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.

CHAPTER NINE

Viral proteases: Structure, mechanism and inhibition

Jacqueto Zephyr, Nese Kurt Yilmaz, and Celia A. Schiffer*
Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA, United States
*Corresponding author: e-mail address: celia.schiffer@umassmed.edu

Contents

Abstract

Viral proteases are diverse in structure, oligomeric state, catalytic mechanism, and substrate specificity. This chapter focuses on proteases from viruses that are relevant to human health: human immunodeficiency virus subtype 1 (HIV-1), hepatitis C (HCV), human T-cell leukemia virus type 1 (HTLV-1), flaviviruses, enteroviruses, and coronaviruses. The proteases of HIV-1 and HCV have been successfully targeted for therapeutics, with picomolar FDA-approved drugs currently used in the clinic. The proteases of HTLV-1 and the other virus families remain emerging therapeutic targets at different stages of the drug development process. This chapter provides an overview of the current knowledge on viral protease structure, mechanism, substrate recognition, and inhibition. Particular focus is placed on recent advances in understanding the molecular basis of diverse substrate recognition and resistance, which is essential toward designing novel protease inhibitors as antivirals.

The Enzymes, Volume 50
ISSN 1874-6047
https://doi.org/10.1016/bs.enz.2021.09.004

Copyright © 2021 Elsevier Inc.
All rights reserved.

1. Introduction

Many viruses encode one or more proteases as a common strategy to sustain replication with a compacted genome. The viral genome encodes a polyprotein with an embedded viral protease that cleaves the polyprotein at several specific sites to generate mature viral proteins. Viral proteases are therefore essential for replication, which makes them ideal therapeutic targets. The cleavage sites recognized by a given viral protease are generally diverse in amino acid sequence and processed at different rates, as polyprotein processing is an ordered and obligatory sequential process to produce infectious virus. This ability to specifically recognize and cleave diverse substrate sequences is conserved for proteases of different families of viruses. Although overall folds are shared between eukaryotic and viral proteases, diverse substrate recognition represents a unique feature of viral proteases. The unique substrate preference of viral proteases can be exploited in inhibitor design, toward developing selective and potent drug-like molecules. This has been achieved for HIV-1 and HCV where protease inhibitors have been part of antiviral combination therapies. However, drug resistance has emerged as a major problem and therefore needs to be considered as part of the drug-design process for evolving targets. The error-prone nature of replication allows viruses to rapidly select for resistance mutations under drug pressure. This chapter will focus on the viral protease structure, mechanism, substrate recognition, inhibition and resistance mutations for human immunodeficiency virus subtype 1 (HIV-1) and hepatitis C (HCV) virus where viral proteases have proven to be successful therapeutic targets. Furthermore, viral proteases from human T-cell leukemia virus type 1 (HTLV-1), flaviviruses, enteroviruses, and coronaviruses are discussed as emerging targets for drug development.

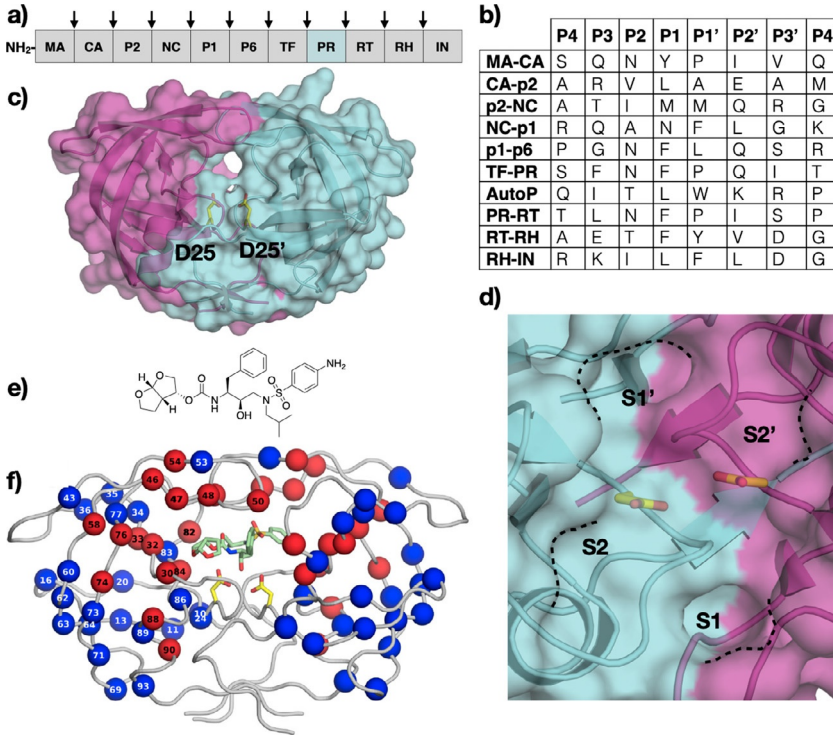
2. Viral proteases targeted with antiviral drugs

2.1 HIV-1 protease

HIV-1 protease is the most well-studied viral protease and was the first to be targeted for therapeutics. The success in targeting HIV-1 protease has paved the way for inhibiting other viral proteases in drug development. There are numerous reviews and book chapters highlighting the progress in HIV-1 protease research over the years [1–3]. Therefore, in this chapter, a general overview is provided with recent advances in understanding the molecular basis of drug-resistance mutations.

2.1.1 HIV-1 protease structure

HIV-1 protease is expressed as part of a larger Gag-Pro-Pol polyprotein (Fig. 1A). The initial cleavage occurs in *cis* to release the protease itself, and the current model suggests that the polyprotein precursor must dimerize



	P4	P3	P2	P1	P1'	P2'	P3'	P4'
MA-CA	S	Q	N	Y	P	I	V	Q
CA-p2	A	R	V	L	A	E	A	M
p2-NC	A	T	I	M	M	Q	R	G
NC-p1	R	Q	A	N	F	L	G	K
p1-p6	P	G	N	F	L	Q	S	R
TF-PR	S	F	N	F	P	Q	I	T
AutoP	Q	I	T	L	W	K	R	P
PR-RT	T	L	N	F	P	I	S	P
RT-RH	A	E	T	F	Y	V	D	G
RH-IN	R	K	I	L	F	L	D	G

Fig. 1 HIV-1 Protease. (A) Schematic representation of HIV-1 polyprotein (Gag-Pro-Pol), with viral protease (PR) cleavage sites between structural proteins (MA: matrix; CA: capsid; NC: nucleocapsid), enzymes (RT: reverse transcriptase; RH: RNase H; IN: integrase) and peptides (TF: *trans*-frame; spacer peptides P1, P2) indicated by arrows. (B) Amino acid sequence of polyprotein cleavage sites recognized by HIV-1 protease. The protease cleaves between P1-P1′. (C) The crystal structure of HIV-1 protease (PDB ID: 1T3R) with the two monomers colored cyan and magenta and the catalytic residue (D25) depicted in yellow sticks and labeled, and (D) a close-up view of the active site with S2-S2′ pockets annotated. (E) The chemical structure of the latest FDA-approved HIV-1 protease inhibitor, darunavir. (F) Cocrystal structure of darunavir (green sticks) bound to HIV-1 protease, with location of primary (red) and secondary (blue) resistance mutations indicated by spheres. *Panel F: Reprinted with permission from A.N. Matthew, F. Leidner, G.J. Lockbaum, M. Henes, J. Zephyr, S. Hou, et al., Drug design strategies to avoid resistance in direct-acting antivirals and beyond. Chem. Rev. 121(6) (2021) 3238–3270. Copyright (2021) American Chemical Society.*

to activate the protease. Additionally, the protease cleaves nine other sites on the viral polyprotein (Fig. 1B). HIV-1 protease is composed of two identical chains of 99-amino acids each and functions as a pseudo-symmetric homodimer (Fig. 1C). The overall structure of HIV-1 protease is related to that of monomeric cellular aspartic proteases. However, HIV-1 protease contains two distinct flaps that cover the active site formed at the dimer interface. These flaps must open up to allow substrates to bind and then close to stably bind substrates to perform catalysis.

2.1.2 Substrate recognition and cleavage

HIV-1 protease cleaves the Gag-Pro-Pol polyprotein at a total of ten sites that are of diverse amino acid sequence which are processed at different rates (Fig. 1B). The active site cavity covers the S2-S2' pockets with the S3/S3' being partially solvent-exposed. The hydrophobic S2-S2' pockets recognize mostly hydrophobic side chains, especially at P1/P1' positions (Fig. 1B). Although the sequences differ, cleavage sites share a conserved shape upon binding to the protease active site. The model explaining diverse substrate recognition is termed the substrate envelope, which represents the consensus volume occupied by the substrates. Hence, HIV-1 protease recognizes peptide substrates based not on the sequence but the shape the peptide can adopt [4–6]. HIV-1 protease uses two catalytic aspartic acids (D25 / D25'), one from each monomer, to activate a water molecule for catalysis. Each monomer contributes one aspartic acid as part of the general aspartyl protease Asp-Thr-Gly catalytic motif located at the base of the active site. The hydroxyl ion then attacks the scissile carbonyl carbon to generate a tetrahedral intermediate that collapses to generate new N- and C-termini. The tetrahedral transition state has been exploited in HIV-1 protease inhibitor design, following the transition state theory which states that enzymes bind tighter to the transition state intermediates compared to substrate or product.

2.1.3 HIV-1 protease inhibitors as drugs

There are a total of ten FDA-approved drugs targeting HIV-1 protease [1]. The initial peptidomimetic HIV-1 protease inhibitors (saquinavir, ritonavir, indinavir, nelfinavir, and amprenavir) resulted in high metabolic clearance, poor oral bioavailability, and had a low barrier to resistance. Second-generation inhibitors, aimed at improving on these limitations, led to the development of darunavir (DRV), the latest and most potent FDA-approved HIV-1 protease inhibitor (Fig. 1E). DRV exhibits activity against a broad range of clinical isolates that demonstrated resistance against the earlier generation drugs.

2.1.4 Resistance mutations

HIV-1 readily develops resistance under drug pressure because of the error-prone replication and reverse transcription of the viral RNA. Resistance mutations occur throughout the HIV-1 protease and are broadly classified into two categories. Primary resistance mutations that directly alter contact with the inhibitor occur in the active site and are largely explained by being outside the substrate envelope. In contrast, secondary resistance mutations are located mostly distal from the active site and impact the dynamic ensemble of the protease altering the balance of substrate turnover versus inhibitor binding (Fig. 1F). [7–19]. Recent studies combining parallel molecular dynamics simulations and machine learning have provided insights into how remote changes can confer resistance and which interactions are most indicative of these changes [7,20]. Each of the FDA-approved HIV-1 protease inhibitors has specific signature mutations that cause resistance; detailed reviews can be found elsewhere [3,19,21–25].

2.1.5 Substrate envelope and mechanisms of resistance

HIV-1 protease substrate envelope, which explains the molecular mechanism of diverse substrate recognition, is also a useful tool to explain the mechanism of drug resistance and thus can guide the design of more robust protease inhibitors [26]. The substrate envelope explains how resistance mutations lead to a decrease in inhibitor binding but do not impair substrate processing. Comparing cocrystal structures of substrate and inhibitor-bound structures revealed that inhibitors protrude outside of the substrate envelope at specific locations that correspond closely with the site of drug resistance mutations. Moreover, inhibitors that fit within the envelope have a higher barrier for drug resistance. The HIV-1 protease substrate envelope has been used to guide the design of inhibitors that are more robust against resistance [27–30].

The success in developing drugs against HIV-1 protease has demonstrated that viral proteases are viable therapeutic targets. Lessons learned from the drug design efforts targeting HIV-1 have been applied to other systems with varying levels of success, including the HCV NS3/4A protease.

