Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
The SARS-coronavirus papain-like protease: Structure, function and inhibition by designed antiviral compounds

Yahira M. Báez-Santos 1, Sarah E. St. John 1, Andrew D. Mesecar *

Department of Biological Sciences, Purdue University, West Lafayette, IN, USA
Department of Chemistry, Purdue University, West Lafayette, IN, USA
Center for Drug Discovery, Purdue University, West Lafayette, IN, USA
Center for Cancer Research, Purdue University, West Lafayette, IN, USA

Abstract

Over 10 years have passed since the deadly human coronavirus that causes severe acute respiratory syndrome (SARS-CoV) emerged from the Guangdong Province of China. Despite the fact that the SARS-CoV pandemic infected over 8500 individuals, claimed over 800 lives and cost billions of dollars in economic loss worldwide, there still are no clinically approved antiviral drugs, vaccines or monoclonal antibody therapies to treat SARS-CoV infections. The recent emergence of the deadly human coronavirus that causes Middle East respiratory syndrome (MERS-CoV) is a sobering reminder that new and deadly coronaviruses can emerge at any time with the potential to become pandemics. Therefore, the continued development of therapeutic and prophylactic countermeasures to potentially deadly coronaviruses is warranted. The coronaviral proteases, papain-like protease (PLpro) and 3C-like protease (3CLpro), are attractive antiviral drug targets because they are essential for coronaviral replication. Although the primary function of PLpro and 3CLpro are to process the viral polyprotein in a coordinated manner, PLpro has the additional function of stripping ubiquitin and ISG15 from host-cell proteins to aid coronaviruses in their evasion of the host innate immune responses. Therefore, targeting PLpro with antiviral drugs may have an advantage in not only inhibiting viral replication but also inhibiting the dysregulation of signaling cascades in infected cells that may lead to cell death in surrounding, uninfected cells. This review provides an up-to-date discussion on the SARS-CoV papain-like protease including a brief overview of the SARS-CoV genome and replication followed by a more in-depth discussion on the structure and catalytic mechanism of SARS-CoV PLpro, the multiple cellular functions of SARS-CoV PLpro, the inhibition of SARS-CoV PLpro by small molecule inhibitors, and the prospect of inhibiting papain-like protease from other coronaviruses. This paper forms part of a series of invited articles in Antiviral Research on “From SARS to MERS: 10 years of research on highly pathogenic human coronaviruses.”

© 2014 Elsevier B.V. All rights reserved.
### 1. Introduction

#### 1.1. SARS-CoV genome and replication

Coronavirus (CoV) replication is a highly orchestrated process involving complex replication machineries to protect the virus genome, the largest known RNA genome, and the viral proteins from the host’s antiviral defense mechanisms (Brian and Baric, 2005). Before the zoonotic emergence of the first human coronavirus that caused the 2002/2003 epidemic of severe acute respiratory syndrome (SARS-CoV), there was a paucity of information pertaining to CoV genomes and mechanisms of replication. Since then, the numbers of CoVs sequenced and studied have increased dramatically and several potential drug targets have emerged. Many of the SARS-CoV drug targets are encoded in the 5’-terminal two thirds of the genome, within the two open reading frames (ORF) that encode for the replicase or non-structural proteins (nsps) (see Fig. 1) (Tong, 2009).

Production of the replicase proteins is initiated by the translation of ORF1a and ORF1ab via a -1 ribosomal frame-shifting mechanism (Bredenbeek et al., 1990). This mechanism produces two large viral polyproteins, pp1a and pp1ab, that are further processed by two virally encoded cysteine proteases (Fig. 1), the papain-like protease (PLpro) and a 3C-like protease (3CLpro), which is sometimes referred to as main protease (Mpro) (Thiel et al., 2003; Ziebuhr, 2004; Ziebuhr et al., 2000, 2001). Processing of the viral polyprotein is required for the release and maturation of 16 viral polyproteins (non-structural proteins or nsps) involved in the formation of a membrane-associated, cytoplasmic enzyme complex, the replicase complex, which is responsible for directing the replication and transcription of the viral genome. It is thought that the establishment of viral replication sites is initiated by the recruitment of replicase proteins to host membranes, a process mediated by several viral transmembrane domain-containing proteins such as the nsp3, nsp4 and nsp6 (Kanjanahaluethai et al., 2007; Oostra et al., 2007, 2008; van Hemert et al., 2008).

Electron tomography and three-dimensional reconstruction imaging studies have revealed that the replicase complex utilizes a reticulovesicular network of double-membrane vesicles (DMVs) with interconnected outer membranes originating from the endoplasmic reticulum (ER) (Angelini et al., 2013; Gosert et al., 2002; Knoops et al., 2008). As a result, the viral replication mechanism is localized, increasing the concentration and cooperation of viral macromolecules and forming a framework of RNA synthesis. Most importantly, this mechanism provides a microenvironment for the protection of viral RNA from host nucleases, and of double-stranded RNA intermediates from the host cell's innate immune surveillance. Among the components of the replicase complex, nsp3 is of special interest since it is believed to be part of the central scaffolding protein of the replicase complex due to the large number of interactions with other nsps (Angelini et al., 2013; Imbert et al., 2008; Snijder et al., 2003).

#### 1.2. The multi-domain protein nsp3

The SARS-CoV nsp3 multi-domain protein is the largest replicase subunit at 1,922 amino acids (Snijder et al., 2003; Thiel et al., 2003). Nsp3 is thought to play an essential role during the formation of virus replication complexes via its insertion into host membranes and its numerous interactions with other nsps (Imbert et al., 2008) especially nsp4 and nsp6 (Angelini et al., 2013). Numerous domains have now been identified in nsp3 (Fig. 1) and many are predicted to be conserved in all CoV (Neuman et al., 2008). Due to the large size of the nsp3 multi-domain protein, in vitro and in cellular studies have mainly utilized truncated nsp3 constructs that represent the predicted domains boundaries (Fig. 1). Using this approach, the three dimensional structures of most of the domains from the nsp3 of SARS-CoV have been determined by X-ray crystallography or NMR spectroscopy (Chatterjee et al., 2009; Egloff et al., 2006; Johnson et al., 2010; Ratia et al., 2006; Saikatendu et al., 2005; Serrano et al., 2007, 2009; Tan et al., 2009).

The N-terminal region of the nsp3 (181–1000) is highly conserved among CoVs, containing a ubiquitin-like (Ubl) globular fold followed by a flexible, extended acidic-domain (AC domain) rich in glutamic acid (38%) (Serrano et al., 2007). Next to the AC domain is a catalytically active ADP-ribose-1'-phosphatase (ADRP,
1.3. The SARS-CoV PLpro domain within the nsp3

The SARS-CoV PLpro catalytic domain is flanked by numerous catalytically active enzymes, transmembrane domains, and domains of unknown function, and the entire nsp3 is localized to the ER membranes where the majority of the domains reside in the cytosol of the cell (Fig. 1) (Hagemeijer et al., 2010; Oostra et al., 2008). In the cytosol, the membrane associated PLpro domain recognizes the P4–P1 consensus cleavage sequence LXGG, found in the boundaries of nsp1/2, nsp2/3 and nsp3/4 where membrane association is required for cleavage of the nsp3/4 (Han et al., 2005; Harcourt et al., 2004). Proteolytic cleavage of the peptide bond after the glycine at position P1 results in the release of nsp1, nsp2 and nsp3 from the viral polyprotein (Fig. 1, left bottom panel), a process that is essential for viral replication. Therefore, SARS-CoV PLpro is proposed to be an excellent candidate as a drug target for the development of anti-CoV therapeutics.

1.4. SARS-CoV PLpro is a protease, a deubiquitinating (DUB) and deISGylating enzyme

Reminiscent of the overall architecture of human deubiquitinating enzymes (DUBs) in the ubiquitin specific protease family (USP), the molecular structure of the PLpro catalytic domain consists of a canonical, right-handed thumb–palm–fingers architecture which is flanked at the N-terminus by an additional ubiquitin-like (UBL) domain of unknown function (Fig. 2a and b) (Ratia et al., 2006). The in vitro characterization of PLpro enzymatic activities reveal that PLpro can recognize and hydrolyze the cellular proteins ubiquitin (Ub) and ISG15 from host cell proteins.

