-chain N atoms of the region 158-SVGIF-162. During the refinement process for the tetragonal crystal form, additional electron density ($wF_o - DF_c$ synthesis) was found in the binding cleft of molecules *A*, *B* and *C*. Glycerol-3-phosphate (G3P), which is a very abundant metabolite in bacterial cells, could be fitted to this density (Fig. 4). A comparison with the ADP-ribose-bound molecule *B* revealed that the phosphate of G3P is in the same position as the β-phosphate of ADP-ribose. Similarly to the ADP-ribose-bound structure, the region 158-SVGIF-162 forms hydrogen bonds between its backbone amide groups and the O atoms of the G3P phosphate. Interestingly, despite high concentrations of sulfate ion in the crystallization buffer, the binding site of the FCoV X domain is apparently selective for phosphate.

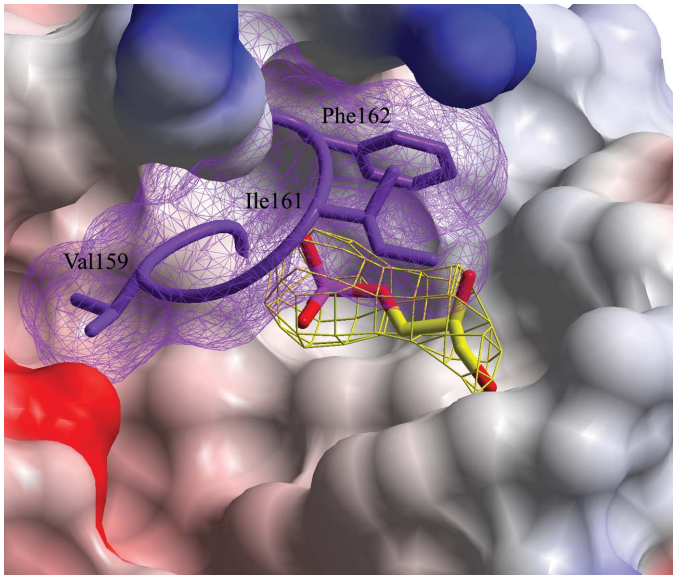
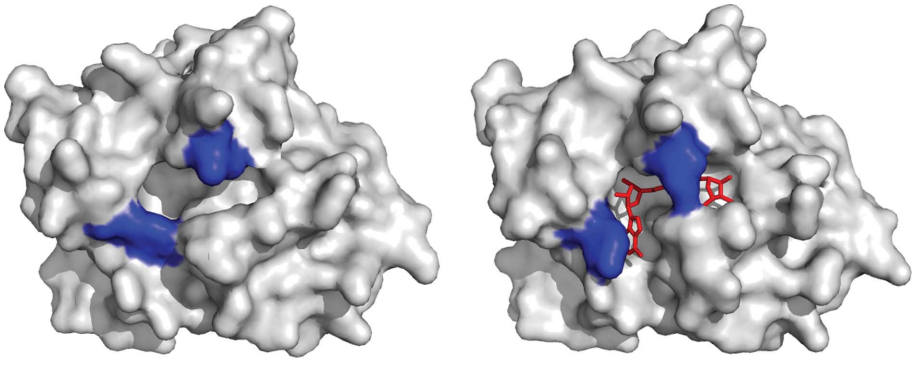


Figure 4
Glycerol-3-phosphate in the binding site. The O atoms of G3P phosphate are hydrogen bonded to the main-chain N atoms of the region 158-SVGIF-162. The surface of the protein is coloured according to electrostatic potential (negative charge, red; neutral, white; positive charge, blue); residues 158–162 are shown in purple, the O atoms of G3P are shown in red, the phosphate in pink and the C atoms in yellow.