2.2 HCV NS3/4A protease

The positive-sense single-stranded RNA genome of HCV encodes a viral polyprotein within which exists subsists the NS3/4A bifunctional protein (Fig. 2A). The 70 kDa NS3 protein contains two domains, a protease located at N-terminus and a helicase domain at the C-terminus [31]. Each enzymatic domain of NS3 can function independently of the other. The NS3 protease

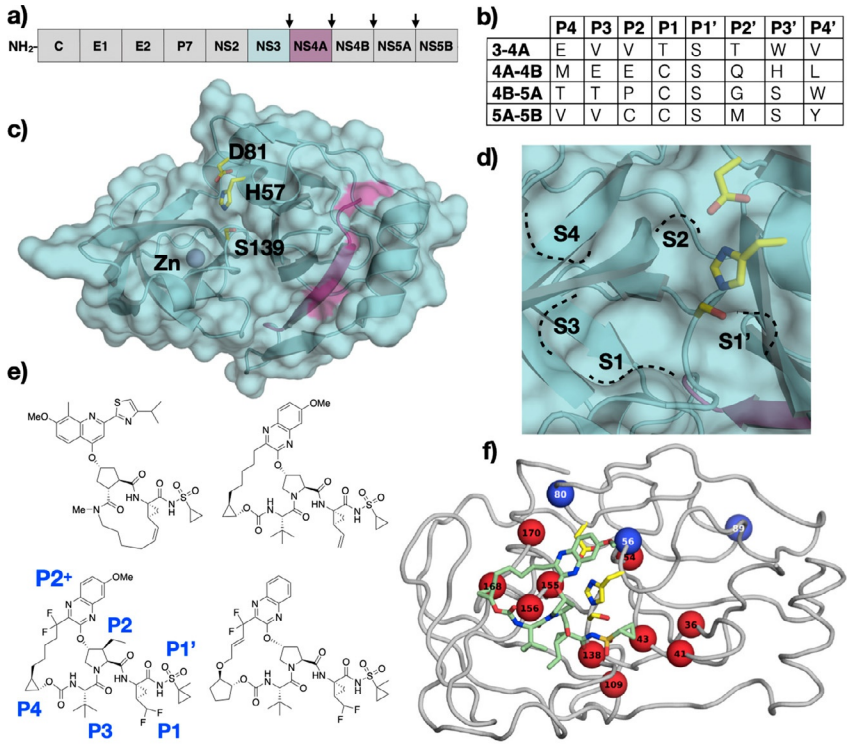


Fig. 2 HCV NS3/4A Protease. (A) Schematic representation of HCV polyprotein, with viral protease cleavage sites between non-structural proteins (NS) indicated by arrows. (B) Amino acid sequence of polyprotein cleavage sites recognized by HCV NS3/4A protease. The protease cleaves between P1-P1′. (C) The crystal structure of HCV NS3/4A protease (PDB ID: 3M5O) with the catalytic residues depicted as yellow sticks and labeled, and (D) a close-up view of the active site with S4-S1′ pockets annotated. (E) The chemical structure of HCV NS3/4A protease inhibitors (top, left) simeprevir, (top, right) grazoprevir, (bottom, left) voxilaprevir, and (bottom, right) glecaprevir. (F) Cocrystal structure of grazoprevir (green sticks) bound to HCV NS3/4A protease, with primary (red) and secondary (blue) resistance-associated substitutions labeled. *Panel F: Reprinted with permission from A.N. Matthew, F. Leidner, G.J. Lockbaum, M. Henes, J. Zephyr, S. Hou, et al., Drug design strategies to avoid resistance in direct-acting antivirals and beyond. Chem. Rev. 121(6) (2021) 3238–3270. Copyright (2021) American Chemical Society.*

domain contains all the catalytic residues and the substrate binding pocket. However, the protease activity is significantly enhanced by the NS4A protein which serves as a cofactor. In total, four sites in the polyprotein are cleaved by the NS3/4A protease between non-structural (NS) viral proteins (Fig. 2B) [32–37]. Cleavage at the NS2/NS3 junction is carried out by a

second viral encoded protease, the NS2 autoprotease, that is responsible for cleaving only that site [38]. Similar to HIV-1 protease, there are multiple reviews on HCV NS3/4A protease as a therapeutic target. Here, we will provide a concise overview and highlight progress made over the past decade, specifically in direct-acting antivirals (DAAs) against the HCV NS3/4A protease, which represents a triumph in drug design.

2.2.1 HCV NS3/4A protease structure

The protease domain of NS3 is composed of 180 amino acids and has a chymotrypsin-like fold with a zinc binding site opposite to the active site that stabilizes the protease tertiary structure (Fig. 2C). A minimal domain at the C-terminus of the NS4A cofactor is sufficient to increase the activity of the NS3/4A protease. In complex with NS3, NS4A forms a β-strand that extends through the core of the NS3 protease domain, more than 10 Å away from the active site. The NS4A β-strand constitutes one of the eight β-strands that make up one of the two β-barrels of the NS3/4A protease. The Ser-His-Asp catalytic triad is located between the two β-barrels (Fig. 2D). Unlike cellular chymotrypsin-like serine proteases with similar overall folds, the HCV NS3/4A protease lacks extended surface loops leading to an active site that is relatively solvent exposed, except when the helicase domain covers the active site.

2.2.2 Substrate recognition

HCV NS3/4A protease cleaves the viral polyprotein at four sites and recognizes a ten amino acid sequence (P6-P4'), of which only three positions are conserved. All the substrates bind in a conserved β-sheet conformation in the active site with similar hydrogen bonding patterns. Furthermore, the P1 residue of the substrate, which is either a cysteine or threonine, forms van der Waals interactions with an aromatic residue (F154) of the S1 pocket. Despite only the P2 position of the NS4/NS5A cleavage site being a proline, the P2 residues at the other cleavage sites adopt a phi torsion angle to maintain a binding mode similar to that of the proline. The P6 residue of the substrate is either aspartate or glutamate and forms a salt bridge interaction with K165 in the S6 pocket or with R123 in the case of the NS4B/NS5A cleavage site [39]. The other sites in the substrate have no sequence conservation. The approach used in the HIV-1 protease to understand diverse substrate recognition was successfully applied to the HCV NS3/4A protease to determine the substrate envelope.

2.2.3 The HCV NS3/4A substrate envelope

The HCV NS3/4A protease also recognizes diverse substrate amino acid sequences in a conserved shape despite having a solvent exposed active site, further validating the substrate envelope hypothesis as a general approach to understanding diverse substrate recognition [39,40]. Lack of cocrystal structures of the protease bound to peptide substrates spanning the non-prime and prime side limit our knowledge to only the non-prime side of the substrate. The prime side binds stronger to the protease active site where also all current inhibitors bind.

2.2.4 Enzyme catalysis

The HCV NS3/4A protease contains the canonical Ser-His-Asp catalytic triad and follows the general acyl transfer mechanism of serine proteases. Together, the catalytic histidine (H57) and aspartate (D81) increase the nucleophilicity of the serine (S139), thereby increasing the reaction rate with the scissile carbonyl carbon of the substrate [41]. The deprotonated S139 hydroxyl attacks the carbonyl carbon of the substrate, generating an acylated tetrahedral intermediate. The high-energy intermediate is stabilized by the oxyanion hole composed of backbone amide nitrogen atoms of residues 137–139 and the Nϵ of the catalytic histidine. The scissile amide nitrogen is then protonated by H57 generating a new N-terminus and leaving an ester acylated protease that is hydrolyzed by a water molecule to generate a new C-terminus and completing the catalytic cycle [42].

2.2.5 Inhibitors of HCV NS3/4A protease as drugs

The nucleophilic nature of the catalytic serine has been exploited in drug design leading to the development of the first-generation covalent HCV NS3/4A protease inhibitor drugs (telaprevir, boceprevir, and narlaprevir). The second-generation inhibitors departed from the ketoamide warhead and introduced bulky aromatic $P2^{+}$ moieties that are attached via an ether linker to the P2 group (asunaprevir). Macrocyclization as a strategy to rigidify and pre-organize the inhibitors for binding gave rise to P1-P3 (simeprevir) and P2-P4 (grazoprevir) macrocyclic protease inhibitors (Fig. 2E). The latest generation inhibitors are P2-P4 macrocycles and addressed the need to treat genotype 3 (GT-3) infected patients [43]. These compounds (voxilaprevir and glecaprevir) are very similar to grazoprevir and incorporate fluorine atoms that are in part responsible for the pan-genotypic activity [44]. Only five of the twelve FDA-approved HCV NS3/4A inhibitors are currently used in the clinic primarily due to drug resistance [45].

2.2.5.1 P1-P3 macrocycles

P1-P3 macrocyclic inhibitors were designed to overcome the low barrier to resistance of the previous generation drugs [46]. The P1-P3 macrocyclic scaffold links the P1 and P3 groups that reside proximal to each other in the cocrystal structures. The P1-P3 macrocyclic protease inhibitors share the same cycloprolylsulfonamide P1', P1, and five-membered cyclic P2 moieties found in the linear competitive inhibitors. However, the P3 *tert*-butyl group of the linear inhibitors is replaced with a alkyl linker that is connected to the terminal olefin of the P1 group. Simeprevir was the first macrocyclic inhibitor approved for the treatment of GT-1 and GT-4. The P1' cyclopropyl sulfonamide group of simeprevir fits well into the S1' pocket of the HCV NS3/4A protease. The P1-P3 macrocycle complements the groove formed by the S1 and S3 pockets. The $P2^+$ moieties of these inhibitors adopt a binding mode similar to that of the linear competitive inhibitors, interacting with H57, D81, and R155. Simeprevir lacks a P4 group and therefore does not interact with the S4 pocket but still makes van der Waals interactions with A156 and A157 that form one side of the S4 pocket. Danoprevir, a P1-P3 macrocyclic inhibitor that contains a *tert*-butyl P4 group makes additional interactions with D168 in the S4 pocket.

2.2.5.2 P2-P4 macrocycles

Although the P1-P3 macrocyclic HCV NS3/4A inhibitors represented a significant improvement in potency and pharmacokinetic/pharmacodynamic properties, the lack of efficacy against the GT-3 variant and susceptibility to resistance drove the development of the P2-P4 macrocycles. The macrocycle of these inhibitors links the $P2^+$ aromatic moiety and the P4 group. While the P2-P4 macrocycle in grazoprevir is an unsubstituted alkyl linker, that of voxilaprevir and glecaprevir contain two fluorine atoms at the benzylic position. Additionally, glecaprevir's macrocycle contains an alkene in the *E* conformation potentially to decrease flexibility. Voxilaprevir and glecaprevir contain a methyl substituent at the P1' cyclosulfoaminde moiety that was present in previous generation drugs, which orients the P1' group deeper into the S1' pocket [47,48]. Grazoprevir has the same terminal olefin unnatural amino acid at the P1 moiety found in the linear inhibitors. The terminal olefin of that P1 group is replaced with a difluoromethylene moiety in voxilaprevir and glecaprevir. All three P2-P4 macrocycles contain a P2 proline core bonded to a quinoxaline moiety. The quinoxaline substitution patterns differ, with grazoprevir and voxilaprevir containing a 7-OMe substituent while glecaprevir is unsubstituted at that position. All three drugs

contain the identical P3 *tert*-butyl group found in the previous generation inhibitors, and the carbamate linkage to attach a cyclopropyl P4 group in the case of grazoprevir and voxilaprevir or cyclopentyl P4 group in the case of glecaprevir. While grazoprevir is approved against GT-1, voxilaprevir and glecaprevir are pan-genotypic inhibitors and are approved against all the genotypes (GT1–6), highlighting the biological significance of the subtle differences between these compounds [44].

These three latest inhibitors bind to the HCV NS3/4A protease in nearly identical binding poses (Fig. 2F). The $P2^+$ quinoxaline adopts a binding pose distinct from the $P2^+$ moieties of earlier generation inhibitors. The quinoxaline avoids contact with R155, a common resistance-associated substitution (RAS) site, and is oriented in such a way to interact primarily with invariant residues D75 and H57 of the catalytic triad (Fig. 2F). The cyclopropyl P4 group of grazoprevir and voxilaprevir binds but does not fully occupy the S4 pocket, while cyclopentyl P4 cap of glecaprevir can better fill in the S4 pocket [48]. In voxilaprevir and glecaprevir, the difluoromethylene moiety contributes fluorine-specific interactions, including two orthogonal multipolar interactions. Furthermore, the high electronegativity of fluorine compared to hydrogen promotes the hydrogen of the difluoromethylene moiety to form a "caged" fluorine-induced hydrogen bond with the backbone carbonyl oxygen of the protease and the carbonyl oxygen of the P1 group thereby pre-organizing the inhibitor for binding. These fluorine-specific interactions occur with the backbone atoms, thereby improving the potency against proteases from all genotypes, and decreasing the susceptibility to resistance [44]. The two fluorine atoms on the macrocycle of voxilaprevir and glecaprevir mainly contribute to improving metabolic stability.

2.2.6 Resistance associated substitutions (RAS) and genotypic differences

Although the chemical structure of the HCV NS3/4A inhibitors differs significantly, there are three major RASs that confer resistance to all, especially earlier generation, inhibitors (Fig. 2F). One of these RAS, R155K, leads to resistance by disrupting the electrostatic network spanning R123, D168, R155, and D81, which is important for constructing a flat binding surface for inhibitors [49,50]. Inhibitors with bulky aromatic $P2^+$ moieties that interact with R155 for potency are susceptible to this RAS because of the loss of a cation-π interaction with the guanidinium group [49]. The $P2^+$ quinoxaline moiety of the P2-P4 macrocycles does not interact with

R155 but instead interacts primarily with the catalytic triad. Therefore, the R155K RAS only causes subtle changes in the binding modes of grazoprevir and glecaprevir [43,44,51]. This unique binding mode of the quinoxaline $P2^+$ moiety is independent of macrocyclization and underlies the improved potency of later generation inhibitors [52]. The RASs at A156 lead to resistance to all inhibitors by causing steric hindrance with either the P2, P4, or the P2-P4 macrocycle moiety. Interestingly, the capacity of the ketoamide warhead group of the first-generation inhibitors to covalently modify the protease is also compromised by the A156T RAS. The RAS and polymorphisms at D168 cause resistance by disrupting the salt bridge with R155 and thereby increasing the dynamics of the S4 pocket that accommodate the P4 group or the P2-P4 macrocycle and also impact genotypic differences [44,51,53,54].