In the context of viral replication, the membrane associated PLpro domain is flanked at the N-terminus by an additional ubiquitin-like (UBL) domain of unknown function, which is conserved in betacoronaviruses and gammacoronaviruses, and one uncharacterized domain termed the marker domain (G2M) (Neuman et al., 2008). Following the G2M are two predicted double-pass transmembrane domains (TM1–2 and TM3–4) (Neuman et al., 2008; Snijder et al., 2003), a putative metal binding region (ZN) and the Y domain of unknown function (subdomains Y1–3) (Snijder et al., 2003; Thiel et al., 2003; Ziebuhr et al., 2001). Interestingly, comparative genome and proteome analyses of two bovine CoV (BCoV) isolates showed a predominant clustering of mutations within the nsp3 multi-domain (Choulenko et al., 2001). Consequently, the multi-functionality of the nsp3, the frequency of point mutations observed in nsp3 domains, and the involvement of nsp3 in structural arrangements of the replicase complex and double-membrane vesicles may engender pleiotropic effects, not only in SARS-CoV pathogenicity, but also on future emerging coronaviruses (Snijder et al., 2003).
Kinetic studies on the catalytic efficiency of SARS-CoV PLpro toward different substrates have shown that ISGylated and ubiquitinated substrates are more readily hydrolyzed than small peptide substrates such as RLRGG-AMC, which represents the C-terminus sequence of Ub and ISG15, suggesting a more complex mechanism for substrate recognition that extends well beyond the S4 to S1 enzyme recognition subsites for the LXGG peptide (Baez-Santos et al., 2014b; Lindner et al., 2005, 2007; Ratia et al., 2014). More recently, a comprehensive analysis of PLpro substrate specificity by X-ray crystallography and mutational analyses demonstrated that two distinct Ub binding subsites (SUB1 and SUB2) exist distant from the catalytic site, providing SARS-CoV PLpro a unique bivalent mechanism of interaction with Ub-like modifier substrates (Ratia et al., 2014). For most USPs and SARS-CoV PLpro, the primary Ub/UBL binding subsite (SUB1) is distal to the isopeptide bond and located at the boundaries of palm domain and fingers regions (Fig. 2c). For SARS-CoV PLpro, a second, distal Ub binding subsite (SUB2) exists which is located in a ridge region of the thumb domain (Fig. 2c). This region provides interactions for a second Ub molecule on a K48-linked di-Ub chain, and for ISG15, which has structural resemblance to K48-linked di-Ub molecules (Ratia et al., 2014).

1.5. SARS-CoV PLpro innate immune functions

The DUB and deISGylating activities of SARS-CoV PLpro have significant functional implications in the innate immune response during SARS-CoV infection. Both Ub and ISG15 are important signaling elements of the host innate immune response against viral infection, which can be negatively regulated by viral DUB and deISGylating enzymes (Calistrì et al., 2014). SARS-CoV PLpro has been shown to act as a strong antagonist of many Ub-dependent cellular responses to viral infection (Mielech et al., 2014a). Although the mechanism of PLpro-mediated antagonism of cellular pathways is not well understood, the evidence strongly suggests that catalytic activity is important for antagonism and, therefore, DUB and deISGylating activities have been proposed as a mechanism.

Because the multiple cellular roles of CoV PLpro DUB/deISGylating activities have been reviewed elsewhere (Mielech et al., 2014a), only a brief overview is presented here and the reader should visit other sources for a complete assessment on this topic. SARS-CoV PLpro antagonistic activities have been shown to block the production of important cytokines involved in the activation of the host’s innate immune response against viral infection, including the Type I interferon (IFN-β) and chemokines such as CXCL10 and CCL5. Based on the in vitro and cell culture data available on SARS-CoV PLpro and other homologous coronaviral PLPs, a summary of the possible mechanisms for how PLpro can interfere with the expression of proinflammatory cytokines, such as interferon β (IFN-β), and block the establishment of the antiviral state is illustrated in Fig. 3. Initially, SARS-CoV PLpro was shown to interfere with the activation of the transcription factors IRF3 (Devaraj et al., 2007) and NF-κB (Clementz et al., 2010; Frieman et al., 2009). Devaraj et al. showed that SARS-CoV PLpro-mediated antagonism of Type I interferon (IFN) production functions upstream of IRF3 activation by blocking IRF3 phosphorylation, homodimerization, and consequently nuclear translocation (Devaraj et al., 2007; Sun et al., 2012). On the other hand, Frieman et al. demonstrated that PLpro stabilizes the NF-κB inhibitor, IκBα, and thereby blocks the activation of the NF-κB signaling pathway (Frieman et al., 2009). More recently, Mielech et al. demonstrated SARS-CoV PLpro ability to decrease endogenous levels of proinflammatory cytokines and chemokines in activated cells (Mielech et al., 2014b).

Interestingly, in some instances, mutation of the active site catalytic cysteine residue of SARS-CoV PLpro only had a small decrease in antagonism of certain pathways, suggesting that the protease activity is not the only requirement for antagonism (Devaraj et al., 2007). However, Clementz et al. demonstrated that the SARS-CoV PLpro-mediated block of NF-κB activation can be reversed by the addition of a PLpro competitive inhibitor,
emphasizing the importance for catalytic activity during antagonism of this cellular pathway (Clementz et al., 2010). Nevertheless, it is important to consider that a mutation of the catalytic machinery of PLpro may not affect its ability to interact with Ub and ISG15 molecules as the competitive inhibitor does. Therefore, a protein-to-protein interaction mediated by PLpro's strong affinity for Ub/ISG15 molecules on ubiquitinated/ISGylated proteins could still lead to the interference of the IRF3/NF-κB signaling pathways.

2. SARS-CoV PLpro structure and function

2.1. Active site structure and catalytic mechanism of SARS-CoV PLpro

2.1.1. Active site structure

SARS-CoV PLpro belongs to the peptidase clan CA (family C16). The active site contains a classic catalytic triad, composed of Cys112–His273–Asp287, that is well-aligned, functional, and located at the interface of the thumb and palm sub-domains (Fig. 4) (Ratia et al., 2006). The catalytic cysteine (Cys112) is situated at the foot (N-terminus) of α-helix α4 in the thumb domain. The side chain sulfur atom of Cys112 is positioned 3.7 Å from the pros(τ)-nitrogen atom of the catalytic histidine (His273) (McNaught and Wilkinson, 1997), which is located at the foot of the palm domain and adjacent to the flexible loop BL2 (also called the G267–G272 loop (Ratia et al., 2008) or β-turn (Baez-Santos et al., 2014a)). One of the oxygen atoms of the side chain of catalytic aspartic acid (Asp287) is located 2.7 Å from the tele(τ)-nitrogen of the catalytic histidine at the foot of the palm domain (Fig. 4). The side chain of Trp107 is located within the oxyanion hole and the indole-ring nitrogen is proposed to participate in the stabilization of negatively-charged tetrahedral intermediates produced throughout catalysis (Ratia et al., 2014; Ratia et al., 2008). The overall tertiary structure of PLpro is remarkably similar to the cellular ubiquitin specific proteases (USP) including USP14 and USP7 (or HAUSP) (Fig. 2b). However, in their unbound states, the catalytic triad of PLpro aligns well structurally only with the catalytic triad of USP14 and not USP7. In order for the catalytic triad of USP7 to adopt the proper orientation for catalysis, a substrate-induced alignment must occur upon association with ubiquitin or other regulatory domains (Hu et al., 2002; Molland et al., 2014; Ratia et al., 2014a).
et al., 2006). The fingers domain of PLpro, which contains a zinc ion that is tetrahedrally coordinated by four cysteines, is essential for catalysis because it maintains the structural integrity of PLpro (Barretto et al., 2005; Ratia et al., 2006).

2.1.2. Catalytic mechanism

SARS-CoV PLpro functions through a proposed cysteine protease catalytic cycle where Cys112 acts as nucleophile, His273 functions as a general acid-base, and Asp287 is paired with histidine helping to align it and promote deprotonation of Cys112 (Fig. 5). Currently, there are no experimental studies defining the protonation state of Cys112 prior to nucleophilic attack on the substrate, or whether the reactive nucleophilic species is the thiolate ion of the thiolate–histidine pair, or the neutral thiol, that acts via a mechanism analogous to that of serine peptidases where a thionium ion must be produced after reaction with the ligand.

