Crystal structure of the papain-like protease of MERS coronavirus reveals unusual, potentially druggable active-site features

Jian Lei a,c, Jeroen R. Mesters a,c, Christian Drosten b,c, Stefan Anemüller a,c, Qingjun Ma a,c, Rolf Hilgenfeld a,c,*

a Institute of Biochemistry, Center for Structural and Cell Biology in Medicine, University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany
b Institute of Virology, University of Bonn Medical School, 53127 Bonn, Germany
c German Center for Infection Research (DZIF), Germany

A R T I C L E   I N F O

Article history:
Received 8 May 2014
Revised 14 June 2014
Accepted 18 June 2014
Available online 30 June 2014

Keywords:
MERS-CoV
Papain-like protease
Oxyanion hole
Zinc finger
Deubiquitinating
Antiviral drug design

A B S T R A C T

The Middle-East Respiratory Syndrome coronavirus (MERS-CoV) causes severe acute pneumonia and renal failure. The MERS-CoV papain-like protease (PLpro) is a potential target for the development of antiviral drugs. To facilitate these efforts, we determined the three-dimensional structure of the enzyme by X-ray crystallography. The molecule consists of a ubiquitin-like domain and a catalytic core domain. The catalytic domain displays an extended right-hand fold with a zinc ribbon and embraces a solvent-exposed substrate-binding region. The overall structure of the MERS-CoV PLpro is similar to that of the corresponding SARS-CoV enzyme, but the architecture of the oxyanion hole and of the S3 as well as the S5 specificity sites differ from the latter. These differences are the likely reason for reduced in vitro peptide hydrolysis and deubiquitinating activities of the MERS-CoV PLpro, compared to the homologous enzyme from the SARS coronavirus. Introduction of a side-chain capable of oxyanion stabilization through the Leu106Trp mutation greatly enhances the catalytic activity of the MERS-CoV PLpro. The unique features observed in the crystal structure of the MERS-CoV PLpro should allow the design of antivirals that would not interfere with host ubiquitin-specific proteases.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/3.0/).

1. Introduction

Ten years after the outbreak of severe acute respiratory syndrome (SARS) of 2002/2003 (Hilgenfeld and Peiris, 2013), another highly pathogenic coronavirus, Middle-East Respiratory Syndrome coronavirus (MERS-CoV), has been recognized to infect humans (Zaki et al., 2012; de Groot et al., 2013). Accumulating evidence suggests camels to act as a zoonotic source of the virus (Reusken et al., 2013; Haagmans et al., 2014; Meyer et al., 2013). Limited human-to-human transmission of the virus has been described (Assiri et al., 2013). As of June 11, 2014, 683 cases of MERS have been reported, with 204 deaths (http://www.who.int). The clinical symptoms of MERS include severe pneumonia and sometimes acute renal failure (Eckerle et al., 2013). However, the majority of MERS patients had comorbidities, such as diabetes, lung disease, or chronic renal disease (Perlman, 2013).

SARS-CoV and MERS-CoV belong to the genus Betacoronavirus but pertain to highly distinct phylogenetic clades termed b and c, respectively (de Groot et al., 2013). In case of SARS-CoV, the best-characterized potential antiviral drug targets are the two viral proteases, the main protease (Mpro, also called 3C-like protease, 3CLpro) (Hilgenfeld and Peiris, 2013; Anand et al., 2003; Yang et al., 2003, 2005; Xu et al., 2005; Lu et al., 2006; Verschueren et al., 2008; Zhu et al., 2011; Kilianski et al., 2013) and the papain-like protease (PLpro) (Hilgenfeld and Peiris, 2013; Kilianski et al., 2013; Barretto et al., 2005; Ratia et al., 2006, 2008; Baez-Santos et al., 2014). The latter enzyme exists in all coronaviruses (Woo et al., 2010) and has been shown to be responsible for releasing non-structural proteins (Nsp) 1, 2, and 3 from the N-terminal part of polyproteins 1a and 1ab. The three cleavage sites contain the sequence motif LXGGXX. In addition, the SARS-CoV PLpro has been shown to have deubiquitinating and interferon antagonism activities, thereby interfering with the host innate immune response (Barretto et al., 2005; Lindner et al., 2005; Devaraj et al., 2007; Frieman et al., 2009). Specifically, it can prevent the activation of IRF3 (interferon-regulatory factor 3) and antagonize the NF-kB (nuclear factor κ-light-chain-enhancer of activated B cells) pathway, but the
detailed mechanisms involved are still unclear. Very recently, the MERS-CoV PL\textsuperscript{pro} has been reported to also have proteolytic, deubiquitinating, and delISG15ylating activities in HEK293T cells (Yang et al., 2013; Mielech et al., 2014) (ISG15 = interferon-stimulated gene 15); it therefore also acts as an interferon antagonist through blocking the IRF3 pathway. Interestingly, these reports differ in their finding that the interferon-antagonizing activity of the MERS-CoV PL\textsuperscript{pro} is either independent of (Yang et al., 2013) or dependent on (Mielech et al., 2014) its proteolytic activity.

In spite of the accumulating knowledge on the essential roles of the coronavirus PL\textsuperscript{pro} in virus replication and evasion of the host-cell innate immune response (Devaraj et al., 2007; Frieman et al., 2009; Yang et al., 2013; Mielech et al., 2014; Lindner et al., 2007; Clementz et al., 2010), the three-dimensional structures of only two of these enzymes have been reported so far, i.e., that of the PL\textsuperscript{pro} from SARS-CoV (Ratia et al., 2006) and that of the PL\textsuperscript{pro} from Transmissible Gastroenteritis Virus (TGEV) (Wojdyła et al., 2010). Here we present the crystal structure of the MERS-CoV PL\textsuperscript{pro} at 2.50 Å resolution, in order to unravel the structural basis of the activities of the enzyme and facilitate structure-based drug design efforts. In addition, we report the in vitro hydrolytic activities of the enzyme towards two synthetic peptide substrates and a fluorogenic ubiquitin derivative.