2.2.7 The HCV NS3/4A substrate envelope and resistance

The HCV NS3/4A substrate envelope explains the molecular basis of diverse substrate recognition and also serves as a tool to elucidate the susceptibility of inhibitors to resistance. As the consensus volume adopted by the four viral polyprotein substrates, the substrate envelope represents the essential recognition feature. Inhibitor-bound structures reveal areas where inhibitors protrude from the substrate envelope. The major RAS occur at A156, R155, and D168 where the inhibitors protrude beyond the envelope and contact the active site, rendering these residues more important for inhibitor binding compared to substrate recognition [21,40,45,49,52]. The $P2^+$ aromatic group of the P1-P3 macrocyclic drugs interact outside the substrate envelope with R155 and in some cases Y56, which are residues not essential for substrate recognition [55]. Although the large aromatic $P2^+$ quinoxaline moiety of the P2-P4 macrocyclic drugs protrude out of the substrate envelope, the interactions are limited to catalytic residues essential for enzymatic activity, D81 and H57, thereby decreasing the susceptibility to RAS at R155. The position of the macrocycle also plays a role in selecting for RASs. The P1-P3 macrocycle resides in the S1 and S3 pockets and fits within the substrate envelope. In contrast, the P2-P4 macrocycle protrudes out to contact A156, and selects for RAS primarily at this position and D168 (sometimes in combination with Y56) [55]. Inhibitors have been designed to decrease the susceptibility to resistance due to the P2-P4 macrocycle by relocating the macrocycle at the P1-P3 position while maintaining the unique binding mode of the $P2^+$ quinoxaline moiety [43]. These P1-P3 macrocyclic inhibitors maintained potency and exhibited a flatter resistance profile compared to the P2-P4 macrocyclic analogs [48,56].

3. Viral proteases as emerging therapeutic targets

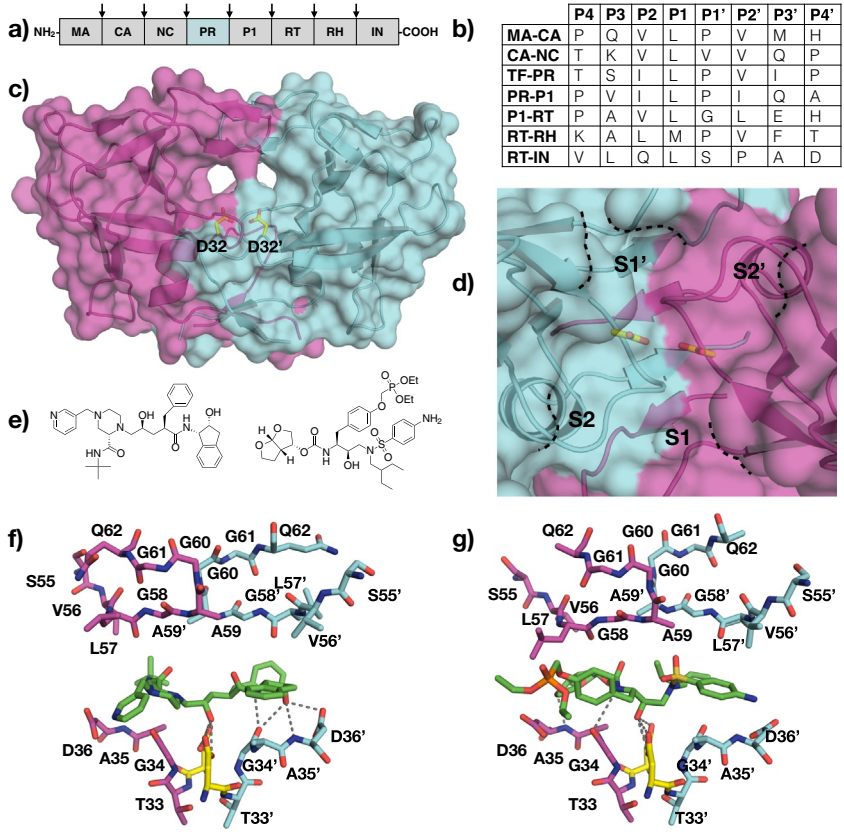
The successful targeting of the HIV-1 protease established viral proteases as therapeutic targets and paved the way to the relatively rapid success in developing potent pan-genotypic inhibitors of the HCV NS3/4A protease. Other viruses of global health concern that are dependent on a viral protease for replication are also potentially tractable therapeutic targets. These emerging targets include proteases from the human T-cell leukemia virus type 1 (HTLV-1), flaviviruses, enteroviruses, and coronaviruses.

3.1 HTLV-1 protease

HTLV is a retrovirus, belonging to the *Deltaretrovirus* genus. HTLV-1 is associated with various diseases including adult T-cell leukemia lymphoma (ATL) and HTLV-1-associated myelopathy. HTLV was the first retrovirus to be identified and currently, an estimated 20 million people are infected worldwide. ATL caused by HTLV-1 has poor prognosis and is resistant to conventional and high-dose chemotherapy [57]. Unfortunately, there are no vaccines or direct-acting antivirals against HTLV-1. Similar to the replication of other retroviruses, the HTLV-1 virus genome encodes an aspartyl protease that is critical for polyprotein processing and viral maturation (Fig. 3A and B) [58]. Given the similarities in structure and function to the HIV-1 protease, the HTLV-1 protease can be a primary target for drug design. Certain aspects of the HTLV-1 protease have been previously reviewed [59,60]; here we will provide an overview of the protease structure and substrate recognition, with recent advances in inhibitor development.

3.1.1 HTLV-1 protease structure

The HTLV-1 protease is a homodimeric aspartyl protease, similar to the HIV-1 protease (Fig. 3C). However, the HTLV-1 protease is larger with each monomer comprising of 125 amino acids instead of 99. The additional amino acids, relative to the HIV-1 protease, are located at the C-terminus and are a distinct feature of the HTLV-1 protease compared to other viral aspartyl proteases. The crystal structure of HTLV-1 protease revealed a very similar fold to that of the HIV-1 protease. Both proteases share a conserved core structure around the active site and at the dimerization interface, with structural divergences in the loops around the active site and the flaps [59].



	P4	P3	P2	P1	P1'	P2'	P3'	P4'
MA-CA	P	Q	V	L	P	V	M	H
CA-NC	T	K	V	L	V	V	Q	P
TF-PR	T	S	I	L	P	V	I	P
PR-P1	P	V	I	L	P	I	Q	A
P1-RT	P	A	V	L	G	L	E	H
RT-RH	K	A	L	M	P	V	F	T
RT-IN	V	L	Q	L	S	P	A	D

Fig. 3 HTLV-1 Protease. (A) Schematic representation of HTLV-1 polyprotein, with viral protease cleavage sites indicated by arrows. (B) Amino acid sequence of polyprotein cleavage sites. The protease cleaves between P1-P1′. (C) The crystal structure of HTLV-1 protease (PDB ID: 6W6Q) with the catalytic residues depicted as yellow sticks and labeled, and (D) a close-up view of the active site with S2-S2′ pockets annotated. (E) (*left*) Indinavir and (*right*) PU6 (e), with their respective cocrystal structures bound to the HTLV-1 protease (PDB ID: 3WSJ and 6W6S) (F and G).

A unique aspect of the HTLV-1 protease is an additional two residues in the flap regions relative to that of HIV-1 that increases the helical conformation of the flap regions [61].

The dimeric structure of HTLV-1 protease is held together by four-stranded antiparallel β sheets that involve the N and C termini of each monomer. The active site is located at the cleft region above the dimer interface where one aspartate from each monomer, positioned at the base of the

cleft, forms the catalytic dyad. Each monomer contributes a β turn that together constitute the flaps and enclose the active site [59]. The additional residues at the flap region create a more extended substrate-binding site compared to that of the HIV-1 protease [58].

3.1.2 Substrate recognition

The substrates of HTLV-1 protease are highly hydrophobic and diverse in sequence, except for the P1 position where there is a preference for leucine (Fig. 3B). The side chain interactions confer specificity while backbone interactions properly orient the substrate in the active site [59]. The strongest interactions between the protease and substrates occurs within S3-S3' pockets [59]. There are no substrate-bound structures of the HTLV-1 protease available to date. However, the cocrystal structure of the protease bound to a statine-containing peptide provides insights into substrate recognition. Each carbonyl oxygen and amide group of the peptidic inhibitor participates in direct hydrogen bond interaction with the enzyme in a manner identical to other viral aspartyl proteases [61]. The HTLV-1 protease can accommodate larger residues in S1/S1' and S2/S2' pockets relative to the HIV-1 protease (Fig. 3D). The flaps also play a crucial role in dictating substrate specificity although hydrogen bonding interactions with the interior of the protease are stronger compared to that of the flaps. The HTLV-1 S3/S3' pockets differ significantly compared to those of HIV-1 protease. The S3 pocket of the HTLV-1 protease is partially solvent exposed; consequently, P3 of the substrates can either be hydrophobic or hydrophilic. The S4/S4'subsites of HTLV-1 are larger and more hydrophobic compared to those of the HIV-1 protease, consistent with the observed substrate preference for more hydrophobic P4 residues. Overall, substrate-binding pockets more distal from the catalytic site are more divergent between the two proteases.

Unlike HIV-1, which utilizes both viral and host molecular machinery to replicate, HTLV-1 is thought to replicate primarily in the DNA form and therefore has to force the host cell to divide [61,62]. This aspect of the HTLV-1 life cycle does not require viral machinery for transcription, which tends to be error-prone; therefore, the HTLV-1 sequence is more conserved compared to that of HIV-1. This implies that HTLV-1 protease will be less likely to evolve under drug pressure to develop resistance.

3.1.3 Inhibitors of HTLV-1 protease

The current treatment options for ATL caused by HTLV-1, including chemotherapy and antibodies against the interleukin 2 receptor, offer limited success. Inhibitors targeting the HTLV-1 protease provide an attractive novel

approach for developing therapeutics. The search for a drug targeting the HTLV-1 protease is in very early stages. The two main strategies being investigated are designing peptidomimetic inhibitors based on the substrate sequence, and repurposing HIV-1 protease inhibitors. Of the HTLV-1 protease peptidomimetic inhibitors reported, statine-containing peptides have reached nanomolar potencies [63,64]. The similarity (28% overall and 45% in the active site) between HTLV-1 and HIV-1 proteases serves as the rationale to leverage FDA-approved HIV-1 protease inhibitors against the HTLV-1 protease. Three of these inhibitors, ritonavir, indinavir, and darunavir were found to inhibit HTLV-1 protease with varying effectiveness (Fig. 3E). These three peptidomimetic inhibitors contain hydrophobic P1/P1' moieties and a secondary hydroxyl group as a transition state mimetic. Ritonavir, with a $K_i > 20\,\mu M$ was found to induced apoptosis in HTLV-1 infected cells [57]. Indinavir weakly inhibits the HTLV-1 protease with a K_i of $3.5\,\mu M$ [65]. A cocrystal structure of indinavir bound to the HTLV-1 protease reveals that the central hydroxyl of indinavir contacts both catalytic aspartates and occupies S3-S3' pocket (Fig. 3F) [66]. However, the larger HTLV-1 protease active site prevented optimal packing with the flap residues. DRV, a picomolar inhibitor of the HIV-1 protease, has a K_i of about $1\,\mu M$ against the HTLV-1 protease [67]. Cocrystal structures of DRV bound to HIV-1 and HTLV-1 proteases revealed a similar binding mode. Given the larger binding pockets of HTLV-1 protease, DRV analogues with bulkier substitutions were also investigated. A DRV analog, UM6, containing a bulky and hydrophobic P1' moiety was 10-fold more potent against HTLV-1 protease compared to DRV. Moreover, substituting a diethylphosphonate group at the P1 phenyl group to generate inhibitor PU6, led to further improvement in the potency to a K_i of $0.03\,\mu M$. The cocrystal structure of PU6 bound to the HTLV-1 protease revealed that the phosphonate moiety extends into the S1 pocket and increases van der Waals contacts with the protease (Fig. 3G). UM6 and PU6 were also effective at inhibiting Gag processing in biochemical assays and maturation of virus-like particles [67]. The cocrystal structures of DRV and analogs bound to HTLV-1 protease provide a promising path toward designing potent HTLV-1 protease inhibitors as therapeutics.

3.2 Flavivirus proteases

The *flavivirus* genus contains over 70 mosquito-borne viruses, including human pathogens such as Zika (ZIKV), dengue (DENV), and West Nile (WNV) virus [68]. Viruses within this genus are highly similar, with 56% sequence identity between ZIKV and dengue serotype 2 (DENV-2) [69].