Fig. 4. Catalytic center of the active site of SARS-CoV PLpro. Unliganded X-ray crystal structure of SARS-CoV PLpro (PDB: 2FEB). The catalytic triad residues, Cys112, His273 and Asp287 are shown along with the oxyanion hole-stabilizing residue Trp107 (right panel). The flexible BL2 loop that interacts with the substrates is indicated as are the zinc ion (purple sphere) and α-helix 9/4 in the thumb domain. Distances between atoms are given in units of Ångströms (Å) and are indicated by dashed lines.

Fig. 5. Catalytic cycle and proposed chemical mechanism of the SARS-CoV PLpro catalyzed reaction. Active site residues of the catalytic triad (Cys112, His273, Asp287) and oxyanion hole residue Trp107 are shown in black. The peptide substrate is shown in green and a catalytic water molecule is shown in blue.
Because of the extremely low pKa of the thionium intermediate that must be produced in this case, we have chosen to show the classically accepted nucleophilic thiolate form in our general mechanism.

In its unliganded form, "E", the catalytic residues of SARS-CoV PLpro are within hydrogen-bonding distance of one another, indicating that the protonation state of Cys112 may exist in equilibrium with that of His273 and that the event of substrate binding drives this equilibrium toward the reactive thiolate (Figs. 4 and 5, step a). Alternatively, deprotonation of Cys112 may occur during substrate binding to form the "ES" complex (Fig. 5, step a). An addition–elimination sequence follows, where attack of the thiolate of Cys112 on the carbonyl carbon of the peptide bond forms the first negatively-charged tetrahedral intermediate "TI-1" or "FP," (Fig. 5, step b). The oxyanion of the tetrahedral intermediate is stabilized via the presence of an oxyanion hole adjacent to the PLpro enzymatic active site (Chou et al., 2014; Ratia et al., 2006, 2014). This oxyanion hole contains a triptophan residue (Trp107), which our group has confirmed is crucial for the enzymatic function of PLpro through site-directed mutagenesis (Baez-Santos, 2012). In addition, we recently showed via the X-ray structure of SARS-CoV PLpro covalently modified with ubiquitin-aldehyde that indole-ring nitrogen of Trp107 acts as a hydrogen-bond donor to the hemithioacetal intermediate (Ratia et al., 2014), confirming its function in the oxyanion hole. Furthermore, an asparagine residue (Asn110, not depicted) has been found to be highly conserved among coronaviral PLP2s and may contribute to stabilization in the oxyanion hole, as it is positioned above the catalytic cysteine. Elimination of the C-terminal amine of the substrate cleaves the peptide bond, subsequently forming the thioester intermediate "F" (Fig. 5, step c). A second addition–elimination sequence follows, where the nucleophilic addition of water to the carbonyl carbon of the thioester forms a new, negatively-charged tetrahedral intermediate "TI-2" or "FQ" (Fig. 5, step d). The oxyanion of this intermediate is again stabilized via the oxyanion hole of the PLpro active site. Elimination of cysteine from the tetrahedral intermediate results in the formation of the N-terminal carboxylic acid "EQ," which may be transiently stabilized in the PLpro active site via a hydrogen-bond between the carbonyl carbon of the newly formed acid and the indole ring nitrogen of Trp107 (Fig. 5, step e) (Chou et al., 2014). Elimination of the cleaved N-terminus of the peptide substrate then completes the catalytic cycle where the final product "Q" is released from the active site and the free enzyme "E" is regenerated (Fig. 5, step f).

3. SARS-CoV PLpro inhibitor reactions with Cys112

Cysteine proteases are known to react with a variety of electrophilic or “warhead” functionalities within covalent inhibitor molecules. These warhead inhibitors typically function by first forming a noncovalent interaction complex within the cysteine protease active site, where the warhead group of the inhibitor is positioned in the proximity of the reactive cysteine nucleophile. The reaction proceeds via a nucleophilic attack of the thiolate on the electrophilic carbon of the warhead group, forming a covalently modified enzyme–inhibitor complex, which subsequently inactivates the enzyme. Examples of such reactive warhead groups that are known to inhibit cysteine proteases include aldehydes, epoxy-ketones, Michael acceptors, activated ketones, activated esters, vinyl sulfones, acrylamides, alkynes, alkylhalides, and nitriles (Fig. 6) (Ghosh et al., 2005, 2006; Jacobs et al., 2013; Jain et al., 2004; Xue et al., 2008; Yang et al., 2006; Zhang et al., 2007).

Our lab successfully solved the first X-ray crystal structure of SARS-CoV PLpro covalently modified with an electrophilic “warhead.” (Ratia et al., 2014). The SARS-CoV PLpro enzyme was co-crystallized with a semisynthetic version of ubiquitin, where the C-terminus of ubiquitin was modified with an aldehyde functional group (UbAl). This intentional modification of ubiquitin was introduced as a way to modify the catalytic cysteine with a covalent-reversible group thereby trapping the PLpro–UbAl complex and mimicking the reaction intermediate "TI-2" in Fig. 5. The SARS-CoV PLpro–UbAl structure was determined to a resolution of 2.5 Å and is the first X-ray crystal structure of the hemithioacetal, tetrahedral intermediate in SARS-CoV PLpro (Fig. 7a and b). The catalytic cysteine is positioned 1.6 Å from the carbon of the tetrahedral intermediate, a distance irrefutably indicating a covalent C–S bond. The tetrahedral, hemithioacetal intermediate is stabilized via the oxyanion hole of SARS-CoV PLpro by a hydrogen bond to the indole nitrogen of Trp107, and is also within hydrogen-bonding distance to the pros(φ)-nitrogen atom of His273.

4. SARS-CoV PLpro inhibitor classes

4.1. Inhibitors identified via a designed yeast-based screen

In 2011, two new PLpro inhibitors were reported by Frieman and coworkers (Frieman et al., 2011). These inhibitors were identified via a yeast-based screening assay that was developed to discover inhibitors of SARS-CoV PLpro. The assay involved the expression of SARS-CoV PLpro in Saccharomyces cerevisiae that results in a slow-growth phenotype. A 2000 compound NIH Diversity Set library was screened for molecules capable of reversing the PLpro-induced slow-growth phenotype. Compounds identified as hits in this primary assay were then tested for efficacy against SARS-CoV replication and inhibition of PLpro function in cell culture models. This screening method resulted in the identification of NSC158362, a compound capable of specifically inhibiting SARS-CoV replication in cell culture without cytotoxic effects. However, NSC158362 was found incapable of inhibiting the protease, deubiquitase and anti-interferon functions of PLpro – pointing to a potential new mechanism for inhibition of SARS-CoV replication. A second compound, NSC158011 was found capable of inhibiting the protease activity of PLpro in cell-based assays; however, it was unsuccessful at blocking SARS-CoV viral replication. The structures of NSC15832 and NSC158011 are shown in Fig. 8a (Frieman et al., 2011).

4.2. Thiopurine compounds

Chou and coworkers screened a small library of 960 compounds for the inhibition of SARS-CoV PLpro using a FRET-based assay and purified SARS-CoV PLpro enzyme (Chen et al., 2009; Chou et al., 2008). Two thiocarbonyl-containing compounds, 6MP and 6TG, were found to be effective inhibitors of PLpro with IC50 values of 21.6 and 5 μM, respectively. In their assay, zinc ion (Zn2+) was also found to be an effective inhibitor of PLpro (vide infra). Both 6MP and 6TG were found to be slow-binding, competitive, reversible, and selective inhibitors of SARS-CoV PLpro. Through comparison to molecular analogs of 6MP and 6TG that lacked the thiocarbonyl functionality, the thiocarbonyl was found to be the crucial moiety for these compound’s inhibition of SARS-CoV PLpro and it is postulated to react covalently with the active site cysteine. Currently, both 6MP and 6TG are used clinically in the treatment of leukemia, and the anti-cancer action of 6TG is dependent upon its conversion to 6TG nucleotides and incorporation into DNA, preventing replication. This, and the acute toxicities associated with both 6MP and 6TG, makes them feasible as leads for future PLpro inhibitors, but not as SARS-CoV drugs themselves (Fig. 8b) (Chen et al., 2009; Chou et al., 2008).
4.3. Natural product inhibitors

A variety of natural products have been found to inhibit SARS-CoV PLpro activity including tanshinones derived from Salvia miltiorrhiza (Park et al., 2012b), diarylheptanoids derived from Alnus japonica (Park et al., 2012a), and geranylated flavonoids isolated from the fruits of the Paulownia tomentosa tree (Cho et al., 2013).