2. Materials and methods

2.1. Recombinant production of MERS-CoV papain-like protease (PL\textsuperscript{pro})

The PL\textsuperscript{pro} of MERS-CoV (strain 2c EMC/2012; GenBank: AVF09327.1) comprises 320 amino-acid residues, corresponding to Gln1482 – Asp1801 of pp1a, and is part of non-structural protein AFV09327.1) comprises 320 amino-acid residues, corresponding to Gln1482 – Asp1801 of pp1a, and is part of non-structural protein

2.2. Crystallization and diffraction-data collection

Purified PL\textsuperscript{pro} was concentrated to ~11 mg/ml in buffer C. Crystallization was performed at 18 °C by using a Phoenix crystallization robot (Art Robbins) employing the sitting-drop vapor-diffusion method, with mixing 0.25 μl of protein and 0.25 μl of reservoir to equilibrate against 75 μl reservoir solution. The following commercially available screens were used: SaltRxTM, PEG/IonTM 1 & 2 Screen, IndexTM, and PEG RxTM 1 & 2 (Hampton Research). Crystals were observed under condition 19 of IndexTM. Optimized crystals were subsequently obtained within one day using 0.056 M NaH2PO4, 1.344 M K2HPO4, pH 8.0, and 15% glycerol as reservoir, with mixing 2 μl of protein and 2.5 μl of reservoir to equilibrate against 500 μl reservoir solution.

Crystals were flash-cooled in a 100-K nitrogen-gas stream. A dataset to 2.50 Å resolution was collected using synchrotron radiation at wavelength 0.98 Å at beamline P11 of DESY, Hamburg. Diffraction data were processed with the program XDS (Kabsch, 2010). The space group was determined as C2, with unit-cell parameters a = 100.89 Å, b = 47.67 Å, c = 88.43 Å, β = 122.35°. Diffraction data statistics are given in Table 1.

2.3. Phase determination, model building and refinement

The structure of the MERS-CoV PL\textsuperscript{pro} was solved by molecular replacement using the program BALBES (Long et al., 2008). The program selected molecule A of the SARS-CoV PL\textsuperscript{pro} (PDB: 2FE8, Ratia et al., 2006) as the most suitable search model. The resulting model for the MERS-CoV PL\textsuperscript{pro} was inspected and rebuilt using Coot (Emsley et al., 2010), and refined using autoBUSTER (Bricogne et al., 2011). The final refinement statistics are presented in Table 1. Atomic coordinates and structure factors have been deposited in the PDB with accession code 4P16. All figures except Fig. 3 and
Supplementary Fig. 1 have been prepared using Pymol (Schrödinger; http://www.pymol.org/).

2.4. Site-directed mutagenesis

Using the pET28a-PLpro plasmid as template, site-directed mutagenesis (L106W mutation) was performed by PCR with the following primers (mutated codons shown in bold and underlined):
gtagcttcctatggtggaatat and acaattattacccattggagaaac. The PCR products were digested by DpnI (Thermo Scientific) and transformed to DH5α directly. The positive clones were incubated in LB medium overnight and the bacterial culture was harvested next day. The plasmids were purified from the bacteria using the GeneJET Plasmid Miniprep Kit (Thermo Scientific). All DNA plasmids were sequenced and the correctness of the mutation was verified. The procedures for the expression of the construct carrying the L106W mutation and the purification of the corresponding protein were the same as described above for the wild-type protein.

2.5. Assays of MERS-CoV PLpro in vitro activity

Before usage of the freshly prepared MERS-CoV PLpro for determination of the enzyme kinetics, the number of free (non-oxidized) cysteine residues was determined by titration with Ellman’s reagent (Riddles et al., 1983). All procedures followed the standard protocol of the Ellman’s Reagent kit (Thermo Scientific). The resulting number of 13.4 ± 0.3 free cysteine residues (out of 13 in the amino-acid sequence) showed that in bulk of our enzyme preparation, the catalytic Cys111 was in the free state and fully reduced.

We also examined the influence of ethylene diamine tetraacetic acid (EDTA) on the enzyme’s activity, but found no effect of concentrations up to 10 mM. However, at a concentration of 20 mM, a decrease of enzymatic activity was observed, presumably due to the removal of Zn2+ from the PLpro zinc finger. Accordingly, the kinetic assays were carried out in the absence of EDTA.

All enzymatic assays were performed in 20 mM Tris–HCl, 150 mM NaCl, pH 7.9, 2 mM dithiothreitol (DTT), using a 96-well microtiter plate. Three fluorogenic substrates, Cbz-Arg-Leu-Arg-Gly-Gly–7-amino-4-methylcoumarin (Z-LRGG-AMC) (Bachem), Z-LRGG-AMC and ubiquitin-AMC (Ub-AMC) (BostonBiochem), were used. The enzymatic cleavage reactions were monitored at 25 °C by measuring the increased fluorescence (λex: 360 nm; λem: 460 nm) resulting from AMC release, using an Flx800 fluorescence spectrophotometer (BioTek). Reactions were started by addition of the substrate to the microtiter plate. The kinetic assays were run under the following conditions: 1 μM MERS-CoV PLpro with different concentrations (10–100 μM) of Z-LRGG-AMC in a final volume of 100 μl, 1 μM MERS-CoV PLpro with different concentrations (20–160 μM) of Z-LRGG-AMC in a final volume of 100 μl, or 100 nM PLpro with different concentrations (0.2–1.2 μM) of Ub-AMC in a final volume of 50 μl. In case of the PLpro carrying the L106W mutation, the following conditions were used: 125 nM enzyme with different concentrations (10–100 μM) of Z-LRGG-AMC in a final volume of 100 μl, 250nM enzyme with different concentrations (40–100 μM) of Z-LRGG-AMC in a final volume of 100 μl, or 125 nM PLpro with different concentrations (0.1–1.0 μM) of Ub-AMC in a final volume of 50 μl. Initial velocities were determined from the linear section of the curve. Since no saturation could be observed, the data were fitted to the equation $v = [E]_{tot} - \frac{k_{app}}{K_{M}}$, where $k_{app}$ approximates $k_{cat}/K_{M}$ as described in Barretto et al. (2005) and Wojdyła et al. (2010). A calibration curve was generated by measuring the fluorescence of free AMC in reaction buffer at concentrations ranging from 0.005 μM to 2.5 μM.