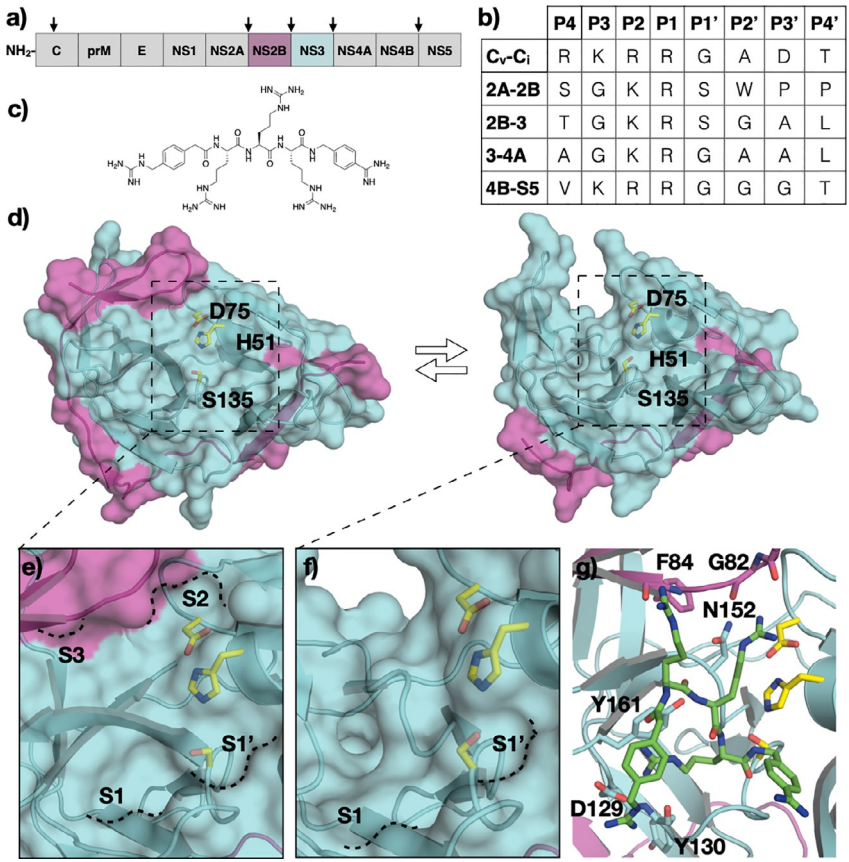
The ZIKV outbreak in 2015 raised public concern due to sexual and maternal-fetal mode of transmission and infections causing neurological disorders, especially microcephaly in neonates [70,71]. Meanwhile, the persistent DENV infections have been a global health concern since the 1950s [72]. Hemorrhagic fever is the most severe pathology of DENV-infected patients [73]. Similar to other positive-strand RNA viruses, flaviviruses encode a viral protease that processes the viral polyprotein at multiple sites. The flaviviral proteases share a high level of sequence identity (~30–80%) and are composed of two viral non-structural proteins, the NS2B cofactor and the NS3 protease (Fig. 4A) [74]. Flaviviral proteases require an Arg/Lys at the P1 and P2 positions of the substrate (Fig. 4B), which greatly influences chemical features of peptidomimetic inhibitors that are being developed (Fig. 4B and C). We will focus on the viral proteases from DENV and ZIKV, the two viruses that have caused the most health burden within the genus.

3.3 Structure of flavivirus NS2B/NS3 protease

The flavivirus NS2B/NS3 protease is a chymotrypsin-like viral protease, with a canonical Ser-His-Asp catalytic triad (Fig. 4D). Similar to HCV, the NS3 of flaviviruses contains the protease domain at the N-terminus and a helicase domain at the C-terminus. However, unlike HCV NS3 protease, the flaviviral NS3 protease is inactive without the NS2B cofactor, which stabilizes the NS3 protease structure and is essential for proper folding [75–78]. Crystal structures of flaviviral proteases revealed that the NS2B is dynamic and can adapt an open or closed conformation. In the open conformation, only one β strand portion of the NS2B interacts with the NS3 protease, while the rest of the NS2B is disordered. In the closed conformation, the NS2B fully wraps around NS3 protease domain and forms an additional β turn that contributes residues to construct part of the active site S2 pocket and interact with P2 of the substrates. In the open conformation, the S2 pocket is not present; therefore, the protease cannot stably bind substrates (Fig. 4E and F). The active site of the flaviviral proteases are relatively flat and solvent exposed. Upon substrate binding, the reaction mechanism of flaviviral proteases follows that of the general chymotrypsin-like serine protease family, similar to that of the HCV NS3/4A protease.

3.3.1 Substrate recognition

The flaviviral polyprotein contains transmembrane domains that insert the polyprotein into the ER membrane as translation by the host ribosomes



	P4	P3	P2	P1	P1'	P2'	P3'	P4'
C_v-C_i	R	K	R	R	G	A	D	T
2A-2B	S	G	K	R	S	W	P	P
2B-3	T	G	K	R	S	G	A	L
3-4A	A	G	K	R	G	A	A	L
4B-S5	V	K	R	R	G	G	G	T

Fig. 4 Flavivirus Proteases. (A) Schematic representation of flavivirus polyprotein, with viral protease cleavage sites indicated by arrows (B) Amino acid sequence of polyprotein cleavage sites. The protease cleaves between P1-P1′. (C) The chemical structure of flaviviral protease inhibitor, compound **1**. (D) The crystal structure of Zika virus (ZIKV) NS2B/3 protease in the (left) closed and (right) open conformation (PDB ID: 5LC0 and 5GXJ), with the catalytic residues labeled. (E) The active site topology of the closed (E) and open (F) conformation with S3-S1′ pockets annotated. (G) The cocrystal structure of compound **1** bound to the ZIKV NS2B/3 protease in the closed conformation.

takes place. The NS2B/NS3 protease is responsible for processing all the cleavage sites on the cytosolic side of the polyprotein, including NS2A/NS2B, NS2B/NS3, NS3/NS4A, and NS4B/NS5 (5, 6). The binding mode of peptides corresponding to the non-prime side of the substrates has been revealed through a crystal structure of ZIKV protease caught in *cis* cleavage and other cocrystal structures with covalent peptidic inhibitors [79]. The C-terminus generated from scissile bond cleavage is stabilized in the

oxyanion hole through backbone hydrogen bonding interactions. The P1 arginine extends into the S1 pocket and forms a salt bridge with Asp129 at the base of the pocket. The P1 guanidinium group also participates in hydrogen bonding interaction with the backbone carbonyl of Tyr130, and forms cation-π interaction with Tyr161. In the S2 pocket, the positive charge of P2 lysine is stabilized by side chain of Asp75 of the catalytic triad and Asp83 of the NS2B cofactor. In the *cis* cleavage structure, the P3 glycine residue adopts a rotamer that orients the P4 and P5 residues toward solvent. Therefore, this structure does not provide information on interactions at S3 and S4 pockets. The cocrystal structure with bound compound **1** revealed that the P3 arginine makes polar contacts with the carbonyl of F84 of NS3 and the sidechain of S85 of NS2B (Fig. 4C and G). The P4 group of compound **1** does not extend into the S4 pocket but instead is solvent exposed, with the compound binding in a cyclic conformation [80].

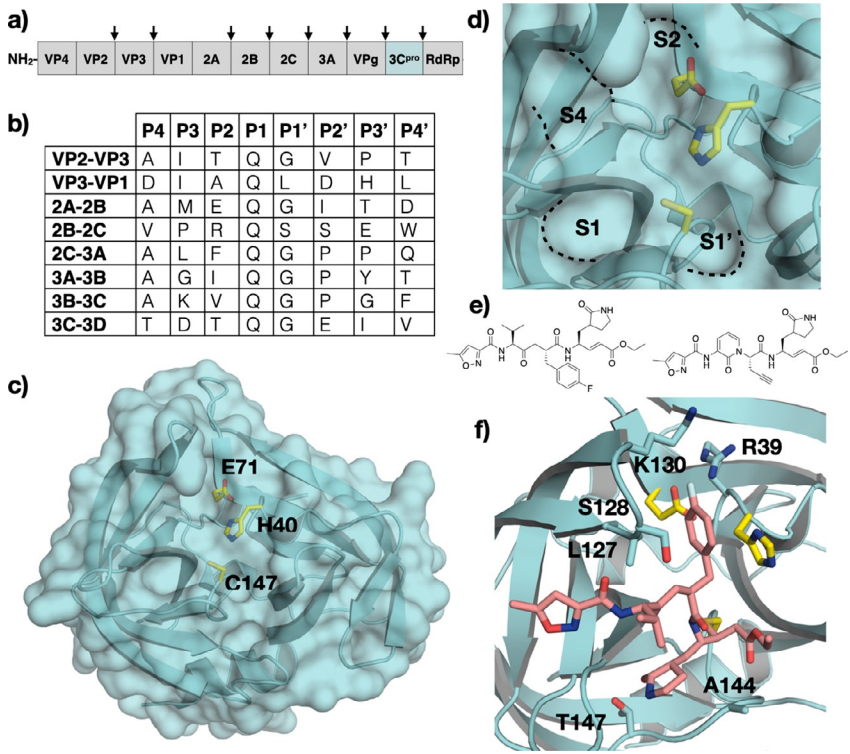
3.3.2 Inhibitors of flavivirus proteases

Given the overall high similarities between flaviviral proteases, inhibitors identified against one flaviviral protease should be generally effective against others, and support the ultimate aim of discovering pan-flaviviral protease inhibitors. Despite significant efforts, none of the flaviviral protease inhibitors have yet reached the clinical stage. The flat and highly negatively charged active site led to the design of inhibitors with poor cellular activity and pharmacokinetics [81–83]. The traditional approach of designing inhibitors against serine proteases, which involves the addition of electrophilic warhead groups to the non-prime side of the substrate, led to potent inhibitors with poor selectivity, cellular activity, and pharmacokinetics [84]. The initial peptidomimetic inhibitors led to the development of compound **1** ($K_i = 0.43\,\mu M$) which showcased potency could be achieved without a warhead group; however, with five guanidinium moieties compound **1** does not have desired drug-like properties [80]. The macrocyclization drug design strategy, which has proven successful in inhibiting the HCV NS3/4A protease, was applied to the ZIKV protease with limited success [85]. Non-peptidic and small aromatic active site inhibitors with low levels of potency have also been identified [86]. Other methods of identifying inhibitors such as high-throughput screening and fragment-based drug design have also been investigated with limited success [87–89]. There are several reviews available on the inhibition of flaviviral proteases [74,81,84,90,91]. The binding loop of aprotinin, a general protease inhibitor

that is also active against flaviviral proteases, has been used as the basis for cyclic peptidic inhibitors, and to understand the contribution of prime side recognition to potency [92,93]. Although there is significant structural information on the binding mode of the peptidomimetic inhibitors, the bias toward cationic scaffolds limits the progress of drug design to the clinic. To overcome this limitation, allosteric inhibitors have also been investigated, with reports of inhibitors aiming to interfere with the NS3 and NS2B binding interface [90,94–101]. However, there is no clear evidence to validate the binding mode of these inhibitors. Drug repurposing as an alternative approach has also been investigated. Specifically, efforts targeting the ZIKV protease have led to the identification of micromolar inhibitors, including bromocriptine, temoporfin, and hydroxychloroquine [89,102,103]. Similar to other allosteric inhibitors identified, detailed characterization of inhibitor binding mode is lacking, but is essential for structure-based drug design and to guide medicinal chemistry efforts.

3.4 Enterovirus proteases

The genus *Enterovirus* (EV) of the *Picornaviridae* family includes poliovirus, numbered rhinoviruses and enteroviruses. Poliovirus is the best-known and one of the most studied enterovirus for causing the debilitating poliomyelitis disease [104]. Enteroviruses that do not cause poliomyelitis are classified as non-polio enteroviruses (NPEV), many of which are important human pathogens. EV-A71 and EV-D68 have recently emerged as serious public health threats. EV-A71 causes hand-foot-and-mouth disease (HFMD) with life-threatening neurological complications, such as brainstem encephalitis. EV-D68 can cause mild respiratory disorders as well as more severe symptoms such as acute flaccid paralysis in children [105,106]. Human rhinoviruses are the predominant agents responsible for the common cold [107]. Two EV-A71 inactivated enteroviruses are approved in China as vaccines and represent the only preventive measure against NPEV. Enterovirus has a positive-strand RNA genome that is translated into a single large polyprotein to be proteolytically processed into ten structural and replication proteins by two virally encoded proteases, $2A^{pro}$ and $3C^{pro}$ (Fig. 5A and B). Inhibition of these viral proteases represents a tractable approach to treat enterovirus infections, as both proteases play crucial roles in polyprotein processing: $2A^{pro}$ cleaves only one site on the polyprotein while the $3C^{pro}$ cleaves all the other sites. The $3C^{pro}$ is a highly conserved cysteine protease among enteroviruses possessing a non-canonical Cys-His-Glu catalytic triad.



	P4	P3	P2	P1	P1'	P2'	P3'	P4'
VP2-VP3	A	I	T	Q	G	V	P	T
VP3-VP1	D	I	A	Q	L	D	H	L
2A-2B	A	M	E	Q	G	I	T	D
2B-2C	V	P	R	Q	S	S	E	W
2C-3A	A	L	F	Q	G	P	P	Q
3A-3B	A	G	I	Q	G	P	Y	T
3B-3C	A	K	V	Q	G	P	G	F
3C-3D	T	D	T	Q	G	E	I	V

Fig. 5 Enterovirus 3C Protease ($3C^{pro}$). (A) Schematic representation of the enterovirus polyprotein, with $3C^{pro}$ cleavage sites indicated by arrows. (B) Amino acid sequence of polyprotein cleavage sites. The protease cleaves between P1-P1′. (C) The crystal structure of the EV-A71 $3C^{pro}$ (PDB ID: 7DNC) with the catalytic residues labeled, and (D) a close-up view of the active site with S4-S1′ pockets annotated. (E) Chemical structures of rupintrivir and AG7404. (F) Cocrystal structure of rupintrivir bound to EV-A71 $3C^{pro}$ (PDB ID: 3SJO).

3.4.1 Structure of enterovirus $3C^{pro}$ proteases

Enterovirus $3C^{pro}$ proteases belong to the chymotrypsin-related endopeptidase protease family and are biologically active as monomers [108,109]. The catalytic triad and the surrounding residues of $3C^{pro}$ are highly conserved among the *Enterovirus* genus. The EV-A71 $3C^{pro}$ consists of two twisted β-ribbon folds packed in a perpendicular orientation (Fig. 5C). The substrate-binding pocket lies at the groove between the two β-ribbon folds. The catalytic cysteine in the active site acts as a nucleophile and attacks the carbonyl carbon of the scissile bond following a mechanism similar to that of the HCV NS3/4A serine protease.