4.3.1. Tanshinones

Park and coworkers assayed seven different tanshinones for SARS-CoV PLpro inhibition using a continuous fluorometric assay (Park et al., 2012b). The tanshinones were found to inhibit SARS-CoV PLpro activity with IC50 values ranging from 0.8 to 30.0 μM. Interestingly, the authors found SARS-CoV PLpro inhibition increased as the pre-incubation time of PLpro with tanshinone increased, suggesting a slow-binding mechanism where inhibition of PLpro occurs through the rapid formation of an enzyme–inhibitor complex that slowly isomerizes to form a modified enzyme complex and possible covalent inhibition. The most potent tanshinone inhibitor of SARS-CoV PLpro enzymatic activity is shown in Fig. 8c, left panel (Park et al., 2012b).

4.3.2. Diarylheptanoids

The diarylheptanoid inhibitors were identified via activity-guided fractionation of the ethanolic extracts from the air-dried bark of A. japonica and assayed for inhibition of SARS-CoV PLpro using a continuous fluorometric assay (Park et al., 2012a). The most active diarylheptanoid inhibitor was found to contain an α,β-unsaturated carbonyl group – suggesting that inhibition of SARS-CoV PLpro may occur through formation of a covalent bond with the active site cysteine (vide supra). The most potent diarylheptanoid inhibitor of SARS-CoV PLpro enzymatic activity is shown in Fig. 8c, middle panel (Park et al., 2012a).

4.3.3. Geranylated flavonoids

Using activity-guided fractionation of the methanolic extracts of P. tomentosa fruits, Cho and coworkers identified five new geranylated flavonoid derivatives and seven known flavonoids that exhibit micromolar inhibition of SARS-CoV PLpro (Cho et al., 2013). Using a fluorogenic assay, the 12 flavonoids were found to inhibit SARS-CoV PLpro with IC50 values between 5.0 and 14.4 μM. The most active inhibitor was found to be a novel geranylated flavonoid. In general, it was shown that compounds containing the dihydro-2H-pyran moiety displayed better PLpro inhibition than their parent compounds. The most potent geranylated flavonoid inhibitor of SARS-CoV PLpro enzymatic activity is shown in Fig. 8c, right panel (Cho et al., 2013).

4.4. Zinc Ion (Zn2+) and zinc conjugate inhibitors

Using a fluorogenic inhibitor-screening platform that utilized nanomolar concentrations of SARS-CoV PLpro, Han and coworkers established that zinc ion (Zn2+) was capable of inhibiting PLpro protease activity with an IC50 value of 1.3 μM (Han et al., 2005). Additionally, they found the zinc conjugates N-ethyl-N-phenylthiocarbamic acid–Zn(II) and hydroxyypyridine-2-thione–Zn(II)
to inhibit PLpro protease activity with IC$_{50}$ values of 3.3 and 3.7 µM, respectively. In this study, inhibition of PLpro activity by Zn(II) appeared to be specific as the other divalent metals tested (Mg(II), Mn(II), Ca(II), Ni(II), and Co(II)) showed no inhibitory effect on the activity of SARS-CoV PLpro when tested at a concentration of 10 µM; though Cu(II) was found to inhibit SARS-CoV PLpro protease activity to 70% at a concentration of 10 µM (Lindner et al., 2005).

4.5. Naphthalene inhibitors

4.5.1. Compound 24

The class of naphthalene inhibitors was first reported in 2008, and was discovered by our group through the high-throughput screen (HTS) of 50,080 compounds for inhibitors of SARS-CoV PLpro (Ratia et al., 2008). Two promising SARS-CoV PLpro inhibitors were identified, 7724772 and 6577871.
having IC\textsubscript{50} values of 20.1 \( \mu \)M and 59 \( \mu \)M, respectively (Fig. 9a1 and a2). Interestingly, both of these hits from the primary HTS contained a naphthylmethylamine moiety. The more potent naphthyl inhibitor identified, compound 7724772, was identified as a racemate, and upon further investigation the (R)-enantiomer was found to be a more potent inhibitor, having IC\textsubscript{50} value of 8.7 ± 0.7 \( \mu \)M (Fig. 9a1). Analogs of (R)-7724772 were synthesized to explore the effects of adding different substituents at the ortho- and meta-positions of the benzene ring and the effects of changing the linkage to either the 1- or 2-naphthyl position. Preliminary structure–activity relationship (SAR) analysis indicated that substitutions at the ortho-benzene position other than a methyl group were unfavorable and decreased binding affinity, while changing the linkage to the 1-naphthyl position increased inhibitor potency by 4-fold. Addition of a nitro-substituent at the meta-position to the amide on the benzene ring decreased inhibitory potency, while addition of an amine at that same position increased inhibition of PL\textsubscript{pro} by 4-fold (Fig. 9b1). These SAR studies guided the construction of inhibitor 24, which has a significantly increased inhibitory potency against SARS-CoV PL\textsubscript{pro} (IC\textsubscript{50} = 0.6 ± 0.1 \( \mu \)M vs. 20.1 ± 1.1 \( \mu \)M). Compound 24 was also found to inhibit SARS-CoV viral replication in Vero cells with an EC\textsubscript{50} of 14.5 ± 0.8 \( \mu \)M (Fig. 9c1) (Ghosh et al., 2009; Ratia et al., 2008). More in depth kinetic and biochemical studies demonstrated 24 to be a potent, noncovalent, competitive inhibitor of SARS-CoV PL\textsubscript{pro} (\( K_i \) = 0.49 ± 0.08 \( \mu \)M).

The X-ray structure of 24 in complex with SARS-CoV PL\textsubscript{pro} was determined to a resolution of 2.5 Å, where 24 was found to bind to SARS-CoV PL\textsubscript{pro} within the S3 and S4 subsites in a cleft leading to the active site (Ratia et al., 2008). The amide group of 24 forms two hydrogen bonds with the Asp165 and Gln270 residues of SARS-CoV PL\textsubscript{pro}, while the rest of the interactions between the inhibitor and SARS-CoV PL\textsubscript{pro} were hydrophobic. The naphthyl group of the inhibitor forms hydrophobic interactions with the aromatic rings of the Tyr265 and Tyr269 residues and with the side chains of Pro248 and Pro249. The (R)-methyl group of 24 is positioned directly into a polar cavity, between residues Tyr265 and Thr302. The ortho-methyl substituent on the benzene ring is directed toward the floor of this cavity, and surrounded by residues Tyr265, Tyr274, and Leu163. The amino-substituent of the benzene ring is surrounded by several polar groups that may serve as hydrogen-bond acceptors (Fig. 9d1).

4.5.2. Compound 15g

The second, less potent naphthylmethylamine hit identified from our primary HTS campaign (6577871, Fig. 9a2) was also found to be a competitive, noncovalent inhibitor of SARS-CoV PL\textsubscript{pro} (Ratia et al., 2008). Interestingly, 6577871 had a 1-substituted naphthalene ring, which was found to be the more active linkage in the 7724772 series (vide supra). A small library of 6577871 analogs was synthesized to investigate the effect of substitution at the 1- or 2-naphthyl positions, at the benzylic-naphthyl position, and at positions on the benzene ring (Ghosh et al., 2010). Similar to the SAR trends found for 7724772, a 1-naphthyl linkage was found to be most favorable for increased PL\textsubscript{pro} inhibition. Again, substitution of a methyl group at the benzylic-naphthyl position was found to increase inhibition; however, in the case of the 6577871 series, the stereochemistry of this chiral center had a much less dramatic effect, where the preference for an (R)- or (S)-stereochemistry changed depending upon the substituents on the benzene ring. A 2-methoxy substituent on the benzene ring was found to have a 3.5-fold decrease in inhibition, while...
positioning the methoxy substituent at the 3- or 4-positions on the benzene ring did not measurably change the IC_{50} values. A series of benzodioxole analogs was also investigated and found to have minimally increased inhibitory potency relative to the methoxy-substituted benzene analogs (Fig. 9b2). These studies and SAR analyses ultimately led to the generation of inhibitor 15g, which has a significantly increased inhibitory potency against SARS-CoV PLpro relative to 6577871 (IC_{50} = 0.32 ± 0.01 μM vs. 59 ± 7.8 μM). Compound 15g was found to inhibit SARS-CoV viral replication in Vero cells with an EC_{50} = 9.1 ± 0.5 μM. Interestingly, the (S)-enantiomer of 15g, compound 15h, was found to have only slightly decreased inhibition of SARS-CoV PLpro, having an IC_{50} value of 0.56 ± 0.03 μM. Moreover, 15g showed better inhibition of SARS-CoV PLpro and better inhibition of SARS-CoV replication than the previously identified compound, 24 (Fig. 9c2). Additional kinetic and biochemical studies demonstrated that 15g is a potent, noncovalent, competitive inhibitor of PLpro (Ghosh et al., 2010).