3. Results and discussion

3.1. Overall structure of MERS-CoV PLpro

The MERS-CoV PLpro molecule is divided into two parts, the N-terminal ubiquitin-like (Ubl) domain and the catalytic domain (Fig. 1A). The Ubl domain consists of the 62 N-terminal amino-acid residues. It comprises five β-strands, one α-helix, and one 3_4-helix (η) in the order β1–β2–η1–β3–β4–η1–β5. The catalytic domain forms an extended right-hand scaffold, which comprises three distinct subdomains: thumb, fingers, and palm. The thumb domain consists of six α-helices and four β-strands with the order α2–α3–β6–β7–α4–β8–α5–α6–η7–β9. The fingers domain contains β10–β13, the N-terminal 16 residues (Glu231–Leu246) of β14, η2, η8, and the C-terminal 7 residues (Phe312–Ser318) of β19. It also includes a zinc finger, in which four cysteine residues (Cys191, Cys194, Cys226, and Cys228) coordinate a zinc ion (Fig. 1A and B). The palm domain comprises six β-strands: β15–β18, the C-terminal 10 residues (Ser247–Thr256) of β14, and the N-terminal 10 residues (Asp302–Leu311) of β19. In the loop Gly271–Gly277 between β15 and β16, residues Ile272–Ala275 are not defined by electron density, indicative of high flexibility. The substrate-binding site is a solvent-exposed region between the palm and thumb

Fig. 1. Structure of the MERS-CoV papain-like protease (PLpro). (A) Cartoon view of the enzyme’s overall structure. α-Helices (cyan) and β-strands (purple) are numbered, polypeptide segments devoid of repetitive secondary structure, including loops and turns, are brown. The ubiquitin-like (Ubl) domain is encircled by a red dashed line. The catalytic domain consists of the thumb, fingers, and palm subdomains. The structural zinc ion in the fingers domain is indicated by a gray sphere. The Cα atoms of the catalytic-site cysteine (111), histidine (278), and aspartate (293) residues are also shown (yellow, blue, and red sphere, respectively). The red arrow indicates the substrate-binding region and points to the catalytic site. (B) The four cysteine ligands (Cys191, Cys194, Cys226 and Cys228) and the structural zinc ion (gray sphere) in the zinc ribbon of the fingers domain. An F_{s}o-F_{v} omit density (green; contoured at 5 σ above the mean) for the zinc is shown. Sulfur atoms are shown in yellow, oxygen in red, nitrogen in blue, and carbon in light blue. The coordinate bonds between the sulfur atoms and the zinc ion are indicated by dashed red lines. (C) The catalytic triad: Cys111, His278, and Asp293. Atom colors: carbon, yellow; oxygen, red; nitrogen, blue. A 2Fo–Fv electron density (gray; contoured at 1.0 σ above the mean) is also displayed. CME111: Cys111 covalently modified by β-mercaptoethanol. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
domains. At the center of this region, the catalytic triad consisting of Cys111, His278, and Asp293 is located.

Interestingly, the segment 283-RLKGG/Li-289, located in the connection between the active-site residues His278 and Asp293, constitutes a potential autocleavage site for the MERS-CoV PLpro. Whereas we have some preliminary evidence for partial autoprocessing of our protease preparation at this site (data not shown), the electron density maps suggest that no cleavage has occurred in the crystallized protein at this position. Residues 283–285 are at the C-terminus of strand b16 and the two glycines are the central residues of a b-turn that leads into strand b17, which starts with Leu288. Thus, the residues concerned are part of well-defined secondary-structure elements and therefore, although largely accessible to solvent, cannot be accommodated by the substrate-binding site of the protease. However, whether or not in-trans PLpro autocleavage occurs at this position in the viral polyprotein, remains to be investigated.

3.2. Comparisons of the overall fold

3.2.1. The Ubl domain

As the Ubl domain and the catalytic domain of the MERS-CoV PLpro are two independent domains, we treated them separately in searching the Protein Data Bank (PDB) for structural similarity using the DALI server (Holm and Rosenström, 2010). However, we note that there are no variations in the relative orientation of these domains with respect to each other, as the Ubl domain is anchored to the core domain by two strong salt-bridges in MERS-CoV PLpro (Arg16...Glu64, 2.9 Å; Asp39...His81, 3.5 Å) as well as in SARS-CoV PLpro (the TGEV PL1pro lacks the Ubl domain). The structural comparisons show that the MERS-CoV Ubl domain is similar to other proteins or domains featuring the ubiquitin fold (Table 2, Fig. 2A). The Ubl domain of SARS-CoV PLpro is essential for downregulating IRF3 or the NF-kB antiviral signaling pathway, but the Ubl domain alone is not sufficient to do so (Frieman et al., 2009). On the other hand, the Ubl domain is often involved in protein–protein interactions (Mueller and Feigon, 2003; Su and Lau, 2009; Hartmann-Petersen and Gordon, 2004) and it is conceivable that the Ubl domain of MERS-CoV PLpro could interfere with host signaling-pathway proteins by providing a binding scaffold, which is necessary for the catalytic domain to accomplish its function of antagonizing the host’s innate immune response.