Comparison of ligand-free molecules with the ADP-ribose-bound molecules reveals two major conformational changes that occur upon ligand binding. Egloff *et al.* (2006) showed that the X domain undergoes conformational changes upon binding to ADP-ribose, leading to the formation of a 'bridge' over the binding cleft. This 'bridge' is formed by close contacts between Ile132 and Gly48 (SARS-CoV numbering). Gly48 is part of a largely conserved 'triple-glycine' sequence 47-GGG-49 (FCoV numbering). In the X domains of HCoV-229E and IBV strain M41, in which all three positions are occupied by glycine residues, the 'bridge' between Ile126–Gly44 (HCoV-229E numbering) and Ile133–Gly49 (IBV strain M41 numbering) is formed upon ADP-ribose binding (Xu *et al.*, 2009). However, in the case of IBV strain Beaudette the second glycine is substituted by serine (46-GSG-48). This leads to large structural changes in the region 128-SCGIF-132 (IBV strain Beaudette numbering), which might explain why this X domain does not bind ADP-ribose (Piotrowski *et al.*, 2009). In the FCoV X domain the first glycine is replaced by valine (75-VGG-77). In the ligand-free cubic crystal form the shortest interatomic distance between residues Ile161 and Gly76 is greater than 6.0 Å (with one exception). In the tetragonal crystal form with G3P bound the shortest interatomic distance between residues Ile161 and Gly76 is 4.5 Å (in molecules *A* and *C*) or 5.7 Å (molecule *B*). However, in ADP-ribose-bound molecules *B*, *C* and *D* the shortest distance between those two residues is smaller than 4 Å, which results in formation of the 'bridge' (Fig. 5). Interestingly, the G3P-bound form seems to be an intermediate between closed and open conformations in which residues Ile161 and Gly76 are becoming closer to each other than in the ligand-free form but in which the 'bridge' is still not formed.

In the ligand-free FCoV X domain the size of the binding pocket is too small to accept ADP-ribose. This largely arises from the conformation of Tyr187, which lies across the binding pocket. In ADP-ribose-bound molecules the side chain of Tyr187 rotates by 90° and orients along the ADP-ribose binding site, significantly expanding the binding cavity (Fig. 5). Cleft analysis using the *PDBsum* server (http://www.ebi.ac.uk) showed that conformational change of Tyr187 leads to the enlargement of the binding cleft by almost 600 Å^3, which creates sufficient space for the binding of ADP-ribose.



(*a*) (*b*)

Figure 5
Surface of the FCoV X domain in the ligand-free conformation (*a*) and with ADP-ribose bound (*b*). Residues Ile161 and Tyr187, which undergo significant conformational changes during ADP-ribose binding, are labelled in blue. ADP-ribose is shown as red sticks.

In order to further examine the binding properties of the FCoV X domain, an ADP-ribose-binding assay was performed *in vitro* (see §2). The retention of the ligand on immobilized

X domain was 7.1% (compared with 0.8% retention in the case of the control), showing that the FCoV X domain binds ADP-ribose very weakly under the conditions used in the experiment. The determined upper limit for the dissociation constant was around 400 μ*M*. This value is one order of magnitude higher than those obtained for orthologues from SARS-CoV (24 μ*M*; Egloff *et al.*, 2006) and HCoV-229E (28.9 μ*M*; Piotrowski *et al.*, 2009). The binding affinity of the FCoV X domain is also markedly lower than that of AF1521 (126 n*M*; Karras *et al.*, 2005) and human macroH2A1.1 (2.7 μ*M* for *O*-acetyl-ADP-ribose; Kustatscher *et al.*, 2005).

Macro domains have two related but distinct properties: ligand binding and enzymatic activity. Two highly divergent macro-domain folds have recently been identified in the nsp3 'SARS-unique domain' (SUD; Tan *et al.*, 2009) that was shown to bind single-stranded poly(A) (Chatterjee *et al.*, 2009) and oligo(G) strings (Tan *et al.*, 2007). Several macro domains, including coronaviral X domains, have been shown to bind ADP-ribose and related ligands. It has also been suggested that the function of viral X domains may be the ability to bind poly(ADP)-ribose, by which they interact with host-cell pathways (Egloff *et al.*, 2006). Catalytic adenosine diphosphate-ribose-1″-phosphatase activity has been shown for a number of macro domains, including the X domain of another member of coronavirus subgroup 1a, TGEV. The ADRP reaction turnover was found to be very low *in vitro* and this led to uncertainty as to the physiological relevance of this enzymatic activity (Egloff *et al.*, 2006). On the other hand, this activity has been implicated in the rate control of a yet-to-be-identified RNA-processing pathway in infected cells (Snijder *et al.*, 2003). In addition, it has been reported that MHV liver pathology is affected in an engineered virus mutant carrying an X domain with a mutation in the putative active site (Eriksson *et al.*, 2008). Despite the available information, the function of the coronaviral X domain and its importance in the viral life cycle have not been fully elucidated and further functional studies are therefore required in order to understand its role in detail.

4. Conclusions

The FCoV X-domain structure has a macro-domain-like fold with an ADP-ribose-binding pocket. It reveals high similarity to the structures of other members of the macro-domain family, especially to X domains from other coronaviruses. We were also able to show that the binding affinity of the FCoV X domain for ADP-ribose is noticeably lower than in the case of some other members of the family.