3.4.2 Substrate recognition

Enterovirus $3C^{pro}$ cleaves the viral polyprotein at eight sites, each constituting a unique P4-P4' sequence with a strong preference for glutamine at the P1 position and glycine at the P1' position (Fig. 5B). The other positions within the P4 to P4' substrate can accommodate diverse amino acids. The interactions at S3-S1' pockets of the protease contribute most to substrate binding. There are no available crystal structure of enterovirus $3C^{pro}$ bound to a natural substrate peptide. However, cocrystal structures of EV-D68 $3C^{pro}$ bound to peptidomimetic inhibitors provide insight into the molecular basis of substrate recognition [110,111]. The S2 pocket is hydrophobic, hence the preference for hydrophobic P2 moieties. The EV-D68 $3C^{pro}$, similar to other homologous viral proteases, does not possess an S3 pocket which leads to a diverse set of amino acids at the P3 position of the substrates. The S4 pocket is very shallow and can accommodate only small residues (Fig. 5D).

3.4.3 Inhibitors of enterovirus $3C^{pro}$ proteases

There are no FDA-approved inhibitors targeting the enterovirus $3C^{pro}$. Rupintrivir, a peptidomimetic inhibitor of rhinovirus $3C^{pro}$ and an analogue (AG7404) have advanced to clinical trials, but failed due to limited efficacy (Fig. 5E) [112,113]. Cocrystal structures of rupintrivir bound to the EV-A71 $3C^{pro}$ reveal the structural basis for antiviral activity [114,115]. The warhead group of rupintrivir forms a covalent bond with nucleophilic C147 and the rupintrivir scaffold binds via an extensive set of hydrogen bonds and hydrophobic interactions that span the S1'-S4 pockets (Fig. 5F). The majority of $3C^{pro}$ inhibitors are of similar peptidomimetic scaffolds, although non-peptidomimetic inhibitors have also been investigated [116].

3.5 Coronavirus proteases

The coronavirus disease 2019 (COVID-19) pandemic is caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [117]. Over 170 million confirmed infection cases have been reported worldwide as of mid-2021. The 2019 pandemic marks the third coronavirus outbreak in the 21st century [118,119]. SARS-CoV, Middle East respiratory syndrome (MERS-CoV), and SARS-CoV-2 belong to the *Coronaviridae* family under the *Betacoronavirus* genus. Similar to the other positive-sense single-stranded RNA viruses discussed in this chapter, coronaviruses encode viral proteases,

main protease (M^{pro}, also known as $3CL^{pro}$, or nsp5), and papain-like protease (PL^{pro}, also known as nsp3), to cleave the viral polyprotein [120]. M^{pro} is seen as a more tractable therapeutic target because of its significant role in polyprotein processing (Fig. 6A and B) [121]. Coronaviruses share a high overall sequence identity. Consequently, M^{pro} from SARS-CoV-1 and SARS-CoV-2 have ~96% sequence identity [122,123]. Here we will focus on the structure, mechanism, and inhibition of M^{pro} from SARS-CoV-2.

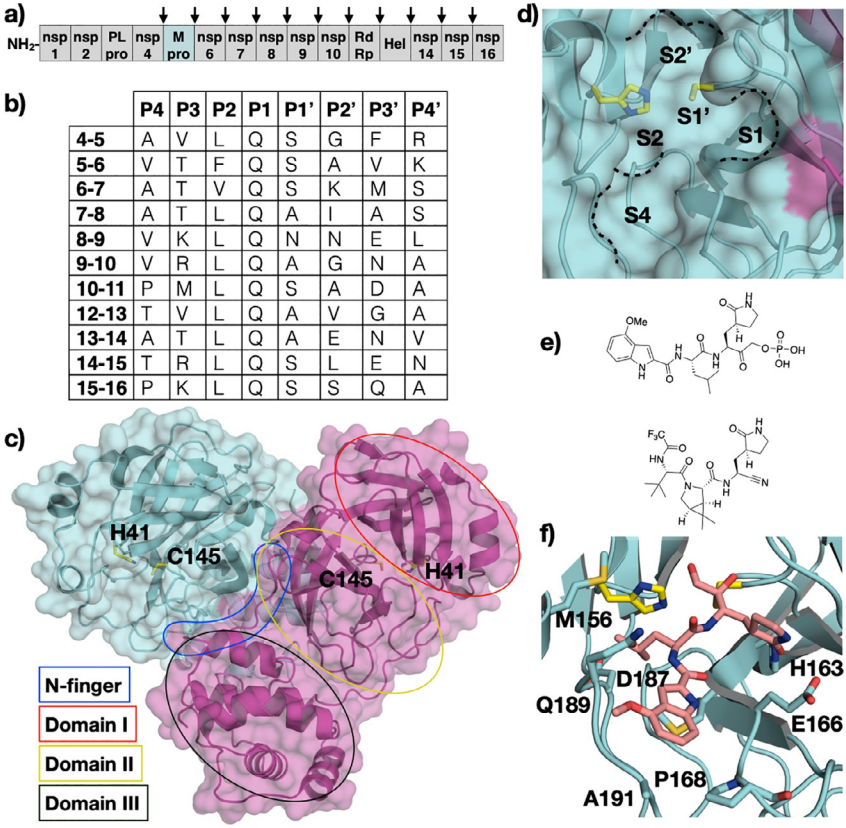


Fig. 6 Coronavirus Main Protease (M^{pro}). (A) Schematic representation of the coronovirus polyprotein, with M^{pro} cleavage sites indicated by arrows. (B) Amino acid sequence of polyprotein cleavage sites. The protease cleaves between P1-P1′. (C) The crystal structure of SARS-CoV-2 M^{pro} homodimer (PDB ID: 7KHP) with the catalytic residues labeled and the N-finger, domains I, II, and III colored accordingly on one of the monomers, and (D) a close-up focused view of the active site with S4-S2′ pockets annotated. (E) Chemical structures of M^{pro} inhibitors PF-07304814 and PF-07321332. (F) Cocrystal structure of PF-00835231 bound to SARS-CoV-2 M^{pro} (PDB ID: 6XHM).

3.5.1 Structure of coronavirus main protease

M^{pro} is a 306 residue protease with a catalytic dyad consisting of cysteine and histidine. The biologically active state of M^{pro} is dimeric, where two monomers are arranged in an orthogonal fashion (Fig. 6C) [124]. Each monomer is made up of three domains; domain I and domain II together construct two anti β-parallel barrels with a fold similar to that of trypsin-like serine proteases, and domain III contains five α-helices that are connected via a linker to domain II. Dimerization of the protomers occurs at the N-terminus (N-finger) located between domains II and III. The N-finger also contributes to the formation of the substrate-binding site between domains I and II [125,126].

3.5.2 Substrate recognition

In addition to the catalytic dyad composed of H47 and C145, M^{pro} contains a buried water molecule in the active site that substitutes for the aspartate residue that usually accompany cysteine and histidine in the canonical catalytic triad of other proteases (Fig. 6D) [121,127]. The catalytic process follows a multi-step mechanism, where H47 deprotonates the catalytic cysteine generating a nucleophilic thiolate that attacks the carbonyl carbon of the scissile bond and forms a tetrahedral intermediate. The tetrahedral intermediate collapses leaving an acylated intermediate. The histidine activates a water molecule that hydrolyzes the intermediate regenerating the thiol sidechain [128,129]. M^{pro} cleaves the viral polyproteins (pp1a and pp1ab) at 11 sites (Fig. 6B). The amino acid sequences at the cleavage sites are diverse, except at P1 which is always glutamine, and P2 which has a preference for leucine. The P3 position can accommodate hydrophobic and hydrophilic residues, and P4 has a preference for amino acids with small sidechains. The prime side of the cleavage sites lacks sequence conservation except at P1' where there is a preference for small amino acids. M^{pro} is expected to recognize substrates spanning P4-P1' based on peptide substrate analysis [130].

An acyl-enzyme intermediate crystal structure of SARS-CoV-2 M^{pro} bound to the C-terminal autoproteolysis product (nsp5-nsp6) provides details of non-prime side substrate recognition [131]. Cocrystals structures of SARS-CoV-1 M^{pro} bound to an 11-mer peptide (TSAVLQSGFRK) representing the nsp4-nsp5 elucidated the interactions that drive recognition of the prime side [132]. Comparing apo and substrate-bound structures revealed an induced-fit binding mechanism, where the width of the binding groove increases to accommodate the side chains of the substrates to bind in an extended β-sheet conformation. The backbone nitrogen of C145 and G143 stabilize the acyl-enzyme tetrahedral intermediate.

The P1 glutamine side chain is stabilized through hydrogen bonding with polar sidechains in the S1 pocket. The hydrophobic P2 group forms packing interactions with M165 in the deep S2 pocket. The P3 residue is partially solvent exposed, which may explain the lack of specificity at this position. The shallow S1' pocket explains the preference for small amino acids at the P1' position of the substrates. The S2' pocket is long and deep, explaining the diverse amino acids seen at the P2' position of the substrates.

3.5.3 Inhibitors of coronavirus main protease

Inhibitors of SARS-CoV-1 and -2 M^{pro} have been reviewed extensively [123,128,133,134]. Peptidomimetics and small molecules that contain Michael acceptors such as aldehydes and ketones make a significant portion of the designed inhibitors. A common feature of these inhibitors is a P1 γ-lactam group as a glutamine mimetic. Two dipeptidyl inhibitors, PF-07304814 and PF-07321332, are entering phase I clinical trials (NCT04535167 and NCT04756531) (Fig. 6E) [135,136]. PF-07304814, the prodrug form of PF-00835231, contains a phosphonate group for better solubility that gets cleaved by alkaline phosphatase enzymes in tissue. The orally available PF-07321332 was designed to overcome the intravenous method of delivery for PF-07304814. In addition, extensive high-throughput screening (HTS) of FDA-approved drugs and fragment-based libraries has been investigated against M^{pro} [133]. Notably, boceprevir, an FDA-approved HCV NS3/4A protease inhibitor, was found to inhibits SARS-CoV-2 M^{pro} with single digit IC_{50} value [133]. The rapid discovery of SARS-CoV-2 M^{pro} inhibitors as clinical candidates surpass that of other viral protease drug targets, and underlies the importance of understanding the structure and mechanism of viral proteases.

4. Conclusion

Viruses utilize encoded proteases as a general strategy to survive with a compact genome. The viral proteases are essential to replication and the viral life cycle. Proteases of various viral families are diverse in structure, catalytic mechanism, and substrate preference. The success in targeting the proteases of HIV-1 and HCV proved that viral proteases are tractable drug targets, and simultaneously revealed significance of drug resistance as one of the major limitations of current approaches in drug design against rapidly evolving targets. The substrate envelope model provides a tool to aid in overcoming this limitation. The strategies and lessons of the past will be valuable in developing protease inhibitors as drugs against emerging viral therapeutic targets.

Acknowledgments

Research in the Schiffer lab on viral proteases is supported by NIH grants R01-GM135919, R21-AI149716, and R01-AI085051. JZ is supported by F3-GM131635.

References

[1] A.K. Ghosh, H.L. Osswald, G. Prato, Recent progress in the development of HIV-1 protease inhibitors for the treatment of HIV/AIDS, J. Med. Chem. 59 (11) (2016) 5172–5208.
[2] A. Brik, C.H. Wong, HIV-1 protease: mechanism and drug discovery, Org. Biomol. Chem. 1 (1) (2003) 5–14.
[3] J. Anderson, C. Schiffer, S.K. Lee, R. Swanstrom, Viral protease inhibitors, Handb. Exp. Pharmacol. 189 (2009) 85–110.
[4] M. Prabu-Jeyabalan, E. Nalivaika, C.A. Schiffer, Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal structures of six substrate complexes, Structure 10 (3) (2002) 369–381.
[5] S. Chellappan, V. Kairys, M.X. Fernandes, C. Schiffer, M.K. Gilson, Evaluation of the substrate envelope hypothesis for inhibitors of HIV-1 protease, Proteins 68 (2) (2007) 561–567.
[6] A. Ozen, T. Haliloglu, C.A. Schiffer, Dynamics of preferential substrate recognition in HIV-1 protease: redefining the substrate envelope, J. Mol. Biol. 410 (4) (2011) 726–744.
[7] F. Leidner, N. Kurt Yilmaz, C.A. Schiffer, Deciphering complex mechanisms of resistance and loss of potency through coupled molecular dynamics and machine learning, J. Chem. Theory Comput. 17 (4) (2021) 2054–2064.
[8] A. Özen, C.A. Schiffer, Integrating evolution of drug resistance into drug discovery, in: Structural Biology in Drug Discovery, 2020, pp. 521–543.
[9] M. Henes, K. Kosovrasti, G.J. Lockbaum, F. Leidner, G.S. Nachum, E.A. Nalivaika, D.N.A. Bolon, N. Kurt Yilmaz, C.A. Schiffer, T.W. Whitfield, Molecular determinants of epistasis in HIV-1 protease: elucidating the interdependence of L89V and L90M mutations in resistance, Biochemistry 58 (35) (2019) 3711–3726.
[10] M. Henes, G.J. Lockbaum, K. Kosovrasti, F. Leidner, G.S. Nachum, E.A. Nalivaika, S.K. Lee, E. Spielvogel, S. Zhou, R. Swanstrom, D.N.A. Bolon, N. Kurt Yilmaz, C.A. Schiffer, Picomolar to micromolar: elucidating the role of distal mutations in HIV-1 protease in conferring drug resistance, ACS Chem. Biol. 14 (11) (2019) 2441–2452.
[11] R. Ishima, N. Kurt Yilmaz, C.A. Schiffer, NMR and MD studies combined to elucidate inhibitor and water interactions of HIV-1 protease and their modulations with resistance mutations, J. Biomol. NMR 73 (6–7) (2019) 365–374.
[12] F. Leidner, N. Kurt Yilmaz, J. Paulsen, Y.A. Muller, C.A. Schiffer, Hydration structure and dynamics of inhibitor-bound HIV-1 protease, J. Chem. Theory Comput. 14 (5) (2018) 2784–2796.
[13] D.A. Ragland, E.A. Nalivaika, M.N. Nalam, K.L. Prachanronarong, H. Cao, R.M. Bandaranayake, Y. Cai, N. Kurt-Yilmaz, C.A. Schiffer, Drug resistance conferred by mutations outside the active site through alterations in the dynamic and structural ensemble of HIV-1 protease, J. Am. Chem. Soc. 136 (34) (2014) 11956–11963.
[14] Y. Cai, W. Myint, J.L. Paulsen, C.A. Schiffer, R. Ishima, N. Kurt Yilmaz, Drug resistance mutations Alter dynamics of inhibitor-bound HIV-1 protease, J Chem Theory Comput 10 (8) (2014) 3438–3448.
[15] J.E. Foulkes-Murzycki, C. Rosi, N. Kurt Yilmaz, R.W. Shafer, C.A. Schiffer, Cooperative effects of drug-resistance mutations in the flap region of HIV-1 protease, ACS Chem. Biol. 8 (3) (2013) 513–518.