3.2.2. The catalytic domain: comparison with other coronavirus PLpros

The catalytic domain of the MERS-CoV PLpro is more similar to that of the SARS-CoV (Ratia et al., 2006) than to the PL1pro of Transmissible Gastroenteritis Virus (TGEV; Wojdyla et al., 2010) (Table 2). There are four regions of significant structural difference between the coronaviral PLpros (Fig. 2B). The first (Region I) concerns the two helices a2 and a3 in the N-terminal region of the thumb domain of the MERS-CoV and SARS-CoV enzymes, which are absent in the TGEV PL1pro (Figs. 2B and 3). The second region (II in Fig. 2B) of

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structural comparisons of MERS-CoV PLpro with other proteins.</strong></td>
</tr>
<tr>
<td><strong>MERS-CoV PLpro ubiquitin-like (Ubl) domain</strong></td>
</tr>
<tr>
<td>PDB/chain ID</td>
</tr>
<tr>
<td>SARS-CoV PLpro Ubl</td>
</tr>
<tr>
<td>DSK2</td>
</tr>
<tr>
<td>hHR23A</td>
</tr>
</tbody>
</table>

| **MERS-CoV PLpro catalytic domain (CD)** |
| PDB/chain ID | Z score | RMSD (Å) | Ca1 | % id2 | References |
| SARS-CoV PLpro CD | 2FE8/A | 27.4 | 2.4 | 246/251 | 32 | Ratia et al. (2006) |
| TGEV PL1pro | 2MP2/A | 18.7 | 3.1 | 198/211 | 23 | Wojdyla et al. (2010) |
| USP14 | 2AYN/B | 13.0 | 3.1 | 199/337 | 14 | Hu et al. (2005) |
| USP21 | 3M1U/A | 12.8 | 3.0 | 196/303 | 15 | Ernst et al. (2013) |
| USP2 | 2HDS/A | 12.8 | 3.1 | 198/315 | 15 | Renatus et al. (2006) |
| USP7 | 1N88/B | 10.5 | 3.5 | 198/333 | 15 | Hu et al. (2002) |

1. Aligned Ca atoms/total Ca atoms.
2. Sequence identity.
major structural differences concerns the four β-strands of the fingers domain. Two β-hairpins, which provide the cysteine residues for binding the structural zinc ion in this region, are twisted to different degrees among the three structures. In the MERS-CoV PL$^{\text{pro}}$, the C-terminal 10 residues of β14 extend into the palm domain and an 8-residue loop (Thr257–Val264) connects β14 and β15. In the SARS-CoV enzyme, the corresponding strand is divided into two separate β-strands, β12 and β13, the latter of which is mostly part of the palm domain, and a shorter loop formed by 5 residues (Gln256–Leu260) connects β13 and β14. In the TGEV PL$^{\text{pro}}$, there are two loops (Ser157–Thr160 and Val166–Val170) and one 3_10-helix (Pro161–Phe165) connecting the fingers and palm domains. The different connections in the various PL$^{\text{pro}}$s might lead to different mutual orientations of the two domains, which in turn might affect the enzymatic activities of the individual PL$^{\text{pro}}$s. Finally, Region IV (Fig. 2B) concerns the loop between β15 and β16 (Gly271–Gly277), four residues of which are not defined by electron density in MERS-CoV PL$^{\text{pro}}$. The corresponding region in TGEV PL$^{\text{pro}}$ (Gly267–Gly272) and in SARS-CoV PL$^{\text{pro}}$ (Gly267–Gly272) is shorter by one residue compared to the MERS-CoV enzyme; it forms a loop between β8 and β9 in TGEV PL$^{\text{pro}}$ and a short 3_10-helix between β14 and β15 in the SARS-CoV enzyme. However, this region is very flexible and adopts different positions in the three copies of the SARS-CoV PL$^{\text{pro}}$ in the asymmetric unit of the crystal (Ratia et al., 2006). In the recently reported crystal structure of the SARS-CoV PL$^{\text{pro}}$ (C125 mutant) in complex with ubiquitin (Chou et al., 2014), the β14–β15 loop shows large conformational differences compared to its position in the free enzyme. The two glycines framing this loop are absolutely conserved among the coronavirus PL$^{\text{pro}}$s, but the residues between them are different in each of the enzymes, suggesting that there must be differences in the interaction between the loop and the substrate.

3.2.3. Comparison with cellular ubiquitin-specific proteases

Even though the catalytic domain of the MERS-CoV PL$^{\text{pro}}$ only shares 12–15% sequence identity with the cellular ubiquitin-specific proteases (USPs), it features largely the same fold as the USPs with known three-dimensional structures (Table 2). A side-by-side comparison of the catalytic domain of the MERS-CoV PL$^{\text{pro}}$ with USP14 is shown in Fig. 2C. The most significant differences are located in the connecting region between the fingers and palm domains. The two “blocking loops”, BL1 and BL2, of USP14 regulate the deubiquitinating activity (Hu et al., 2005) (Fig. 2C). BL1 of USP14 connects the fingers and palm domains. It is a 22-residue loop between β8 and β9, exposed to the substrate-binding surface. The corresponding loop in MERS-CoV PL$^{\text{pro}}$ is much shorter (8 residues) and connects β14 and β15. Furthermore, this short loop does not face the substrate-binding region but rather points to the bottom of the thumb domain of MERS-CoV PL$^{\text{pro}}$ (Fig. 2C), and can thus not be considered a “blocking loop”. As a consequence, the substrate-binding site (indicated by the black arrow in Fig. 2C) is larger in MERS-CoV PL$^{\text{pro}}$ than in the USPs, probably enabling the enzyme to not only bind ubiquitin but also viral polyprotein. Connecting β10 and β11, the other blocking loop, BL2 (residues Gly427–Gly433), of USP14 is near the active site and undergoes conformational change upon substrate binding (Hu et al., 2005); this loop corresponds to β15–β16 (residues Gly271–Gly277) of MERS-CoV PL$^{\text{pro}}$ which unfortunately lacks electron density for four of its seven residues (see above).

3.3. The active site of MERS-CoV PL$^{\text{pro}}$

The MERS-CoV PL$^{\text{pro}}$ possesses a catalytic triad consisting of Cys111, His227, and Asp293 (Figs. 1C and 4A). Cys111 in the MERS-CoV PL$^{\text{pro}}$ (C112S mutant) in complex with ubiquitin (Chou et al., 2014), the β14–β15 loop shows large conformational differences compared to its position in the free enzyme. The two glycines framing this loop are absolutely conserved among the coronavirus PL$^{\text{pro}}$s, but the residues between them are different in each of the enzymes, suggesting that there must be differences in the interaction between the loop and the substrate.