The X-ray structure of 15g in complex with SARS-CoV PLpro was determined to a resolution of 2.63 Å and 15g was found to bind in the same cleft leading to the active site as 24 (Fig. 9d1 and d2). Both compounds, 15g and 24, bind with their naphthymethyleamine moieties positioned in the same orientation and location within the enzyme active site. The naphthyl moieties sit in a hydrophobic pocket surrounded by PLpro residues Tyr269, Tyr265, Pro248, Pro249, and Thr302. The (R)-configured methyl group of 15g extends into a small pocket that is lined with both hydrophobic and polar groups that are bound to water molecules. The piperidine nitrogen of 15g makes a hydrogen bond with the backbone carbonyl oxygen of Tyr269, further anchoring the inhibitor in the binding site. The presence of these additional interactions within the PLpro binding cleft is likely responsible for the increased inhibitory ability of 15g relative to 24 (Fig. 9d2).

4.6. Second generation naphthalene inhibitors

4.6.1. Compounds 3k and 3j

Though compounds 24 and 15g had low micromolar inhibitory potency and minimal cytotoxicity in SARS-CoV infected Vero E6 cells, further optimization of their antiviral potencies and physicochemical properties was still required to make them therapeutically viable. Using the SARs previously described, and the structural information obtained from the X-ray crystal structures of 24 and 15g, a second generation of 30 inhibitors was synthesized and evaluated for SARS-CoV PLpro inhibition. This second generation library was designed to probe the steric and electronic demands of the benzylic-naphthyl position, benzylic position, substituent and bioisoster tolerance of the benzyl and naphthyl rings, and the effect of distance between the aromatic rings. The common features of 15g and 24 required for good PLpro were held constant, namely the 1-naphthyl substitution and position of the hydrogen-donating nitrogen (Fig. 10a).

Increasing the size of the benzylic-naphthyl substituent from a methyl group to anything with greater steric demands decreased SARS-CoV PLpro inhibition in a trend proportional to substituent size, indicating that the pocket accessed by the (R)-methyl group was less accessible than initially indicated by the crystal structures of 24 and 15g (Fig. 10b, position R1). Adding a variety of substituents to the benzylic position did not increase inhibitory potency beyond that of 15g, indicating that additional functionalization at this position does not further engage the residues of the PLpro binding cavity (Fig. 10b, position R2). The effect of substitution on the benzene ring was found to be completely substituent and position dependent, with no clear SAR trend; however two analogs with meta- and para-fluoro substituents were found to have the greatest PLpro inhibition, 3k and 3j (Fig. 10b, position R3). The effect of extending the linkage between the amide and benzene ring was examined and lengthening the chain by even one carbon atom was found to weaken PLpro inhibition. A variety of bioisosteres were also investigated where the W, X, Y, and Z positions (Fig. 10b) were substituted with nitrogen. In general, these bioisosteres did not increase the inhibitory potency of the compounds, with the exception of 5c, which had a similar IC_{50} and EC_{50} to 15g (Fig. 10b and c).

The X-ray structures of SARS-CoV PLpro in complex with lead inhibitors 3k and 3j were determined to resolutions of 2.1 Å and 2.5 Å, respectively (Baez-Santos et al., 2014a). Similar to the X-ray structures of SARS-CoV PLpro in complex with lead inhibitors 24 and 15g, 3k and 3j were found to bind adjacent to the active site, within the S3 and S4 sub-sites of the SARS-CoV PLpro enzyme and in the same orientation. The inhibitors bind to SARS-CoV PLpro through an induced-fit mechanism, where the BL2 loop, which contains Tyr269, adopts a closed conformation in order to interact strongly with the inhibitors. The interactions between the BL2 loop

Fig. 10. Discovery and design pipeline of compounds 3k, 3j, 3e, and 5c. (a) Initial SARS-CoV PLpro inhibitor leads, 15g and 24. (b) SARs derived from small 15g analog library, (c) top SARS-CoV PLpro inhibitor candidates, 3k, 3j, 3e, and 5c (Baez-Santos et al., 2014a).
and the inhibitor occur through different bonds including a hydrogen-bond between the backbone carbonyl of Tyr269 and the amide nitrogen of the inhibitors. The motions associated with the BL2 loop and its different conformations are further discussed below in Section 5.

The naphthyl-rings of 3k and 3j are located in identical chemical space to the inhibitors 15g and 24 within the hydrophilic pocket hedged by Pro248 and Pro249 and lined by Tyr265, Tyr269, and Thr302 (Fig. 11a). The higher resolution X-ray structures of 3k and 3j allowed for a more definitive placement of three conserved water molecules present in the small cavity occupied by the (K)-methyl substituent. These water molecules decrease the effective size of the cavity and support our previously described SAR for substituents at the benzyl-naphthyl position. The superposition of compounds 3k, 3j, and 15g shows a 1.0 A difference in the position of the amide nitrogen of 15g compared to 3k and 3j, allowing a slight tilt in the orientation of the benzene ring to accommodate the different substituents (Fig. 11b). This tilt allows 3k to form stronger interactions with Tyr269 and Gln270 and is postulated to explain the increased inhibitory potency of 3k.

4.6.2. Metabolically stable naphthalene-based SARS-CoV Plpro inhibitors

The most potent identified inhibitor of Plpro, compound 3k, was expected to have greater metabolic stability than the previously identified lead compound, 15g, as the metabolically labile dioxolane of 15g had been replaced with a fluoro-substituent (Baez-Santos et al., 2014a). Additionally, as 3k has a lower topological polar surface area relative to 15g, it was expected to have increased cell permeability and subsequently, greater antiviral activity in comparison to 15g. The most potent analogs of 15g were subjected to anti-SARS-CoV assays in cell culture to determine EC_{50} values, which resulted in the identification of four top compounds with good antiviral activity: 3k, 3j, 3e, and 5c (Baez-Santos et al., 2014a). To assess the potential stability of these compounds in vivo, their stability to Phase I metabolism by mouse liver microsomes was evaluated. Surprisingly, 3k was found to be rapidly metabolized; however, the compounds 3e and 5c showed significantly improved metabolic stability relative to the best first generation inhibitor, 15g (Baez-Santos et al., 2014a). Because of their improved metabolic stability, inhibitors 3e and 5c are promising candidates for advancing to animal efficacy models.

4.6.3. Selectivity of naphthalene-based SARS-CoV Plpro inhibitors

There are at least 98 functional deubiquitinating enzymes in the human genome which fall into one of 6 different classes, where the number of members within each class are indicated in parenthesis; USP (56), OTU (15), JAMM (12), MCPIP (7), MJD (4) and UCH (4) (Fraile et al., 2012). The USP class, to which coronaviral Plpro’s belong, represent the largest class with 56 members. As the catalytic domains among the USP class are highly conserved, the question immediately arises as to how selective the designed inhibitors are for SARS-CoV Plpro and against human USPs? At least two approaches have been used so far in an attempt to answer this question. The first was a chemical biology approach pioneered by Hidde Ploegh and colleagues (Borodovsky et al., 2001, 2002; Love et al., 2007). In this approach, the active-site-directed probe HA–Ub–vinyl sulfone (VS) was used to treat cellular lysates of untransfected cells and cells transfected with SARS-CoV Plpro both in the absence and presence of inhibitors (Ratia et al., 2008). Since HA–Ub–VS potently inhibits a number of human USPs, the ability of compounds to block modification of USPs or Plpro from cellular lysates can be determined via Western blot analysis by using an anti-HA antibody. Using this approach, we found that the naphthalene-based class of inhibitors do not compete against HA–Ub–VS modification of cellular DUBs but does compete against modification of SARS-CoV Plpro.