[16] Y. Cai, N.K. Yilmaz, W. Myint, R. Ishima, C.A. Schiffer, Differential flap dynamics in wild-type and a drug resistant variant of HIV-1 protease revealed by molecular dynamics and NMR relaxation, J. Chem. Theory Comput. 8 (10) (2012) 3452–3462.
[17] N.M. King, M. Prabu-Jeyabalan, R.M. Bandaranayake, M.N. Nalam, E.A. Nalivaika, A. Ozen, T. Haliloglu, N.K. Yilmaz, C.A. Schiffer, Extreme entropy-enthalpy compensation in a drug-resistant variant of HIV-1 protease, ACS Chem. Biol. 7 (9) (2012) 1536–1546.
[18] Y. Cai, C. Schiffer, Decomposing the energetic impact of drug-resistant mutations: the example of HIV-1 protease–DRV binding, in: R. Baron (Ed.), Computational Drug Discovery and Design, Springer New York, New York, NY, 2012, pp. 551–560.
[19] A. Ali, R.M. Bandaranayake, Y. Cai, N.M. King, M. Kolli, S. Mittal, J.F. Murzycki, M.N. Nalam, E.A. Nalivaika, A. Ozen, M.M. Prabu-Jeyabalan, K. Thayer, C.A. Schiffer, Molecular basis for drug resistance in HIV-1 protease, Viruses 2 (11) (2010) 2509–2535.
[20] D.A. Ragland, T.W. Whitfield, S.K. Lee, R. Swanstrom, K.B. Zeldovich, N. Kurt-Yilmaz, C.A. Schiffer, Elucidating the interdependence of drug resistance from combinations of mutations, J. Chem. Theory Comput. 13 (11) (2017) 5671–5682.
[21] N.M. King, M. Prabu-Jeyabalan, E.A. Nalivaika, C.A. Schiffer, Combating susceptibility to drug resistance: lessons from HIV-1 protease, Chem. Biol. 11 (10) (2004) 1333–1338.
[22] A.N. Matthew, F. Leidner, G.J. Lockbaum, M. Henes, J. Zephyr, S. Hou, D.N. Rao, J. Timm, L.N. Rusere, D.A. Ragland, J.L. Paulsen, K. Prachanronarong, D.I. Soumana, E.A. Nalivaika, N. Kurt Yilmaz, A. Ali, C.A. Schiffer, Drug design strategies to avoid resistance in direct-acting antivirals and beyond, Chem. Rev. 121 (6) (2021) 3238–3270.
[23] N.K. Yilmaz, C.A. Schiffer, Drug resistance to HIV-1 protease inhibitors: molecular mechanisms and substrate coevolution, in: D.L. Mayers, et al. (Eds.), Antimicrobial Drug Resistance: Mechanisms of Drug Resistance, Volume 1, Springer International Publishing, Cham, 2017, pp. 535–544.
[24] N. Kurt Yilmaz, R. Swanstrom, C.A. Schiffer, Improving viral protease inhibitors to counter drug resistance, Trends Microbiol. 24 (7) (2016) 547–557.
[25] A. Özen, C.A. Schiffer, Substrate-envelope-guided design of drugs with a high barrier to the evolution of resistance, in: A. Berghuis, et al. (Eds.), Handbook of Antimicrobial Resistance, 149-173, Springer New York, New York, NY, 2017.
[26] M.N. Nalam, A. Ali, M.D. Altman, G.S. Reddy, S. Chellappan, V. Kairys, A. Ozen, H. Cao, M.K. Gilson, B. Tidor, T.M. Rana, C.A. Schiffer, Evaluating the substrate-envelope hypothesis: structural analysis of novel HIV-1 protease inhibitors designed to be robust against drug resistance, J. Virol. 84 (10) (2010) 5368–5378.
[27] L.N. Rusere, G.J. Lockbaum, M. Henes, S.K. Lee, E. Spielvogel, D.N. Rao, K. Kosovrasti, E.A. Nalivaika, R. Swanstrom, N. Kurt Yilmaz, C.A. Schiffer, A. Ali, Structural analysis of potent hybrid HIV-1 protease inhibitors containing Bis-tetrahydrofuran in a Pseudosymmetric dipeptide Isostere, J. Med. Chem. 63 (15) (2020) 8296–8313.
[28] L.N. Rusere, G.J. Lockbaum, S.K. Lee, M. Henes, K. Kosovrasti, E. Spielvogel, E.A. Nalivaika, R. Swanstrom, N.K. Yilmaz, C.A. Schiffer, A. Ali, HIV-1 protease inhibitors incorporating Stereochemically defined P2' ligands to optimize hydrogen bonding in the substrate envelope, J. Med. Chem. 62 (17) (2019) 8062–8079.
[29] G.J. Lockbaum, F. Leidner, L.N. Rusere, M. Henes, K. Kosovrasti, G.S. Nachum, E.A. Nalivaika, A. Ali, N.K. Yilmaz, C.A. Schiffer, Structural adaptation of Darunavir analogues against primary mutations in HIV-1 protease, ACS Infect. Dis. 5 (2) (2019) 316–325.

[30] M.N. Nalam, A. Ali, G.S. Reddy, H. Cao, S.G. Anjum, M.D. Altman, N.K. Yilmaz, B. Tidor, T.M. Rana, C.A. Schiffer, Substrate envelope-designed potent HIV-1 protease inhibitors to avoid drug resistance, Chem. Biol. 20 (9) (2013) 1116–1124.
[31] B.D. Lindenbach, C.M. Rice, Molecular biology of flaviviruses, Adv. Virus Res. 59 (2003) 23–61.
[32] R. Bartenschlager, L. Ahlborn-Laake, J. Mous, H. Jacobsen, Nonstructural protein 3 of the hepatitis C virus encodes a serine-type proteinase required for cleavage at the NS3/4 and NS4/5 junctions, J. Virol. 67 (7) (1993) 3835–3844.
[33] M.R. Eckart, M. Selby, F. Masiarz, C. Lee, K. Berger, K. Crawford, C. Kuo, G. Kuo, M. Houghton, Q.L. Choo, The hepatitis C virus encodes a serine protease involved in processing of the putative nonstructural proteins from the viral polyprotein precursor, Biochem. Biophys. Res. Commun. 192 (2) (1993) 399–406.
[34] A. Grakoui, D.W. McCourt, C. Wychowski, S.M. Feinstone, C.M. Rice, Characterization of the hepatitis C virus-encoded serine proteinase: determination of proteinase-dependent polyprotein cleavage sites, J. Virol. 67 (5) (1993) 2832–2843.
[35] A. Grakoui, C. Wychowski, C. Lin, S.M. Feinstone, C.M. Rice, Expression and identification of hepatitis C virus polyprotein cleavage products, J. Virol. 67 (3) (1993) 1385–1395.
[36] M. Hijikata, H. Mizushima, T. Akagi, S. Mori, N. Kakiuchi, N. Kato, T. Tanaka, K. Kimura, K. Shimotohno, Two distinct proteinase activities required for the processing of a putative nonstructural precursor protein of hepatitis C virus, J. Virol. 67 (8) (1993) 4665–4675.
[37] M. Hijikata, H. Mizushima, Y. Tanji, Y. Komoda, Y. Hirowatari, T. Akagi, N. Kato, K. Kimura, K. Shimotohno, Proteolytic processing and membrane association of putative nonstructural proteins of hepatitis C virus, Proc. Natl. Acad. Sci. U. S. A. 90 (22) (1993) 10773–10777.
[38] A. Grakoui, D.W. McCourt, C. Wychowski, S.M. Feinstone, C.M. Rice, A second hepatitis C virus-encoded proteinase, Proc. Natl. Acad. Sci. U. S. A. 90 (22) (1993) 10583–10587.
[39] K.P. Romano, A. Ali, W.E. Royer, C.A. Schiffer, Drug resistance against HCV NS3/4A inhibitors is defined by the balance of substrate recognition versus inhibitor binding, Proc. Natl. Acad. Sci. U. S. A. 107 (49) (2010) 20986–20991.
[40] A. Ozen, W. Sherman, C.A. Schiffer, Improving the resistance profile of hepatitis C NS3/4A inhibitors: dynamic substrate envelope guided design, J. Chem. Theory Comput. 9 (12) (2013) 5693–5705.
[41] L. Hedstrom, Serine protease mechanism and specificity, Chem. Rev. 102 (12) (2002) 4501–4524.
[42] R.A. Copeland, Enzymes : a Practical introduction to structure, mechanism, and data analysis, Wiley–VCH (Wiley–India), New Delhi, India, 2008.
[43] D.I. Soumana, N. Kurt Yilmaz, K.L. Prachanronarong, C. Aydin, A. Ali, C.A. Schiffer, Structural and thermodynamic effects of macrocyclization in HCV NS3/4A inhibitor MK-5172, ACS Chem. Biol. 11 (4) (2016) 900–909.
[44] J. Timm, K. Kosovrasti, M. Henes, F. Leidner, S. Hou, A. Ali, N. Kurt Yilmaz, C.A. Schiffer, Molecular and structural mechanism of Pan-genotypic HCV NS3/4A protease inhibition by Glecaprevir, ACS Chem. Biol. 15 (2) (2020) 342–352.
[45] A. Ozen, K. Prachanronarong, A.N. Matthew, D.I. Soumana, C.A. Schiffer, Resistance outside the substrate envelope: hepatitis C NS3/4A protease inhibitors, Crit. Rev. Biochem. Mol. Biol. 54 (1) (2019) 11–26.
[46] X. Zhang, Direct anti-HCV agents, Acta Pharm. Sin B 6 (1) (2016) 26–31.
[47] A.N. Matthew, J. Zephyr, C.J. Hill, M. Jahangir, A. Newton, C.J. Petropoulos, W. Huang, N. Kurt-Yilmaz, C.A. Schiffer, A. Ali, Hepatitis C virus NS3/4A protease inhibitors incorporating flexible P2 Quinoxalines target drug resistant viral variants, J. Med. Chem. 60 (13) (2017) 5699–5716.