Fig. 3. Structure-based alignment of MERS-CoV PL$^{\text{pro}}$ (GenBank: AFV9327.1), SARS-CoV PL$^{\text{pro}}$ (GenBank: AY278741.1), and TGEV PL$^{\text{pro}}$ (GenBank: AJ271965.2). Secondary-structure elements of MERS-CoV PL$^{\text{pro}}$ (top) and SARS-CoV PL$^{\text{pro}}$ (bottom) are indicated. Residues of the catalytic triad are marked by black asterisks. The four cysteine residues coordinating the zinc ion are marked by black inverted triangles. The mobile and a short 3_10-helix between (Gly267–Gly272) is shorter by one residue compared to the MERS-CoV enzyme; it forms a loop between β8 and β9 in TGEV PL$^{\text{pro}}$ and a short 3_10-helix between β14 and β15 in the SARS-CoV enzyme. However, this region is very flexible and adopts different positions in the three copies of the SARS-CoV PL$^{\text{pro}}$ in the asymmetric unit of the crystal (Ratia et al., 2006). In the recently reported crystal structure of the SARS-CoV PL$^{\text{pro}}$ (C125 mutant) in complex with ubiquitin (Chou et al., 2014), the β14–β15 loop shows large conformational differences compared to its position in the free enzyme. The two glycines framing this loop are absolutely conserved among the coronavirus PL$^{\text{pro}}$s, but the residues between them are different in each of the enzymes, suggesting that there must be differences in the interaction between the loop and the substrate.

The MERS-CoV PL$^{\text{pro}}$ possesses a catalytic triad consisting of Cys111, His227, and Asp293 (Figs. 1C and 4A). Cys111 in the MERS-CoV PL$^{\text{pro}}$ is situated at the N-terminus of α4 and points into the substrate-binding cleft between the palm and thumb domains.
There is clear electron density indicating that the catalytic cysteine has been modified by disulfide bond formation with β-mercaptoethanol (BME) during crystallization of the enzyme. His278 is located at the N-terminus of β16; its N61 atom is 4.4 Å from the

---

Fig. 4. Active site of MERS-CoV PL\textsuperscript{pren}. (A) Superposition (stereo view) of catalytic-triad residues of MERS-CoV PL\textsuperscript{pren} (thick green sticks), SARS-CoV PL\textsuperscript{pren} (thin purple sticks, PDB: 2FE8, chain A, Ratia et al., 2006), and USP14 (thin orange sticks, PDB: 2AYN, chain B, Hu et al., 2005). The distance between N61 of His278 and the sulfur of Cys111 in MERS-CoV PL\textsuperscript{pren} is indicated. CME111: Cys111 covalently modified by β-mercaptoethanol. (B) Superposition (stereo view) of the polypeptide segment preceding the catalytic cysteine in MERS-CoV PL\textsuperscript{pren} (106-LSDNN(CME)-111, green, shown as thick sticks, with bold green labels) and SARS-CoV PL\textsuperscript{pren} (107-WADNNC-112, purple, thin sticks, smaller purple labels). (C) Stereo view of the same β-turn, Ser107-Asp108-Asn109-Asn110, in MERS-CoV PL\textsuperscript{pren}, with the corresponding 2Fo-Fc electron density (blue; contoured at 1.5 σ above the mean). The main chain of the β-turn and of the (modified) catalytic residue CME111 is shown in pink, while side-chains are green. Gly161, Lys105 (main chain), and Trp93 (side-chain) are displayed as well. Hydrogen bonds are indicated by dashed red lines. (D) Stereo view of a superposition of 108-DNN(CME)-111 in MERS-CoV PL\textsuperscript{pren} (green) and 142-NGSC-145 in SARS-CoV M\textsuperscript{pren} (gray, PDB: 2BX3, Tan et al., 2005). Residues of MERS-CoV PL\textsuperscript{pren} and SARS-CoV M\textsuperscript{pren} (in brackets) are labeled. The RMSD of all main-chain atoms between this pair of four residues is ~0.62 Å. The two black arrows point to the presumable position of the oxyanion in the transition state of the proteolytic reaction. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
sulfur of Cys111. This large distance is likely due to the BME-modification of Cys111; in most papain-like protease structures, the distances are between 3.7 Å and 4.0 Å (Fig. 4A). Although the side-chain of His278 is somewhat displaced from its regular position in the catalytic triad, Asp293 and Cys111 align well with their counterparts in SARS-CoV PL\textsuperscript{pro} and USP14 (Fig. 4A).

The transition state of the proteolytic reaction catalyzed by papain-like cysteine proteases is stabilized in the oxyanion hole, through two hydrogen bonds usually donated by the main-chain amide of the catalytic cysteine residue and by the side-chain amide of a glutamine or asparagine residue five or six positions N-terminal to the catalytic cysteine. For example, the latter residue is Asn in USP14 and HAUSP (Hu et al., 2002, 2005), and Gln in the ubiquitin C-terminal hydrolases (UCH-L3 and UCH-L1 (Johnston et al., 1997). In TGEV PL\textsuperscript{pro}, there is also a glutamine residue (Gln27) at this position (Wojdyla et al., 2010). In SARS-CoV PL\textsuperscript{pro}, the corresponding residue is Trp107, the indole NH of which has been proposed to stabilize the oxyanion transition state (Fig. 4B). Accordingly, replacing Trp by Ala abrogates the protease activity (Ratia et al., 2006). Surprisingly, the corresponding residue in the MERS-CoV PL\textsuperscript{pro} is Leu106, the side-chain of which is incapable of hydrogen bonding (Fig. 4B).