We thank Dr Stuart Siddell (University of Bristol, England) for kindly providing feline coronavirus and Linda Boomaars-van der Zanden for excellent technical assistance. This work was supported by the European project 'VIZIER' ('Comparative Structural Genomics of Viral Enzymes Involved in Replication') funded by the Sixth Framework Programme of the European Commission under reference LSHG-CT-2004-511960.

References

Brunn, A. von, Teepe, C., Simpson, J. C., Pepperkok, R., Friedel, C. C., Zimmer, R., Roberts, R., Baric, R. & Haas, J. (2007). *PLoS ONE*, **2**, e459.
Chatterjee, A., Johnson, M. A., Serrano, P., Pedrini, B., Joseph, J. S., Neuman, B. W., Saikatendu, K., Buchmeier, M. J., Kuhn, P. & Wuthrich, K. (2009). *J. Virol.* **83**, 1823–1836.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B. III, Snoeyink, J., Richardson, J. S. & Richardson, D. C. (2007). *Nucleic Acids Res* **35,** W375–W383.
Egloff, M. P., Malet, H., Putics, A., Heinonen, M., Dutartre, H., Frangeul, A., Gruez, A., Campanacci, V., Cambillau, C., Ziebuhr, J., Ahola, T. & Canard, B. (2006). *J. Virol.* **80**, 8493–8502.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Eriksson, K. K., Cervantes-Barragan, L., Ludewig, B. & Thiel, V. (2008). *J. Virol.* **82**, 12325–12334.
Evans, P. R. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 114–122. Warrington: Daresbury Laboratory.
Gorbalenya, A. E., Enjuanes, L., Ziebuhr, J. & Snijder, E. J. (2006). *Virus Res.* **117**, 17–37.
Gorbalenya, A. E., Koonin, E. V. & Lai, M. M. (1991). *FEBS Lett.* **288**, 201–205.
Gorbalenya, A. E., Snijder, E. J. & Spaan, W. J. (2004). *J. Virol.* **78**, 7863–7866.
Heras, B. & Martin, J. L. (2005). *Acta Cryst.* D**61**, 1173–1180.
Imbert, I., Snijder, E. J., Dimitrova, M., Guillemot, J. C., Lecine, P. & Canard, B. (2008). *Virus Res.* **133**, 136–148.
Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.
Kanjanahaluethai, A., Chen, Z., Jukneliene, D. & Baker, S. C. (2007). *Virology*, **361**, 391–401.
Karras, G. I., Kustatscher, G., Buhecha, H. R., Allen, M. D., Pugieux, C., Sait, F., Bycroft, M. & Ladurner, A. G. (2005). *EMBO J.* **24**, 1911–1920.
Knoops, K., Kikkert, M., Worm, S. H., Zevenhoven-Dobbe, J. C., van der Meer, Y., Koster, A. J., Mommaas, A. M. & Snijder, E. J. (2008). *PLoS Biol.* **6**, e226.
Koonin, E. V., Gorbalenya, A. E., Purdy, M. A., Rozanov, M. N., Reyes, G. R. & Bradley, D. W. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 8259–8263.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Kumaran, D., Eswaramoorthy, S., Studier, F. W. & Swaminathan, S. (2005). *Protein Sci.* **14**, 719–726.
Kustatscher, G., Hothorn, M., Pugieux, C., Scheffzek, K. & Ladurner, A. G. (2005). *Nature Struct. Mol. Biol.* **12**, 624–625.
Lai, M. M. C. & Holmes, K. V. (2001). *Fields Virology*, 4th ed., edited by D. M. Knipe & P. M. Howley, pp. 1163–1185. Philadelphia: Lippincott Williams & Wilkins.
Malet, H., Coutard, B., Jamal, S., Dutartre, H., Papageorgiou, N., Neuvonen, M., Ahola, T., Forrester, N., Gould, E. A., Lafitte, D., Ferron, F., Lescar, J., de Lamballerie, X. & Canard, B. (2009). *J. Virol.* **83**, 6534–6545.