A second approach for testing selectivity against USPs, other DUBs and cysteine proteases is to test for compound inhibition against purified enzymes, as many of these enzymes are now commercially available. We used this approach to test the naphthalene-based inhibitors against purified USP2, USP7, USP8, USP18, USP20, USP21, USP28, UCH-L1, UCH-L3, DEN1, caspase 3 and cathepsin K in biochemical-based assays (Baez-Santos et al., 2014a; Ratia et al., 2008). We found that none of these human enzymes was significantly inhibited by the naphthalene-based Plpro inhibitors up to concentrations as high as 100 µM. Importantly, the selective inhibition of SARS-CoV Plpro by designed, non-covalent inhibitors is the first example demonstrating that it is possible to design selective inhibitors against the USP class of enzymes which is of universal importance for selectively targeting DUBs involved in promoting different diseases.

5. X-ray crystal structures of SARS-CoV Plpro

5.1. Structure of SARS-CoV Plpro in an unbound state

The X-ray structure of the 35-kDa catalytic domain of Plpro (residues 1541–1855 of the SARS-CoV nsp3 polyprotein domain) was originally determined to a resolution of 1.85 Å in its unbound state, providing the first structure of a coronaviral Plpro (Ratia et al., 2006). Although gel filtration analyses indicate that Plpro is predominately monomeric in solution (Han et al., 2007), the active-site-directed probe Hidde Ploegh and colleagues (Borodovsky et al., 2001, 2002; Love et al., 2007). In this approach, the active-site-directed probe HA–Ub–vinyl sulfone (VS) was used to treat cellular lysates of untransfected cells and cells transfected with SARS-CoV Plpro both in the absence and presence of inhibitors (Ratia et al., 2008). Since HA–Ub–VS potently inhibits a number of human USPs, the ability of compounds to block modification of USPs or Plpro from cellular lysates can be determined via Western blot analysis by using an anti-HA antibody. Using this approach, we found that the naphthalene-based class of inhibitors do not compete against HA–Ub–VS modification of cellular DUBs but does compete against modification of SARS-CoV Plpro.

A second approach for testing selectivity against USPs, other DUBs and cysteine proteases is to test for compound inhibition against purified enzymes, as many of these enzymes are now commercially available. We used this approach to test the naphthalene-based inhibitors against purified USP2, USP7, USP8, USP18, USP20, USP21, USP28, UCH-L1, UCH-L3, DEN1, caspase 3 and cathepsin K in biochemical-based assays (Baez-Santos et al., 2014a; Ratia et al., 2008). We found that none of these human enzymes was significantly inhibited by the naphthalene-based Plpro inhibitors up to concentrations as high as 100 µM. Importantly, the selective inhibition of SARS-CoV Plpro by designed, non-covalent inhibitors is the first example demonstrating that it is possible to design selective inhibitors against the USP class of enzymes which is of universal importance for selectively targeting DUBs involved in promoting different diseases.

5.1. Structure of SARS-CoV Plpro in an unbound state

The X-ray structure of the 35-kDa catalytic domain of Plpro (residues 1541–1855 of the SARS-CoV nsp3 polyprotein domain) was originally determined to a resolution of 1.85 Å in its unbound state, providing the first structure of a coronaviral Plpro (Ratia et al., 2006). Although gel filtration analyses indicate that Plpro is predominately monomeric in solution (Han et al., 2007), the unliganded structure of Plpro crystallized in space group C2 as three

Fig. 11. X-ray crystal structures of SARS-CoV Plpro in complex with inhibitors 3k and 3j and comparison to 15g. (a) An overlay of SARS-CoV Plpro (blue cartoon, gray surface) in complex with 3k (orange stick, PDB: 40W0) and Plpro (cyan cartoon, gray surface) in complex with 3j (pink sticks, PDB: 4OVZ) depicting amino acids important for inhibitor binding (Baez-Santos et al., 2014a). (b) An overlay of Plpro–3k (Plpro in blue cartoon, 3k in orange sticks), Plpro–3j (Plpro in cyan cartoon, 3j in pink sticks) and Plpro–15g (green cartoon and sticks, (PDB: 3MJ5, Ghosh et al., 2010)) complexes.
Fig. 12. X-ray crystal structures of SARS-CoV PLpro in different asymmetric units. PLpro is shown as a ribbon representation with each monomer within an asymmetric unit colored in a different shade of blue. The BL2 loop is highlighted in red and the zinc ion is shown as a pink sphere. (a) The crystallographic trimer of unliganded PLpro (PDB: 2FE8, (Ratia et al., 2006)) is shown in the left panel and a structural superposition of the three monomers is shown in the right panel. (b) The crystallographic dimer of the SARS-CoV PLpro–15g inhibitor complex (PDB: 3MJ5, (Ghosh et al., 2010)) is shown in the left panel and a structural superposition of the two monomers is shown in the right panel. Compound 15g is shown in orange. (c) The crystallographic monomer of the SARS-CoV PLpro–24 inhibitor complex (PDB: 3E9S, (Ratia et al., 2008)) is shown in the left panel. Compound 24 is shown in green. The right panel depicts the structural superposition of the SARS-CoV PLpro–15g and SARS-CoV PLpro–24 inhibitor complexes. (D) The crystallographic dimer of the SARS-CoV PLpro–3k inhibitor complex (PDB: 4OW0, (Baez-Santos et al., 2014a)) is shown in the left panel and a structural superposition of the two monomers in the asymmetric unit is shown in the right panel.
closely packed monomers in the crystallographic asymmetric unit (Fig. 12a, left panel). Structural alignment of the three monomers shows no significant conformational differences in the central body (palm and thumb domains) of the enzyme or within the active site. However, significant conformational differences between all three monomers were detected at the fingers domain and the BL2 loop of the palm domain at the entrance of the active site, suggesting high plasticity at these regions (Fig. 12a, right panel).

5.2. Structure of SARS-CoV PLpro in complex with inhibitors

Generally, SARS-CoV PLpro has been highly challenging to crystallize. The crystallization conditions that yielded the SARS-CoV PLpro crystals in its unliganded state were often irreproducible or did not produce diffraction quality crystals, preventing the crystallization of SARS-CoV PLpro in complex with inhibitors via soaking experiments. Attempts to co-crystallize SARS-CoV PLpro with inhibitors using the crystallization conditions identified for the unliganded SARS-CoV PLpro enzyme were also unsuccessful, presumably due to the tight packing interactions of SARS-CoV PLpro monomers and the positioning of the active site on the interior of the crystallographic trimer (Fig. 12a). Eventually, co-crystals of SARS-CoV PLpro–inhibitor complexes were achieved by the screening of thousands of crystallization conditions using a crystallization robot. Ultimately, the SARS-CoV PLpro–inhibitor complexes were found to crystallize in different crystallization conditions (Table 1). The co-crystal of SARS-CoV PLpro–15g complex diffracted to a slightly lower resolution of 2.63 Å, and similarly for the unliganded SARS-CoV PLpro enzyme were also unsuccessful, presumably due to the tight packing interactions of SARS-CoV PLpro monomers and the positioning of the active site on the interior of the crystallographic trimer (Fig. 12a). Eventually, co-crystals of SARS-CoV PLpro–inhibitor complexes were achieved by the screening of thousands of crystallization conditions using a crystallization robot. Ultimately, the SARS-CoV PLpro–inhibitor complexes were found to crystallize in different crystallization conditions (Table 1). The co-crystal of SARS-CoV PLpro–15g complex diffracted to a slightly lower resolution of 2.63 Å, and similarly for the unliganded SARS-CoV PLpro enzyme, it belongs to the space group C2 but with two monomers per asymmetric unit (Ghosh et al., 2010) (Fig. 12b). Crystals of the SARS-CoV PLpro–inhibitor 24 complex (also known as compound GRL0617) grew in space group I222 with one monomer per asymmetric unit and diffracted to a resolution of 2.5 Å (Ghosh et al., 2009; Ratia et al., 2008) (Fig. 12c). Unfortunately, as observed before for the unliganded SARS-CoV PLpro crystals, the crystallization conditions that yielded the inhibitor-bound SARS-CoV PLpro crystals were irreproducible and not suitable for the crystallization of SARS-CoV PLpro in complex with newly developed inhibitors.