How then is the oxyanion transition state stabilized in MERS-CoV PL\textsuperscript{pro}? Ratia et al. (2006) have discussed Asn110 as a potential additional component of the oxyanion hole in the SARS-CoV PL\textsuperscript{pro}. This residue is highly conserved amongst the coronaviral PL\textsuperscript{pro}s, including the MERS-CoV enzyme (Asn109) and in TGEV PL\textsuperscript{pro} (Asn30). In all three coronavirus PL\textsuperscript{pro}s of known three-dimensional structure as well as in USP14, this Asn residue occupies position i+2 of a β-turn (Ser107-Asp108-Asn109-Asn110 in MERS-CoV) that precedes the catalytic cysteine in the polypeptide chain. However, the side-chain of Asn109 is oriented away from the oxyanion hole and is involved in a strong (2.9 Å) hydrogen-bond with the carbonyl oxygen of conserved Gly161 of the thumb subdomain (Fig. 4C); therefore, it is unlikely to undergo a conformational change that will bring it into the neighborhood of the catalytic center. The other two potential hydrogen-bonding donors in this β-turn, Ser107 and Asn110, are heavily engaged in hydrogen bonds across the turn (Fig. 4C) and therefore equally unlikely to undergo the conformational changes necessary to reorient their side-chains towards the catalytic center.

So if none of the side-chains in this β-turn is a likely component of the oxyanion hole, what about main-chain amides? We notice that the segment between residues 108 (position i+1 of the β-turn) and 111 can be superimposed (with an RMSD of ∼0.62 Å for main-chain atoms) onto the loop (142-NGSC-145) preceding the catalytic nucleophile, Cys145, in SARS-CoV main protease (M\textsuperscript{pro}; PDB: 2BX3, Tan et al., 2005) (Fig. 4D). The M\textsuperscript{pro} is a cysteine protease comprising a chymotrypsin-like fold; in these enzymes, the oxyanion hole is formed by the main-chain amides of the catalytic cysteine (or serine) and of the penultimate residue (Taranto et al., 2008; Wu et al., 2013). The penultimate residue before the catalytic nucleophile in chymotrypsin-like proteases is absolutely conserved as a glycine, whereas in MERS-CoV PL\textsuperscript{pro}, the conserved Asn109 resides at this position. However, this asparagine is in a left-handed (α\textsubscript{L}) conformation, with positive ϕ and ψ angles (56° and 34°, resp.) in the Ramachandran plot, just like the conserved glycine in the catalytically competent conformation of chymotrypsin-like proteases (Verschueren et al., 2008; Tan et al., 2005). In fact, Asn and Asp are the only residues apart from Gly, for which this conformation is observed at a significant rate in protein structures (Hutchinson and Thornton, 1994). (Incidentally, in MERS-CoV PL\textsuperscript{pro}, the i+1 residue of this β-turn, Asp108, is also in an α\textsubscript{L} conformation, so that we have the rare situation here of a β-turn with both i+1 and i+2 residues having positive ϕ/ψ angles). In spite of the Asn109 main-chain N-terminus of Gly143 in SARS-CoV M\textsuperscript{pro}, we note that the N-H vector of this amide does not point towards the position that would likely be assumed by the oxyanion. Whether or not the necessary (minor) rearrangement may occur, remains to be answered by elucidating the structure of a complex between the MERS-CoV PL\textsuperscript{pro} and a transition-state analogue.

In any case, our mutation experiment described below clearly demonstrates that the oxyanion hole of the MERS-CoV PL\textsuperscript{pro} is deficient.

3.4. The substrate-binding site of MERS-CoV PL\textsuperscript{pro}

In the MERS-CoV polyproteins, the three cleavage sites for the PL\textsuperscript{pro}s are KLIGG\textsubscript{DV} (Nsp1-2), RLKGG\textsubscript{AP} (Nsp2-3), and KIVGG\textsubscript{AP} (Nsp3-4). The P1 and P2 positions are strictly conserved as glycine residues. P3 is Lys or Arg and P4 Leu or Ile, whereas P3 can be Ile, Lys, or Val. In the SARS-CoV polyprotein, cleavage by the PL\textsuperscript{pro} also occurs behind LXGG, but X is either Asn or Lys (Barretto et al., 2005), never a hydrophobic residue such as Ile or Val as in MERS-CoV. The LXGG motif is also present at the C-terminus of ubiquitin. In addition to the crystal structure of the complex between the C112S mutant of SARS-CoV PL\textsuperscript{pro} and ubiquitin (Chou et al., 2014), the structure of the wild-type SARS-CoV PL\textsuperscript{pro} with Ubal has very recently been reported (Ratia et al., 2014; Ubal is ubiquitin with the C-terminal carboxylate reduced to an aldehyde, which forms a covalent bond with the catalytic Cys residue). However, as the PDB coordinates for the latter structure have not been released yet, we rely for our further discussion on the structures of the ubiquitin complex of the SARS-CoV PL\textsuperscript{pro} (C112S mutant) and on the structure of USP14-Ubal; (Hu et al., 2005). The structural homology between these two enzymes and MERS-CoV PL\textsuperscript{pro} allowed us to deduce conclusions concerning the substrate-binding site of the latter. When we superimposed the substrate-binding sites of MERS-CoV PL\textsuperscript{pro}, SARS-CoV PL\textsuperscript{pro} (C112S), and USP14, we found a remarkable degree of structural conservation but also some important differences (Fig. 5).

In MERS-CoV PL\textsuperscript{pro}, the substrate-binding site is lined by residues Leu106–Tyr112 and Gly161–Arg168 of the thumb subdomain, and Phe269–Tyr279, Pro250, and Thr308 of the palm subdomain (Fig. 5C and D). Asn109, Cys111, the NH of Tyr112, Gly277, and His278 (all conserved in SARS-CoV PL\textsuperscript{pro} and USP14) form the spatially restricted S1 site (Fig. 5C and D), which can only accommodate glycine as the P1 residue. Pro163 (Leu163 in SARS-CoV PL\textsuperscript{pro}, Gln196 in USP14), Asp164 (main chain), Gly277 (conserved), and Tyr279 (side-chain, conserved) are involved in shaping the equally restricted S2 subsite, which again is specific for glycine (Fig. 5C and D). Replacement of Tyr274 in SARS-CoV PL\textsuperscript{pro} (corresponding to Tyr279 of the MERS-CoV enzyme) by Ala leads to a loss of protease activity (Barretto et al., 2005).