[48] A.N. Matthew, J. Zephyr, D. Nageswara Rao, M. Henes, W. Kamran, K. Kosovrasti, A.K. Hedger, G.J. Lockbaum, J. Timm, A. Ali, N. Kurt Yilmaz, C.A. Schiffer, Avoiding drug resistance by substrate envelope-guided design: toward potent and robust HCV NS3/4A protease inhibitors, mBio 11 (2) (2020) e00172–20.
[49] K.P. Romano, A. Ali, C. Aydin, D. Soumana, A. Ozen, L.M. Deveau, C. Silver, H. Cao, A. Newton, C.J. Petropoulos, W. Huang, C.A. Schiffer, The molecular basis of drug resistance against hepatitis C virus NS3/4A protease inhibitors, PLoS Pathog. 8 (7) (2012), e1002832.
[50] D.I. Soumana, A. Ali, C.A. Schiffer, Structural analysis of asunaprevir resistance in HCV NS3/4A protease, ACS Chem. Biol. 9 (11) (2014) 2485–2490.
[51] D.I. Soumana, N. Kurt Yilmaz, A. Ali, K.L. Prachanronarong, C.A. Schiffer, Molecular and dynamic mechanism underlying drug resistance in genotype 3 hepatitis C NS3/4A protease, J. Am. Chem. Soc. 138 (36) (2016) 11850–11859.
[52] A. Ali, C. Aydin, R. Gildemeister, K.P. Romano, H. Cao, A. Ozen, D. Soumana, A. Newton, C.J. Petropoulos, W. Huang, C.A. Schiffer, Evaluating the role of macrocycles in the susceptibility of hepatitis C virus NS3/4A protease inhibitors to drug resistance, ACS Chem. Biol. 8 (7) (2013) 1469–1478.
[53] D. Bonsall, S. Black, A.Y. Howe, R. Chase, P. Ingravallo, I. Pak, A. Brown, D.A. Smith, R. Bowden, E. Barnes, Characterization of hepatitis C virus resistance to grazoprevir reveals complex patterns of mutations following on-treatment breakthrough that are not observed at relapse, Infect Drug Resist 11 (2018) 1119–1135.
[54] S. Zeuzem, G.R. Foster, S. Wang, A. Asatryan, E. Gane, J.J. Feld, T. Asselah, M. Bourliere, P.J. Ruane, H. Wedemeyer, S. Pol, R. Flisiak, F. Poordad, W.L. Chuang, C.A. Stedman, S. Flamm, P. Kwo, G.J. Dore, G. Sepulveda-Arzola, S.K. Roberts, R. Soto-Malave, K. Kaita, M. Puoti, J. Vierling, E. Tam, H.E. Vargas, R. Bruck, F. Fuster, S.W. Paik, F. Felizarta, J. Kort, B. Fu, R. Liu, T.I. Ng, T. Pilot-Matias, C.W. Lin, R. Trinh, F.J. Mensa, Glecaprevir-pibrentasvir for 8 or 12 weeks in HCV genotype 1 or 3 infection, N Engl J Med 378 (4) (2018) 354–369.
[55] A.N. Matthew, F. Leidner, A. Newton, C.J. Petropoulos, W. Huang, A. Ali, N. KurtYilmaz, C.A. Schiffer, Molecular mechanism of resistance in a clinically significant double-mutant variant of HCV NS3/4A protease, Structure 26 (10) (2018) 1360–1372 (e5).
[56] L.N. Rusere, A.N. Matthew, G.J. Lockbaum, M. Jahangir, A. Newton, C.J. Petropoulos, W. Huang, N. Kurt Yilmaz, C.A. Schiffer, A. Ali, Quinoxaline-based linear HCV NS3/4A protease inhibitors exhibit potent activity against drug resistant variants, ACS Med. Chem. Lett. 9 (7) (2018) 691–696.
[57] M.Z. Dewan, J.N. Uchihara, K. Terashima, M. Honda, T. Sata, M. Ito, N. Fujii, K. Uozumi, K. Tsukasaki, M. Tomonaga, Y. Kubuki, A. Okayama, M. Toi, N. Mori, N. Yamamoto, Efficient intervention of growth and infiltration of primary adult T-cell leukemia cells by an HIV protease inhibitor, ritonavir, Blood 107 (2) (2006) 716–724.
[58] J. Tozser, I.T. Weber, The protease of human T-cell leukemia virus type-1 is a potential therapeutic target, Curr. Pharm. Des. 13 (12) (2007) 1285–1294.
[59] S.B. Shuker, V.L. Mariani, B.E. Herger, K.J. Dennison, Understanding HTLV-I protease, Chem. Biol. 10 (5) (2003) 373–380.
[60] D. Shibata, R.K. Brynes, A. Rabinowitz, C.A. Hanson, M.L. Slovak, T.J. Spira, P. Gill, Human T-cell lymphotropic virus type I (HTLV-I)-associated adult T-cell leukemia-lymphoma in a patient infected with human immunodeficiency virus type 1 (HIV-1), Ann. Intern. Med. 111 (11) (1989) 871–875.
[61] M. Li, G.S. Laco, M. Jaskolski, J. Rozycki, J. Alexandratos, A. Wlodawer, A. Gustchina, Crystal structure of human T cell leukemia virus protease, a novel target for anticancer drug design, Proc. Natl. Acad. Sci. U. S. A. 102 (51) (2005) 18332–18337.

[62] E. Wattel, J.P. Vartanian, C. Pannetier, S. Wain-Hobson, Clonal expansion of human T-cell leukemia virus type I-infected cells in asymptomatic and symptomatic carriers without malignancy, J. Virol. 69 (5) (1995) 2863–2868.
[63] T. Satoh, M. Li, J.T. Nguyen, Y. Kiso, A. Gustchina, A. Wlodawer, Crystal structures of inhibitor complexes of human T-cell leukemia virus (HTLV-1) protease, J. Mol. Biol. 401 (4) (2010) 626–641.
[64] M. Zhang, J.T. Nguyen, H.O. Kumada, T. Kimura, M. Cheng, Y. Hayashi, Y. Kiso, Locking the two ends of tetrapeptidic HTLV-I protease inhibitors inside the enzyme, Bioorg. Med. Chem. 16 (14) (2008) 6880–6890.
[65] P. Bagossi, J. Kadas, G. Miklossy, P. Boross, I.T. Weber, J. Tozser, Development of a microtiter plate fluorescent assay for inhibition studies on the HTLV-1 and HIV-1 proteinases, J. Virol. Methods 119 (2) (2004) 87–93.
[66] M. Kuhnert, H. Steuber, W.E. Diederich, Structural basis for HTLV-1 protease inhibition by the HIV-1 protease inhibitor indinavir, J. Med. Chem. 57 (14) (2014) 6266–6272.
[67] G.J. Lockbaum, M. Henes, N. Talledge, L.N. Rusere, K. Kosovrasti, E.A. Nalivaika, M. Somasundaran, A. Ali, L.M. Mansky, N. Kurt Yilmaz, C.A. Schiffer, Inhibiting HTLV-1 protease: a viable antiviral target, ACS Chem. Biol. 16 (3) (2021) 529–538.
[68] N.M. Ferguson, Z.M. Cucunuba, I. Dorigatti, G.L. Nedjati-Gilani, C.A. Donnelly, M.G. Basanez, P. Nouvellet, J. Lessler, EPIDEMIOLOGY. Countering the Zika epidemic in Latin America, Science 353 (6297) (2016) 353–354.
[69] H.H. Chang, R.G. Huber, P.J. Bond, Y.H. Grad, D. Camerini, S. Maurer-Stroh, M. Lipsitch, Systematic analysis of protein identity between Zika virus and other arthropod-borne viruses, Bull. World Health Organ. 95 (7) (2017) 517–525I.
[70] P.K. Russell, The Zika pandemic—a perfect storm? PLoS Negl. Trop. Dis. 10 (3) (2016) e0004589.
[71] L. Wang, S.G. Valderramos, A. Wu, S. Ouyang, C. Li, P. Brasil, M. Bonaldo, T. Coates, K. Nielsen-Saines, T. Jiang, R. Aliyari, G. Cheng, From mosquitos to humans: genetic evolution of Zika virus, Cell Host Microbe 19 (5) (2016) 561–565.
[72] J.L. Kyle, E. Harris, Global spread and persistence of dengue, Annu. Rev. Microbiol. 62 (2008) 71–92.
[73] S. Bhatt, P.W. Gething, O.J. Brady, J.P. Messina, A.W. Farlow, C.L. Moyes, J.M. Drake, J.S. Brownstein, A.G. Hoen, O. Sankoh, M.F. Myers, D.B. George, T. Jaenisch, G.R. Wint, C.P. Simmons, T.W. Scott, J.J. Farrar, S.I. Hay, The global distribution and burden of dengue, Nature 496 (7446) (2013) 504–507.
[74] C. Kang, T.H. Keller, D. Luo, Zika virus protease: an antiviral drug target, Trends Microbiol. 25 (10) (2017) 797–808.
[75] C.G. Noble, C.C. Seh, A.T. Chao, P.Y. Shi, Ligand-bound structures of the dengue virus protease reveal the active conformation, J. Virol. 86 (1) (2012) 438–446.
[76] G. Robin, K. Chappell, M.J. Stoermer, S.H. Hu, P.R. Young, D.P. Fairlie, J.L. Martin, Structure of West Nile virus NS3 protease: ligand stabilization of the catalytic conformation, J. Mol. Biol. 385 (5) (2009) 1568–1577.
[77] D. Luo, T. Xu, C. Hunke, G. Gruber, S.G. Vasudevan, J. Lescar, Crystal structure of the NS3 protease-helicase from dengue virus, J. Virol. 82 (1) (2008) 173–183.
[78] P. Erbel, N. Schiering, A. D'Arcy, M. Renatus, M. Kroemer, S.P. Lim, Z. Yin, T.H. Keller, S.G. Vasudevan, U. Hommel, Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus, Nat. Struct. Mol. Biol. 13 (4) (2006) 372–373.
[79] W.W. Phoo, Y. Li, Z. Zhang, M.Y. Lee, Y.R. Loh, Y.B. Tan, E.Y. Ng, J. Lescar, C. Kang, D. Luo, Structure of the NS2B-NS3 protease from Zika virus after self-cleavage, Nat. Commun. 7 (2016) 13410.

[80] W.W. Phoo, Z. Zhang, M. Wirawan, E.J.C. Chew, A.B.L. Chew, J. Kouretova, T. Steinmetzer, D. Luo, Structures of Zika virus NS2B-NS3 protease in complex with peptidomimetic inhibitors, Antiviral Res. 160 (2018) 17–24.
[81] C. Nitsche, S. Holloway, T. Schirmeister, C.D. Klein, Biochemistry and medicinal chemistry of the dengue virus protease, Chem. Rev. 114 (22) (2014) 11348–11381.
[82] Z. Yin, S.J. Patel, W.L. Wang, W.L. Chan, K.R. Ranga Rao, G. Wang, X. Ngew, V. Patel, D. Beer, J.E. Knox, N.L. Ma, C. Ehrhardt, S.P. Lim, S.G. Vasudevan, T.H. Keller, Peptide inhibitors of dengue virus NS3 protease. Part 2: SAR study of tetrapeptide aldehyde inhibitors, Bioorg. Med. Chem. Lett. 16 (1) (2006) 40–43.
[83] Z. Yin, S.J. Patel, W.L. Wang, G. Wang, W.L. Chan, K.R. Rao, J. Alam, D.A. Jeyaraj, X. Ngew, V. Patel, D. Beer, S.P. Lim, S.G. Vasudevan, T.H. Keller, Peptide inhibitors of dengue virus NS3 protease. Part 1: warhead, Bioorg. Med. Chem. Lett. 16 (1) (2006) 36–39.
[84] A. Poulsen, C. Kang, T.H. Keller, Drug design for flavivirus proteases: what are we missing? Curr. Pharm. Des. 20 (21) (2014) 3422–3427.
[85] C. Nitsche, H. Onagi, J.P. Quek, G. Otting, D. Luo, T. Huber, Biocompatible macrocyclization between cysteine and 2-Cyanopyridine generates stable peptide inhibitors, Org. Lett. 21 (12) (2019) 4709–4712.
[86] Y. Li, Z. Zhang, W.W. Phoo, Y.R. Loh, R. Li, H.Y. Yang, A.E. Jansson, J. Hill, T.H. Keller, K. Nacro, D. Luo, C. Kang, Structural insights into the inhibition of Zika virus NS2B-NS3 protease by a small-molecule inhibitor, Structure 26 (4) (2018) 555–564 (e3).
[87] H. Lee, J. Ren, S. Nocadello, A.J. Rice, I. Ojeda, S. Light, G. Minasov, J. Vargas, D. Nagarathnam, W.F. Anderson, M.E. Johnson, Identification of novel small molecule inhibitors against NS2B/NS3 serine protease from Zika virus, Antiviral Res. 139 (2017) 49–58.
[88] Z. Zhang, Y. Li, Y.R. Loh, W.W. Phoo, A.W. Hung, C. Kang, D. Luo, Crystal structure of unlinked NS2B-NS3 protease from Zika virus, Science 354 (6319) (2016) 1597–1600.
[89] J.F. Chan, K.K. Chik, S. Yuan, C.C. Yip, Z. Zhu, K.M. Tee, J.O. Tsang, C.C. Chan, V.K. Poon, G. Lu, A.J. Zhang, K.K. Lai, K.H. Chan, R.Y. Kao, K.Y. Yuen, Novel antiviral activity and mechanism of bromocriptine as a Zika virus NS2B-NS3 protease inhibitor, Antiviral Res. 141 (2017) 29–37.
[90] S. Voss, C. Nitsche, Inhibitors of the Zika virus protease NS2B-NS3, Bioorg. Med. Chem. Lett. 30 (5) (2020) 126965.
[91] M.A. Behnam, C. Nitsche, V. Boldescu, C.D. Klein, The medicinal chemistry of dengue virus, J. Med. Chem. 59 (12) (2016) 5622–5649.
[92] K.H. Lin, E.A. Nalivaika, K.L. Prachanronarong, N.K. Yilmaz, C.A. Schiffer, Dengue protease substrate recognition: binding of the prime side, ACS Infect. Dis. 2 (10) (2016) 734–743.
[93] K.H. Lin, A. Ali, L. Rusere, D.I. Soumana, N. Kurt Yilmaz, C.A. Schiffer, Dengue Virus NS2B/NS3 protease inhibitors exploiting the prime Side, J. Virol. (2017) 91(10).
[94] S.A. Shiryaev, C. Farhy, A. Pinto, C.T. Huang, N. Simonetti, A. Elong Ngono, A. Dewing, S. Shresta, A.B. Pinkerton, P. Cieplak, A.Y. Strongin, A.V. Terskikh, Characterization of the Zika virus two-component NS2B-NS3 protease and structure-assisted identification of allosteric small-molecule antagonists, Antiviral Res. 143 (2017) 218–229.
[95] M. Brecher, Z. Li, B. Liu, J. Zhang, C.A. Koetzner, A. Alifarag, S.A. Jones, Q. Lin, L.D. Kramer, H. Li, A conformational switch high-throughput screening assay and allosteric inhibition of the flavivirus NS2B-NS3 protease, PLoS Pathog. 13 (5) (2017) e1006411.