As more SARS-CoV PLpro inhibitors were developed, the need for less costly and labor-intensive crystallization methods increased. Therefore, new crystallization efforts were centered on finding conditions for the rapid and reliable co-crystallization of SARS-CoV PLpro–inhibitor complexes. These efforts included developing a new approach for enzyme preparation and crystallization screening and ultimately resulted in the identification of a new crystallization condition, which after optimization, yielded crystals with improved reproducibility and increased diffraction ability. The structure of SARS-CoV PLpro in complex with compounds 3k and 3j resulted from these new crystallization conditions (Baez-Santos et al., 2014a) (Fig. 12d). As observed for the SARS-CoV PLpro–15g complex, the co-crystals of PLpro–3k/3j complexes belong to space group C2 with two molecules per asymmetric unit. Although poorer electron density is observed for the UBL domain, the SARS-CoV PLpro–3k/3j crystals diffracted to improved resolutions of 2.1 Å and 2.5 Å, respectively, and provided the basis for the crystallization and structure determination of SARS-CoV PLpro in complex with inhibitors from the same series of compounds. A summary of all SARS-CoV PLpro crystal structures and crystallization conditions reported to date is provided in Table 1.

The multiple X-ray crystal structures of SARS-CoV PLpro have consistently shown multiple conformations at the fingers domain and the BL2 loop (Fig. 12). The BL2 loop, which is observed in an open conformation in the unliganded SARS-CoV PLpro structure, closes upon inhibitor binding toward the catalytic cleft allowing the formation of intermolecular interactions with the inhibitors. The more closed conformation is observed with the smallest compounds 24, in which a flip of the peptide bond between Gln270 and Tyr269 allows the formation of a hydrogen bond between the backbone amide of Gln270 and the carbonyl at the center of the inhibitor (Ratia et al., 2008) (Fig. 12c, right panel). A more open conformation is observed upon binding of the larger compounds 15g, 3k and 3j, in which the BL2 loop retains the backbone orientation previously observed in the unliganded PLpro structure (Fig. 12c, right panel). While the closing motions of the BL2 loop are driven by the induced-fit mechanism of inhibitor binding, other inhibitor-independent motions are observed at the tip of the fingers domain, which may account for PLpro’s ability to interact with topologically different poly-Ub chains and the UBL modifier ISG15 (see Section 2, Fig. 2c).

In contrast to the motions observed outside the catalytic site, the catalytic triad undergoes limited movement. However, the
active site cysteine Cys112 is highly reactive, as it has been observed in a fully oxidized state in the inhibitor bound structures (Fig. 12c, right panel). It remains unknown whether oxidation of Cys112 is enhanced by inhibitor-induced closure of the BL2 loop, which may restrict access of the active site to reducing agents but still allow molecular oxygen to enter the active site, or whether the inhibitor-induced conformation increases the reactivity of the cysteine to molecular oxygen. Nevertheless, the observed oxidation of SARS-CoV PLpro Cys112 in various crystal structures demonstrates the high reactivity of the SARS-CoV PLpro active site.

5.3. Structure of SARS-CoV PLpro in complex with ubiquitin

Recently, crystal structures of SARS-CoV PLpro in complex with Ub were determined to resolutions of 1.4 Å (Chou et al., 2014) and 2.75 Å (Ratia et al., 2014) (Fig. 13a). Two different approaches were used to stabilize the SARS-CoV PLpro ubiquitin complex for cocrystallization. One approach used an active site cysteine mutant, C112S, to form a noncovalent complex with ubiquitin. The C112S–PLpro–Ub complex crystallized in the monoclinic space group P2₁ with one PLpro–Ub complex per asymmetric unit (Chou et al., 2014). The wild type SARS-CoV PLpro–ubiquitin complex is very weak as the Kᵢ value of free ubiquitin cannot be determined from competitive inhibition studies (Baez-Santos et al., 2014b; Ratia et al., 2014). The introduction of the serine side chain group of serine and the carboxy terminus of ubiquitin. In the mechanism illustrated in Fig. 5, the C112S complex would represent the product (”E+Q”) complex.

Our lab utilized a different approach to capture a transient, covalent intermediate complex (”FQ”) that forms during the SARS-CoV PLpro catalyzed reaction cycle (Fig. 5). Ubiquitin-aldehyde (Ubal) was synthesized, using methodologies developed by Wilkinson et al. (2005), to form a stable, wild type SARS-CoV PLpro–Ubal complex per asymmetric unit (Chou et al., 2014). Ubiquitin-aldehyde can form a covalent but reversible complex that can increase the binding affinity of a protein for ubiquitin by as much as 300,000-times relative to unmodified ubiquitin (Melandi et al., 1996). The SARS-CoV PLpro–Ubal complex crystallized in the trigonal space group P3₂₁ with one PLpro–Ubal complex per asymmetric unit. The crystals diffracted to well beyond 2.7 Å but the 342 Å unit cell dimension in the c-dimension and the synchrotron setup prohibited data collection beyond this resolution. Nonetheless, the X-ray structure of the SARS-CoV PLpro–Ubal complex clearly revealed the covalent bond between the cysteine 112 sulfur atom and the c-terminal carbonyl carbon of ubiquitin (Figs. 5 and 7).

A structural superposition of the wild type PLpro–Ubal and C112S PLpro–Ubal complexes is shown in Fig. 13 and both structures reveal that the binding of ubiquitin induces a significant shift in the zinc-binding motif toward the Ub molecule and alters the conformation of the BL2 loop so that it interacts with the C-terminal residues (residues 72–76) of Ub (Fig. 13b). The BL2 loop contributes the majority of PLpro–Ub intermolecular hydrogen bonds and plays an important role in aligning and orienting the C-terminus of ubiquitin toward the active site of PLpro. There are observable conformational differences of the BL2 loop between the two crystal structures of PLpro–Ub and C112S–PLpro–Ub. For example, different orientations are observed for the side chain of Tyr269 (Fig. 13b). In addition, Arg72, Arg74 and Gly76 of ubiquitin are found in different positions (Fig. 13b and c). In the C112S PLpro–Ub complex, Gly76 is found within hydrogen bond distance of Cys112 and hydrogen bonded to Tyr107 and His273, corresponding to the intermediate (”F”) state (see also Figs. 5 and 7). The subtle shift in the positioning of the Ub C-terminal Gly76 results from the formation of a covalent (1.6 Å) versus a hydrogen bond (2.5 Å) with PLpro, and perhaps, this shift translates to more prominent differences in the orientations of the preceding C-terminal residues Arg72 and Arg74 of Ub. Thus, these two PLpro–Ub complex crystal structures may provide a snapshot of the transition of PLpro from the ”FQ” to the ”E+Q” states revealing once again the high plasticity of the BL2 loop and its interactions with substrates, intermediates, products and inhibitors.

6. Design of antiviral inhibitors targeting PLpro from other coronaviruses

The success of designing inhibitors against SARS-CoV PLpro with antiviral activity and selectivity suggests that it will be possible to design inhibitors with similar properties against other coronaviruses. In addition, since the naphthalene-based inhibitors are potent, competitive inhibitors and bind within the active site of SARS-CoV PLpro, these scaffolds would appear to be excellent starting points. Toward this goal, we tested over 30 of the naphthalene-based inhibitors (Figs. 9 and 10) for their ability to inhibit the papain-like protease two (PLP2) from the human coronavirus (HCoV) NL63 and PLpro from the human Middle East Respiratory Syndrome coronavirus (MERS-CoV). We found that 6 compounds are able to inhibit the HCoV-NL63 PLP2 enzyme with potencies ranging from 18 to 60 μM (Baez-Santos et al., 2014a). However,
none of the compounds inhibited PLpro from MERS-CoV (Baez-Santos et al., 2014b).

The lack of inhibition of the naphthalene-based inhibitors toward MERS-CoV PLpro and the lower inhibitory potency toward HCoV-NL63 PLP2 may be explained in part by the structural differences within the BL2 loop of these enzymes. In Fig. 14 we show the X-ray structure of SARS-CoV PLpro in complex with inhibitor 3k (Baez-Santos et al., 2014a), which has been superimposed onto the homology models of HCoV-NL63 PLP2 and MERS-CoV PLpro (Baez-Santos et al., 2014b). Although the X-ray structure of unliganded MERS-CoV PLpro was recently determined by Rolf Hilgenfeld’s group (Lei et al., 2014), a number of amino acids in the β-turn/loop were not modeled in because of missing electron density. However, our energy-minimized homology models of MERS-CoV PLpro and HCoV-NL63 PLP2 included the BL2 β-turn/loop and a comparison of our MERS-CoV PLpro homology model and the X-ray crystal structure show that the overall structures were very similar (Baez-Santos et al., 2014b).