The S3 subsite of MERS-CoV PL\textsuperscript{pro} features important differences from the one in the SARS-CoV enzyme. In the SARS-CoV PL\textsuperscript{pro} (C112S)-Ub and SARS-CoV PL\textsuperscript{pro}-Ub complexes (Chou et al., 2014; Ratia et al., 2014), the main-chain amide of P3-Arg (substrate residues are indicated in italics in what follows) forms an H-bond with the side-chain OH of Tyr265 (Fig. 5A and S1A). In MERS-CoV PL\textsuperscript{pro}, this residue is replaced by Phe269 (Fig. 5C and D). Replacement of Tyr265 by Phe in SARS-CoV PL\textsuperscript{pro} reduces the peptidolytic activity of the enzyme by a factor of 2.4 and its deubiquitinating activity by 57% (Chou et al., 2014). Another potentially important difference between the S3 subsites of SARS-CoV PL\textsuperscript{pro} and MERS-CoV PL\textsuperscript{pro} is that Glu162 of the former is replaced by Ala162 in the latter (Fig. 5A and C), reducing the negative electrostatic potential of this region in the MERS-CoV PL\textsuperscript{pro} (Fig. 5E and F). This might enable the MERS-CoV enzyme to accommodate hydrophobic P3
residues such as Ile or Val, as they occur at the PLpro cleavage sites of the MERS-CoV (but not the SARS-CoV) polyprotein.

The S4 subsite in MERS-CoV PLpro is a hydrophobic pocket formed by Thr308, Pro250, Phe269, and the Cβ atom of conserved Asp165. The side-chain carboxylate of the latter is involved in defining specificity subsites in MERS-CoV PLpro, SARS-CoV PLpro, and USP14. Conserved residues are in red. Arg168 of MERS-CoV PLpro (corresponding to Glu168 and Glu201 in SARS-CoV PLpro and USP14, respectively) is in blue. Phe269 of the MERS-CoV enzyme (Tyr265 in SARS-CoV PLpro) is green. Residues contributing to S1, S2, and S4, but not S3 and S5, are conserved. (E, F) Electrostatic surfaces of the substrate-binding regions of SARS-CoV PLpro (E) and MERS-CoV PLpro (F), colored according to electrostatic potential (blue, positive potential; red, negative potential). The electrostatic surfaces were calculated using the APBS plugin in PyMOL (Baker et al., 2001). The contouring level is -8 kBT/e to 8 kBT/e. Orientation is the same as in the cartoon representations in (A and C). The S3 and S5 subsites are indicated by orange labels. (E) Glu162 and Glu168 of SARS-CoV PLpro are marked. (F) Ala162 and R168 of MERS-CoV PLpro are marked. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3

<table>
<thead>
<tr>
<th>Kinetic parameters of MERS-CoV PLpro</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Z-RLRGG-AMC</strong></td>
</tr>
<tr>
<td><strong>Z-RLRGG-AMC</strong></td>
</tr>
<tr>
<td><strong>Ub-AMC</strong></td>
</tr>
<tr>
<td>MERS-CoV PLpro</td>
</tr>
<tr>
<td>WT*</td>
</tr>
<tr>
<td>L106W</td>
</tr>
<tr>
<td>SARS-CoV PLpro</td>
</tr>
<tr>
<td>WT (Lindner et al., 2005)</td>
</tr>
<tr>
<td>WT (Barretto et al., 2005)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
The S5 subsite of MERS-CoV \( \text{PL}^{\text{pro}} \) is also different from that in USP14 and in the SARS-CoV \( \text{PL}^{\text{pro}} \). The side-chain of the PS-Arg forms a salt-bridge with Glu201 in the USP14-Ubal complex (Hu et al., 2005; Fig. 5B and S1B) and the same is observed for Glu168 of SARS-CoV \( \text{PL}^{\text{pro}} \) with Ubal (Ratia et al., 2014). However, in the SARS-CoV \( \text{PL}^{\text{pro}} \) (C112S) complex with ubiquitin (Chou et al., 2014), this same glutamate residue interacts with an arginine of the core of ubiquitin (Arg42) instead of the PS-Arg (Fig. 5A). This may be due to the position of Glu168 and hence the entire S5 site of \( \text{PL}^{\text{pro}} \) at the surface of the protease, near the entrance to the substrate-binding cleft. While not forming a pocket, the S5 is still important for substrate binding, because the Glu168 side-chain is at the core of a region with pronounced negative electrostatic potential (see Fig. 5E). Interestingly, this important glutamate of USP14 and SARS-CoV \( \text{PL}^{\text{pro}} \) is replaced by Arg168 in MERS-CoV \( \text{PL}^{\text{pro}} \) (Fig. 5D), changing the electrostatic potential of the S5 site of the latter to more positive and therefore less ideal for interacting with the PS-Arg or Arg42 of ubiquitin, or the PS Lys/Arg of the MERS-CoV polyprotein cleavage sites (Fig. 5F). Indeed, Chou et al. (2014) have replaced Glu168 of SARS-CoV \( \text{PL}^{\text{pro}} \) by Arg and found a 25-fold decrease in deubiquitinating activity of the enzyme.

Meesecar and colleagues have demonstrated that the SARS-CoV \( \text{PL}^{\text{pro}} \) is a druggable target. They used their crystal structure of the enzyme (Ratia et al., 2006) in a virtual screening campaign and reported crystal structures of the complex with the optimized hit compound, GRD617 (5-amino-2-methyl-N-[1R]-1-naphthyl-1-yethylbenzamide) and derivatives thereof (Ratia et al., 2008; Baez-Santos et al., 2014). Interestingly, the inhibitors do not bind to the catalytic center directly but near the S3 and S4 subsites as well as the mobile \( \beta 14–\beta 15 \) loop (\( \beta 15–\beta 16 \) in MERS-CoV \( \text{PL}^{\text{pro}} \)) (Ratia et al., 2008; Baez-Santos et al., 2014). As there are important differences in the S3 subsite as well as in the mobile loop between the SARS-CoV and the MERS-CoV \( \text{PL}^{\text{pro}} \), it is unlikely that these compounds will inhibit \( \text{PL}^{\text{pro}} \) and the replication of MERS-CoV. However, virtual or real screening approaches might be expected to result in the discovery of different small-molecule compounds binding to a similar site in MERS-CoV \( \text{PL}^{\text{pro}} \).