[96] B. Millies, F. von Hammerstein, A. Gellert, S. Hammerschmidt, F. Barthels, U. Goppel, M. Immerheiser, F. Elgner, N. Jung, M. Basic, C. Kersten, W. Kiefer, J. Bodem, E. Hildt, M. Windbergs, U.A. Hellmich, T. Schirmeister, Proline-based allosteric inhibitors of Zika and dengue virus NS2B/NS3 proteases, J. Med. Chem. 62 (24) (2019) 11359–11382.
[97] L. Lim, M. Dang, A. Roy, J. Kang, J. Song, Curcumin allosterically inhibits the dengue NS2B-NS3 protease by disrupting its active conformation, ACS Omega 5 (40) (2020) 25677–25686.
[98] C. Nitsche, T. Passioura, P. Varava, M.C. Mahawaththa, M.M. Leuthold, C.D. Klein, H. Suga, G. Otting, De novo discovery of nonstandard macrocyclic peptides as noncompetitive inhibitors of the Zika virus NS2B-NS3 protease, ACS Med. Chem. Lett. 10 (2) (2019) 168–174.
[99] M. Yildiz, S. Ghosh, J.A. Bell, W. Sherman, J.A. Hardy, Allosteric inhibition of the NS2B-NS3 protease from dengue virus, ACS Chem. Biol. 8 (12) (2013) 2744–2752.
[100] R. Raut, H. Beesetti, P. Tyagi, I. Khanna, S.K. Jain, V.U. Jeankumar, P. Yogeeswari, D. Sriram, S. Swaminathan, A small molecule inhibitor of dengue virus type 2 protease inhibits the replication of all four dengue virus serotypes in cell culture, Virol. J. 12 (2015) 16.
[101] C. Bodenreider, D. Beer, T.H. Keller, S. Sonntag, D. Wen, L. Yap, Y.H. Yau, S.G. Shochat, D. Huang, T. Zhou, A. Caflisch, X.C. Su, K. Ozawa, G. Otting, S.G. Vasudevan, J. Lescar, S.P. Lim, A fluorescence quenching assay to discriminate between specific and nonspecific inhibitors of dengue virus protease, Anal. Biochem. 395 (2) (2009) 195–204.
[102] Z. Li, M. Brecher, Y.Q. Deng, J. Zhang, S. Sakamuru, B. Liu, R. Huang, C.A. Koetzner, C.A. Allen, S.A. Jones, H. Chen, N.N. Zhang, M. Tian, F. Gao, Q. Lin, N. Banavali, J. Zhou, N. Boles, M. Xia, L.D. Kramer, C.F. Qin, H. Li, Existing drugs as broad-spectrum and potent inhibitors for Zika virus by targeting NS2B-NS3 interaction, Cell Res. 27 (8) (2017) 1046–1064.
[103] A. Kumar, B. Liang, M. Aarthy, S.K. Singh, N. Garg, I.U. Mysorekar, R. Giri, Hydroxychloroquine inhibits Zika virus NS2B-NS3 protease, ACS Omega 3 (12) (2018) 18132–18141.
[104] J. Baggen, H.J. Thibaut, J.R.P.M. Strating, F.J.M. van Kuppeveld, The life cycle of non-polio enteroviruses and how to target it, Nat. Rev. Microbiol. 16 (6) (2018) 368–381.
[105] K. Messacar, T.L. Schreiner, J.A. Maloney, A. Wallace, J. Ludke, M.S. Oberste, W.A. Nix, C.C. Robinson, M.P. Glode, M.J. Abzug, S.R. Dominguez, A cluster of acute flaccid paralysis and cranial nerve dysfunction temporally associated with an outbreak of enterovirus D68 in children in Colorado, USA, Lancet 385 (9978) (2015) 1662–1671.
[106] C.C. Holm-Hansen, S.E. Midgley, T.K. Fischer, Global emergence of enterovirus D68: a systematic review, Lancet Infect. Dis. 16 (5) (2016) e64–e75.
[107] B. Winther, in: H.K. Heggenhougen (Ed.), Rhinoviruses, in International Encyclopedia of Public Health, Academic Press, Oxford, 2008, pp. 577–581.
[108] J. Seipelt, A. Guarne, E. Bergmann, M. James, W. Sommergruber, I. Fita, T. Skern, The structures of picornaviral proteinases, Virus Res. 62 (2) (1999) 159–168.
[109] L. Costenaro, Z. Kaczmarska, C. Arnan, R. Janowski, B. Coutard, M. Sola, A.E. Gorbalenya, H. Norder, B. Canard, M. Coll, Structural basis for antiviral inhibition of the main protease, 3C, from human enterovirus 93, J. Virol. 85 (20) (2011) 10764–10773.

[110] J. Tan, S. George, Y. Kusov, M. Perbandt, S. Anemuller, J.R. Mesters, H. Norder, B. Coutard, C. Lacroix, P. Leyssen, J. Neyts, R. Hilgenfeld, 3C protease of enterovirus 68: structure-based design of Michael acceptor inhibitors and their broad-spectrum antiviral effects against picornaviruses, J. Virol. 87 (8) (2013) 4339–4351.
[111] W. Dai, D. Jochmans, H. Xie, H. Yang, J. Li, H. Su, D. Chang, J. Wang, J. Peng, L. Zhu, Y. Nian, R. Hilgenfeld, H. Jiang, K. Chen, L. Zhang, Y. Xu, J. Neyts, H. Liu, Design, synthesis, and biological evaluation of Peptidomimetic aldehydes as broad-Spectrum inhibitors against enterovirus and SARS-CoV-2, J. Med. Chem. (2021).
[112] F.G. Hayden, R.B. Turner, J.M. Gwaltney, K. Chi-Burris, M. Gersten, P. Hsyu, A.K. Patick, G.J. Smith 3rd, L.S. Zalman, Phase II, randomized, double-blind, placebo-controlled studies of ruprintrivir nasal spray 2-percent suspension for prevention and treatment of experimentally induced rhinovirus colds in healthy volunteers, Antimicrob. Agents Chemother. 47 (12) (2003) 3907–3916.
[113] A.K. Patick, M.A. Brothers, F. Maldonado, S. Binford, O. Maldonado, S. Fuhrman, A. Petersen, G.J. Smith 3rd, L.S. Zalman, L.A. Burns-Naas, J.Q. Tran, In vitro antiviral activity and single-dose pharmacokinetics in humans of a novel, orally bioavailable inhibitor of human rhinovirus 3C protease, Antimicrob. Agents Chemother. 49 (6) (2005) 2267–2275.
[114] G. Lu, J. Qi, Z. Chen, X. Xu, F. Gao, D. Lin, W. Qian, H. Liu, H. Jiang, J. Yan, G.F. Gao, Enterovirus 71 and coxsackievirus A16 3C proteases: binding to rupintrivir and their substrates and anti-hand, foot, and mouth disease virus drug design, J. Virol. 85 (19) (2011) 10319–10331.
[115] J. Wang, T. Fan, X. Yao, Z. Wu, L. Guo, X. Lei, J. Wang, M. Wang, Q. Jin, S. Cui, Crystal structures of enterovirus 71 3C protease complexed with rupintrivir reveal the roles of catalytically important residues, J. Virol. 85 (19) (2011) 10021–10030.
[116] Y. Guo, Y. Wang, L. Cao, P. Wang, J. Qing, Q. Zheng, L. Shang, Z. Yin, Y. Sun, A conserved inhibitory mechanism of a Lycorine derivative against enterovirus and hepatitis C virus, Antimicrob. Agents Chemother. 60 (2) (2016) 913–924.
[117] E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in real time, Lancet Infect. Dis. 20 (5) (2020) 533–534.
[118] Y. Yang, F. Peng, R. Wang, M. Yange, K. Guan, T. Jiang, G. Xu, J. Sun, C. Chang, The deadly coronaviruses: the 2003 SARS pandemic and the 2020 novel coronavirus epidemic in China, J. Autoimmun. 109 (2020) 102434.
[119] V.G. da Costa, M.L. Moreli, M.V. Saivish, The emergence of SARS, MERS and novel SARS-2 coronaviruses in the 21st century, Arch. Virol. 165 (7) (2020) 1517–1526.
[120] C. Wang, P.W. Horby, F.G. Hayden, G.F. Gao, A novel coronavirus outbreak of global health concern, Lancet 395 (10223) (2020) 470–473.
[121] J. Ziebuhr, Molecular biology of severe acute respiratory syndrome coronavirus, Curr. Opin. Microbiol. 7 (4) (2004) 412–419.
[122] C. Ceraolo, F.M. Giorgi, Genomic variance of the 2019-nCoV coronavirus, J. Med. Virol. 92 (5) (2020) 522–528.
[123] S. Ullrich, C. Nitsche, The SARS-CoV-2 main protease as drug target, Bioorg. Med. Chem. Lett. 30 (17) (2020) 127377.
[124] V. Grum-Tokars, K. Ratia, A. Begaye, S.C. Baker, A.D. Mesecar, Evaluating the 3C-like protease activity of SARS-coronavirus: recommendations for standardized assays for drug discovery, Virus Res. 133 (1) (2008) 63–73.
[125] N. Zhong, S. Zhang, P. Zou, J. Chen, X. Kang, Z. Li, C. Liang, C. Jin, B. Xia, Without its N-finger, the main protease of severe acute respiratory syndrome coronavirus can form a novel dimer through its C-terminal domain, J. Virol. 82 (9) (2008) 4227–4234.

[126] B. Xia, X. Kang, Activation and maturation of SARS-CoV main protease, Protein Cell 2 (4) (2011) 282–290.
[127] B.L. Ho, S.C. Cheng, L. Shi, T.Y. Wang, K.I. Ho, C.Y. Chou, Critical assessment of the important residues involved in the dimerization and catalysis of MERS coronavirus Main protease, PLoS One 10 (12) (2015) e0144865.
[128] T. Pillaiyar, M. Manickam, V. Namasivayam, Y. Hayashi, S.H. Jung, An overview of severe acute respiratory syndrome-coronavirus (SARS-CoV) 3CL protease inhibitors: peptidomimetics and small molecule chemotherapy, J. Med. Chem. 59 (14) (2016) 6595–6628.
[129] C. Huang, P. Wei, K. Fan, Y. Liu, L. Lai, 3C-like proteinase from SARS coronavirus catalyzes substrate hydrolysis by a general base mechanism, Biochemistry 43 (15) (2004) 4568–4574.
[130] T. Muramatsu, C. Takemoto, Y.T. Kim, H. Wang, W. Nishii, T. Terada, M. Shirouzu, S. Yokoyama, SARS-CoV 3CL protease cleaves its C-terminal autoprocessing site by novel subsite cooperativity, Proc. Natl. Acad. Sci. U. S. A. 113 (46) (2016) 12997–13002.
[131] J. Lee, L.J. Worrall, M. Vuckovic, F.I. Rosell, F. Gentile, A.T. Ton, N.A. Caveney, F. Ban, A. Cherkasov, M. Paetzel, N.C.J. Strynadka, Crystallographic structure of wild-type SARS-CoV-2 main protease acyl-enzyme intermediate with physiological C-terminal autoprocessing site, Nat. Commun. 11 (1) (2020) 5877.
[132] X. Xue, H. Yu, H. Yang, F. Xue, Z. Wu, W. Shen, J. Li, Z. Zhou, Y. Ding, Q. Zhao, X.C. Zhang, M. Liao, M. Bartlam, Z. Rao, Structures of two coronavirus main proteases: implications for substrate binding and antiviral drug design, J. Virol. 82 (5) (2008) 2515–2527.
[133] S.A. Amin, S. Banerjee, K. Ghosh, S. Gayen, T. Jha, Protease targeted COVID-19 drug discovery and its challenges: insight into viral main protease (Mpro) and papain-like protease (PLpro) inhibitors, Bioorg. Med. Chem. 29 (2021) 115860.
[134] R. Liang, L. Wang, N. Zhang, X. Deng, M. Su, Y. Su, L. Hu, C. He, T. Ying, S. Jiang, F. Yu, Development of small-molecule MERS-CoV inhibitors, Viruses 10 (12) (2018) 721.
[135] R.L. Hoffman, R.S. Kania, M.A. Brothers, J.F. Davies, R.A. Ferre, K.S. Gajiwala, M. He, R.J. Hogan, K. Kozminski, L.Y. Li, J.W. Lockner, J. Lou, M.T. Marra, L.J. Mitchell Jr., B.W. Murray, J.A. Nieman, S. Noell, S.P. Planken, T. Rowe, K. Ryan, G.J. Smith 3rd, J.E. Solowiej, C.M. Steppan, B. Taggart, Discovery of ketone-based covalent inhibitors of coronavirus 3CL proteases for the potential therapeutic treatment of COVID-19, J. Med. Chem. 63 (21) (2020) 12725–12747.
[136] B. Boras, R.M. Jones, B.J. Anson, D. Arenson, L. Aschenbrenner, M.A. Bakowski, N. Beutler, J. Binder, E. Chen, H. Eng, J. Hammond, R. Hoffman, E.P. Kadar, R. Kania, E. Kimoto, M.G. Kirkpatrick, L. Lanyon, E.K. Lendy, J.R. Lillis, S.A. Luthra, C. Ma, S. Noell, R.S. Obach, M.N. O'Brien, R. O'Connor, K. Ogilvie, D. Owen, M. Pettersson, M.R. Reese, T. Rogers, M.I. Rossulek, J.G. Sathish, C. Steppan, M. Ticehurst, L.W. Updyke, Y. Zhu, J. Wang, A.K. Chatterjee, A.D. Mesecar, A.S. Anderson, C. Allerton, Discovery of a novel inhibitor of coronavirus 3CL protease as a clinical candidate for the potential treatment of COVID-19, BioRxiv (2020) 2020.09.12.293498.