The structural models reveal that Y269 in SARS-CoV PLpro, an essential residue for binding of the naphthalene-based inhibitors, is absent in MERS-CoV PLpro resulting in the removal of a T-shaped π–π interaction with the naphthalene group of the inhibitor as well as other van der Waals interactions (Fig. 14). In addition, the BL2 β-turn/loop of MERS-CoV PLpro is one residue longer which may alter the hydrogen-bonding interactions between the inhibitor and backbone atoms of Y269. In contrast, the length of the BL2 β-turn/loop is identical for HCoV-NL63 PLP2 and F271 is a conservative substitution for Y269 in SARS-CoV PLpro. The benzylic group of F271 is still fully capable of forming a T-shaped π–π interaction with the naphthalene group and the van der Waals interactions. However, subtle structural shifts surrounding the binding pocket are likely to cause the decreased inhibitory potency of the naphthalene-based series for HCoV-NL63 PLP2. The sequence and structural variability of the BL2 β-turn/loop among coronavirus suggests that design of a broad-spectrum coronavirus antiviral compound based on the naphthalene scaffold will be difficult, but a more narrow spectrum antiviral may be more readily achievable.

7. Conclusions

Over the past 10 years, the papain-like protease domain of SARS-CoV nsp3 has emerged as a viable drug target for the development of anti-SARS therapeutics. The PLpro enzyme is essential for viral replication and small molecule inhibitors of this enzyme are capable of eliciting a SARS-CoV antiviral response in cell culture. The X-ray structures of SARS-CoV PLpro in complex with inhibitors provide a detailed map of the active site, which can be exploited for structure-based drug design (SBDD) of potent inhibitors. Analyses of the structural differences surrounding the active site BL2 loop indicate that the plasticity of this loop and the amino acid side chains within this loop will have a profound influence on the design of inhibitors against SARS-CoV PLpro and other coronavirus PLpro enzymes including the enzyme from the Middle East Respiratory Syndrome coronavirus. Finally, the SBDD approaches and studies that we have applied to SARS-CoV PLpro prove that it is entirely possible to develop highly-selective, non-covalent inhibitors against other USPs involved in disease processes.

Acknowledgements

The authors wish to express their gratitude to Dr. Kliara Ratia from the Department of Medicinal Chemistry and Pharmacognosy, University of Illinois, Chicago, Prof. Susan Baker from the Department of Microbiology & Immunology, Loyola University Chicago Stritch School of Medicine, Prof. Arun Ghosh from the Department of Chemistry, Purdue University, and Prof. Scott Larsen from the Department of Medicinal Chemistry, University of Michigan for their significant scientific contributions and collaboration in developing SARS-CoV PLpro inhibitors over the past decade. In addition, we would like to acknowledge financial support over the years for this project from the National Institute of Allergy and Infectious Diseases (NIAID) via R01AI085089 and P01AI060915. ADM is supported by a grant from the Walther Cancer Foundation. Support for continued crystallographic and computational modeling of this project is being provided in part by the Purdue Center for Cancer Research, which is supported by a grant from the NCI P30 CA023168. Finally, we would like to acknowledge the beamline support staff at LS-CAT, SER-CAT and LRL-CAT at the Advanced Photon Source, Argonne National Lab over the years for their assistance in our X-ray work, which provided the structures of SARS-CoV PLpro in complex with the antiviral inhibitors.

References

Angelini, M.M., Akhlaghpour, M., Neuman, B.W., Buchmeier, M.J., 2013. Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4 and 6 induce double-membrane vesicles. MBio 4.

Baez-Santos, Y.M., 2012. Insight into the Substrate Specificity and Inhibition of Human Coronavirus Papain-Like Proteases. Pharmacy. University of Illinois, Chicago, pp. 1–230.

Baez-Santos, Y.M., Barraza, S.J., Wilson, M.W., Agius, M.P., Mielech, A.M., Davis, N.M., Baker, S.C., Larsen, S.D., Mesecar, A.D., 2014a. SARS-CoV PLpro and HCoV-NL63 PLP2 included the BL2 β-turn/loop and a comparison of our MERS-CoV PLpro homology model and the X-ray crystal structure show that the overall structures were very similar (Baez-Santos et al., 2014b).

Barrett, A.J., Rawlings, N.D., Woessner, J.F., 2012. In: Handbook of Proteolytic Enzymes, . third ed., Catalytic Mechanisms of Cysteine Peptidases third ed., vol. 1 Elsevier Science (chapter 404).

Barretto, N., Jukneliene, D., Ratia, K., Chen, Z., Mesecar, A.D., Baker, S.C., 2005. The papain-like protease of severe acute respiratory syndrome coronavirus has deubiquitinating activity. J. Virol. 79, 15189–15198.

Barretto, N., Jukneliene, D., Ratia, K., Chen, Z., Mesecar, A.D., Baker, S.C., 2006. Deubiquitinating activity of the SARS-CoV papain-like protease. Adv. Exp. Med. Biol. 581, 37–41.

Borodovsky, A., Kessler, B.M., Casagrande, R., Overkleeft, H.S., Wilkinson, K.D., Ploegh, H.L., 2001. A novel active site-directed probe specific for deubiquitylating enzymes reveals proteasome association of USP14. EMBO J. 20, 5187–5196.
Ratia, K., Saikatendu, K.S., Santarisi, B.D., Barretto, N., Baker, S.C., Stevens, R.C., Mesecar, A.D., 2006. Severe acute respiratory syndrome coronavirus papain-like protease: structure of a viral deubiquitinating enzyme. Proc. Natl. Acad. Sci. U.S.A. 103, 5717–5722.

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. Structural basis of severe acute respiratory syndrome coronavirus ADP-ribose-1'-phosphate dephosphorylation by a conserved domain of nsP3. Structure (Camb) 13, 1665–1675.

Serrano, P., Johnson, M.A., Almeida, M.S., Horst, R., Herrmann, T., Joseph, J.S., Neuman, B.W., Subramanian, V., Saikatendu, K.S., Buchmeier, M.J., Stevens, R.C., Kuhn, P., Wuthrich, K., 2007. Nuclear magnetic resonance structure of the N-terminal domain of nonstructural protein 3 from the severe acute respiratory syndrome coronavirus. J. Virol. 81, 12998–13008.

Serrano, P., Johnson, M.A., Chatterjee, A., Neuman, B.W., Joseph, J.S., Buchmeier, M.J., Kuhn, P., Wuthrich, K., 2009. Nuclear magnetic resonance structure of the nucleic acid-binding domain of severe acute respiratory syndrome coronavirus nonstructural protein 3. J. Virol. 83, e1000428.

Snijder, E.J., Bredenbeek, P.J., Dobbe, J.C., Thiel, V., Ziebuhr, J., Poon, L.L., Hase, K., Mekalanos, J.J., Snijder, E.J., Ziebuhr, J., 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J. Mol. Biol. 331, 991–1004.

Sun, L., Xing, Y., Chen, X., Zheng, Y., Yang, Y., Nichols, D.B., Clementz, M.A., Banach, B.S., Li, K., Baker, S.C., Chen, Z., 2012. Coronavirus papain-like proteases negatively regulate antiviral innate immune response through disruption of STING-mediated signaling. PLoS One 7, e30802.

Tan, J., Kusov, Y., Mutschall, D., Tech, S., Nagarajan, K., Hilgenfeld, R., Schmidt, C.L., 2007. The “SARS-unique domain” (SUD) of SARS coronavirus is an oligo(G)-binding protein. Biochem. Biophys. Res. Commun. 364, 877–882.

Ziebuhr, J., 2004. Molecular biology of severe acute respiratory syndrome coronavirus. Curr. Opin. Microbiol. 7, 412–419.

Ziebuhr, J., Snijder, E.J., Gorbalenya, A.E., 2000. Virus-encoded proteinases and proteolytic processing in the Nidovirales. J. Gen. Virol. 81, 1853–1875.