3.5. In-vitro peptide-hydrolysis and deubiquitinating activities of wild-type MERS-CoV \( \text{PL}^{\text{pro}} \) and its L106W variant

We tested the \( \text{PL}^{\text{pro}} \)-catalyzed hydrolysis of the peptides Z-RLRG7-\( \text{AMC} \) and Z-LRG7-\( \text{AMC} \) as mimics of the cleavage sites in the viral polyproteins. For this assay (and the one described below), only freshly prepared enzyme was used, in which the catalytic cysteine residue was shown to be in a free, reduced (and hence active) state by titration with Ellman's reagent and by the absence of an effect of adding EDTA up to a concentration of 10 mM (see Materials & Methods). We found that the peptides were hydrolyzed by MERS-CoV \( \text{PL}^{\text{pro}} \) in vitro, and that the initial rate of hydrolysis increased with raising the substrate concentration. However, we were unable to observe saturation of the reactions. This might indicate a large \( K_{\text{m}} \) value and/or low enzyme efficiency. We used the pseudo-first-order rate constant, \( k_{\text{app}} \), to estimate an approximate \( k_{\text{cat}}/K_{\text{m}} \) value. The \( k_{\text{app}} \) rates were 1.31 ± 0.14 min⁻¹ mM⁻¹ for Z-RLRG7-\( \text{AMC} \) and 1.00 ± 0.01 min⁻¹ mM⁻¹ for Z-LRG7-\( \text{AMC} \), values significantly lower than those reported for the SARS-CoV \( \text{PL}^{\text{pro}} \) (Table 3). This may be explained by the apparent deficiency of the oxyanion hole of the MERS-CoV \( \text{PL}^{\text{pro}} \) and the inability of the side-chain of Phe269 to form the hydrogen bond with the main-chain amide of P3-Arg observed for the corresponding Tyr265 in SARS-CoV \( \text{PL}^{\text{pro}} \) (Chou et al., 2014). In order to examine whether this relatively low activity is caused by the deficiency of the oxyanion hole, i.e. the lack of a side-chain available for oxyanion stabilization through hydrogen bonding, we prepared the L106W mutant of the MERS-CoV \( \text{PL}^{\text{pro}} \) and found its peptide-hydrolyzing activities on Z-RLRG7-\( \text{AMC} \) and Z-LRG7-\( \text{AMC} \), respectively, to be 40-fold and 60-fold higher than those of the wild-type enzyme (Table 3). This is a clear demonstration of the deficiency of the oxyanion hole in the MERS-CoV \( \text{PL}^{\text{pro}} \) and shows that the contribution to oxyanion stabilization by the main-chain amide of residue 109, if at all existent, is far less efficient than by a tryptophan residue in position 106, as found in the SARS-CoV \( \text{PL}^{\text{pro}} \).

We also tested the deubiquitinating activity of the MERS-CoV \( \text{PL}^{\text{pro}} \) in vitro using Ub-AMC as a substrate. We determined \( k_{\text{app}} \) as (5.06 ± 0.35) x 10⁻⁶ min⁻¹ mM⁻¹. This value is also lower than the rate constants reported for SARS-CoV \( \text{PL}^{\text{pro}} \) by different research groups (Table 3). Again, the less-than-perfect oxyanion hole and the replacement of Tyr265 of SARS-CoV \( \text{PL}^{\text{pro}} \) by Phe269 in MERS-CoV \( \text{PL}^{\text{pro}} \) may be responsible for this. In addition, the replacement of Glu168 of SARS-CoV \( \text{PL}^{\text{pro}} \) by Arg168 in the MERS-CoV enzyme very likely weakens the interaction with ubiquitin. The deubiquitinating activity of the MERS-CoV \( \text{PL}^{\text{pro}} \) carrying the L106W mutation was found to be 3.4-fold higher than that of the wild-type enzyme (Table 3). Our results demonstrate that the MERS-CoV \( \text{PL}^{\text{pro}} \) displays both peptidase and deubiquitinase activities in vitro and is thus suitable for screening chemical libraries for inhibitors.

4. Conclusions

The crystal structure of the MERS-CoV \( \text{PL}^{\text{pro}} \) provides critical information on this important potential drug target. The unique architecture of the oxyanion hole, which differs from all papain-like proteases that have been structurally characterized so far, may be of fundamental interest. Because of unique features of its S3 and S5 subsites as well as in the flexible loop (\( \beta 15–\beta 16 \)) covering the substrate-binding site, we expect that the MERS-CoV \( \text{PL}^{\text{pro}} \) will show differences in ubiquitin recognition, compared to SARS-CoV \( \text{PL}^{\text{pro}} \) and USP14. Both peptidase and deubiquitinizing activities of MERS-CoV \( \text{PL}^{\text{pro}} \) have been demonstrated in vitro. Introduction of the L106W mutation leads to a restoration of the oxyanion hole of the \( \text{PL}^{\text{pro}} \) and an enhancement of both catalytic activities. Furthermore, the structural differences from homologous host enzymes such as USP14 should allow the design of antivirals devoid of too many side effects.

Acknowledgements

Technical assistance by Susanne Zoske is gratefully acknowledged. We thank Linlin Zhang and Guido Hansen for discussion. We also acknowledge the staff at beamline P11 of DESY, Hamburg, Germany. This work was supported by the European Commission through its “SILVER” project (contract No. HEALTH-F3-2010-260644) and by the German Center for Infection Research (DZIF). RH acknowledges support by the DFG Cluster of Excellence “Inflammation at Interfaces” (EXC 306). This work is dedicated to Professor Wolfram Saenger on the occasion of his 75th birthday.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.antiviral.2014.06.011.

References