Martzen, M. R., McCraith, S. M., Spinelli, S. L., Torres, F. M., Fields, S., Grayhack, E. J. & Phizicky, E. M. (1999). *Science*, **286**, 1153–1155.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Mihindukulasuriya, K. A., Wu, G., St Leger, J., Nordhausen, R. W. & Wang, D. (2008). *J. Virol.* **82**, 5084–5088.
Mueller-Dieckmann, J. (2006). *Acta Cryst.* D**62**, 1446–1452.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Neuman, B. W., Joseph, J. S., Saikatendu, K. S., Serrano, P., Chatterjee, A., Johnson, M. A., Liao, L., Klaus, J. P., Yates, J. R. III, Wüthrich, K., Stevens, R. C., Buchmeier, M. J. & Kuhn, P. (2008). *J. Virol.* **82**, 5279–5294.
Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S. & Tucker, P. A. (2005). *Acta Cryst.* D**61**, 449–457.
Pedersen, N. C. (1995). *Feline Pract.* **23**, 7–20.
Pehrson, J. R. & Fuji, R. N. (1998). *Nucleic Acids Res.* **26**, 2837–2842.
Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Piotrowski, Y., Hansen, G., Boomaars-van der Zanden, A. L., Snijder, E. J., Gorbalenya, A. E. & Hilgenfeld, R. (2009). *Protein Sci.* **18**, 6–16.
Poland, A. M., Vennema, H., Foley, J. E. & Pedersen, N. C. (1996). *J. Clin. Microbiol.* **34**, 3180–3184.
Prentice, E., McAuliffe, J., Lu, X., Subbarao, K. & Denison, M. R. (2004). *J. Virol.* **78**, 9977–9986.
Putics, A., Filipowicz, W., Hall, J., Gorbalenya, A. E. & Ziebuhr, J. (2005). *J. Virol.* **79**, 12721–12731.
Putics, A., Gorbalenya, A. E. & Ziebuhr, J. (2006). *J. Gen. Virol.* **87**, 651–656.
Saikatendu, K. S., Joseph, J. S., Subramanian, V., Clayton, T., Griffith, M., Moy, K., Velasquez, J., Neuman, B. W., Buchmeier, M. J., Stevens, R. C. & Kuhn, P. (2005). *Structure*, **13**, 1665–1675.
Shull, N. P., Spinelli, S. L. & Phizicky, E. M. (2005). *Nucleic Acids Res.* **33**, 650–660.
Snijder, E. J., Bredenbeek, P. J., Dobbe, J. C., Thiel, V., Ziebuhr, J., Poon, L. L., Guan, Y., Rozanov, M., Spaan, W. J. & Gorbalenya, A. E. (2003). *J. Mol. Biol.* **331**, 991–1004.
Snijder, E. J., van der Meer, Y., Zevenhoven-Dobbe, J., Onderwater, J. J., van der Meulen, J., Koerten, H. K. & Mommaas, A. M. (2006). *J. Virol.* **80**, 5927–5940.
Stertz, S., Reichelt, M., Spiegel, M., Kuri, T., Martinez-Sobrido, L., Garcia-Sastre, A., Weber, F. & Kochs, G. (2007). *Virology*, **361**, 304–315.
Tan, J., Kusov, Y., Mutschall, D., Tech, S., Nagarajan, K., Hilgenfeld, R. & Schmidt, C. L. (2007). *Biochem. Biophys. Res. Commun.* **364**, 877–882.
Tan, J., Vonrhein, C., Smart, O. S., Bricogne, G., Bollati, M., Kusov, Y., Hansen, G., Mesters, J. R., Schmidt, C. L. & Hilgenfeld, R. (2009). *PLoS Pathog.* **5**, e1000428.
Thiel, V., Ivanov, K. A., Putics, A., Hertzig, T., Schelle, B., Bayer, S., Weissbrich, B., Snijder, E. J., Rabenau, H., Doerr, H. W., Gorbalenya, A. E. & Ziebuhr, J. (2003). *J. Gen. Virol.* **84**, 2305–2315.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Vennema, H., Poland, A., Foley, J. & Pedersen, N. C. (1998). *Virology*, **243**, 150–157.
Weiss, S. R. & Navas-Martin, S. (2005). *Microbiol. Mol. Biol. Rev.* **69**, 635–664.
Woo, P. C., Lau, S. K., Lam, C. S., Lai, K. K., Huang, Y., Lee, P., Luk, G. S., Dyrting, K. C., Chan, K. H. & Yuen, K. Y. (2009). *J. Virol.* **83**, 908–917.
Xu, Y., Cong, L., Chen, C., Wei, L., Zhao, Q., Xu, X., Ma, Y., Bartlam, M. & Rao, Z. (2009). *J. Virol.* **83**, 1083–1092.
Ziebuhr, J., Snijder, E. J. & Gorbalenya, A. E. (2000). *J. Gen. Virol.* **81**, 853–879.