

Molecular mechanisms of protease precursor autoprocessing of RNA viruses: a comprehensive review

Received: 20 November 2025

Accepted: 19 February 2026

Published online: 04 March 2026

Cite this article as: Mótyán J.A., Golda M., Mahdi M. *et al.* Molecular mechanisms of protease precursor autoprocessing of RNA viruses: a comprehensive review. *Viral J* (2026). <https://doi.org/10.1186/s12985-026-03119-z>

János András Mótyán, Mária Golda, Mohamed Mahdi, Nashaat T. Nashed, John M. Louis & József Tózsér

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

Review

Molecular mechanisms of protease precursor autoprocessing of RNA viruses: a comprehensive review

János András Mótýán^{1,*}, Mária Golda¹, Mohamed Mahdi¹, Nashaat T. Nashed², John M. Louis², József Tózsér¹

¹ Department of Biochemistry and Molecular Biology, Faculty of Medicine, University of Debrecen, Debrecen, Hungary; motyjan.janos@med.unideb.hu (J.A.M.), golda.maria@med.unideb.hu (M.G.), mohamed@med.unideb.hu (M.M.), tozser@med.unideb.hu (J.T.)

² Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, United States; nash.nashed@nih.gov (N.T.N.), johnl@nidk.nih.gov (J.M.L.)

* Correspondence: János András Mótýán (motyjan.janos@med.unideb.hu)

Abstract:

Many viruses express their proteins in the form of large polyproteins comprising structural and non-structural (*e.g.* enzymatic) units that are released from the precursor through ordered proteolysis. Proteolytic processing of polyproteins is an indispensable regulatory step for virus maturation and replication that is carried out by the virus-encoded and/or cellular proteases. The activity of a viral protease that is expressed as a part of a polyprotein is controlled in part by the self-cleavage (autoprocessing) from the precursor. The mechanism of protease precursor processing has been established at the molecular level for various RNA virus proteases, including human immunodeficiency virus (HIV) and severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Both viral protease precursors are processed *via* intra- (*in cis*) and intermolecular (*in trans*) cleavages at the N- and C-termini, respectively, yielding the mature enzyme. The remarkably similar activation mechanisms of HIV and SARS-CoV-2 PRs suggest that other viral proteases are activated similarly. In this review, we provide a detailed overview on the protease precursor autoprocessing mechanism of HIV-1 and SARS-CoV-2 proteases and compare those to the activation mechanism of non-viral proteases from their zymogens. Also, we review the activation mechanism of other ss(+)RNA viruses that utilize the polyprotein pathway for their replication. Based on such comparison, it appears that the protease activation mechanisms of most enveloped ss(+)RNA viruses from their precursors share many common features, although they do not correlate directly with the evolutionary relationships, the presence or absence of viral envelope or the catalytic mechanism of the viral protease.

Keywords: RNA virus, protease, protease precursor, autoprocessing, self-cleavage, polyprotein, HIV-1, SARS-CoV-2

Running title: Proteolytic processing of RNA virus protease precursors

1. Introduction

1.1 Viral polyproteins and viral proteases

Viruses are classified based on their morphology, chemical composition and the mode of replication (**Figure 1**). The genomic DNA or RNA is either single (ss) or double stranded (ds), and the single stranded genomes can be either positive or negative sense (ss(+)) and ss(-), respectively). The genomic RNA of the ss(+)RNA viruses (also known as sense-strand) can directly serve as mRNA for the synthesis of viral proteins by the host cell's translation machinery. The genome of ss(-)RNA viruses acts as complementary strand, also referred to as antisense strand, from which a positive-sense strand is synthesized by the viral RNA-dependent RNA polymerase [**Gelderblom, 1996**]. Due to the differences of their compositions, viruses adopt different pathways for protein synthesis. One pathway involves the expression of the proteins as parts of polyproteins. In many cases, the large precursor polyprotein comprises one or more viral proteases (PRs) that subsequently process the polyprotein *via* proteolysis at discrete sites, although, host proteases may also contribute to polyprotein processing [**Rampersad and Tennant, 2018**]. During such processing, the polyproteins are precisely cleaved only at specific sites, releasing the functional units (structural proteins and enzymes) from the large precursor, that ensures proper timing for replication, assembly as well as maturation. The regulation of polyprotein processing is not limited to the cleavage sites, other factors also contribute to the control of proteolysis, *e.g.* i) some proteases need a cofactor, ii) the precursor and mature proteins might exhibit differences in their activity or localization, iii) the sequences of the recognition sites influence the kinetics of proteolytic processing, and thus determine in part the order of the cleavages at the individual sites, and iv) the efficacy of precursor processing is dependent in many cases on the conformational accessibility of the cleavage sites. As discussed later, both precursor and mature forms of the viral PRs contribute to the order of proteolysis of the large polyproteins, including intra- and intermolecular cleavages.

Unlike ss(+)RNA and some DNA viruses, the genomes of other DNA viruses and most ss(-)RNA as well as most dsRNA viruses do not code for a known viral PR [**Majerová and Novotný, 2021**]. The virus-encoded PRs are structurally diverse, but they are well regulated with high sequence specificity for polyprotein processing [**Dougherty and Semler, 1993**] [**Spall et al., 1997**] [**Khan et al., 1999**] [**Tong, 2002**] [**Zephyr et al., 2021**]. PRs make a significant contribution to the replication, at different stages of the viral life-cycle [**Dougherty and Semler, 1993**]. Their function includes not only the cleavage of the viral polyproteins at discrete sites but also performing the cleavages in a proper sequential order, which is essential for virion assembly and maturation, as reviewed previously [**Tözsér and Oroszlan, 2003**] [**Konvalinka et al., 2015**]. There are multiple factors that influence the proteolytic processing, but it is not fully understood how the order of the cleavages is determined by the various factors. Nevertheless, it is known that imprecise cleavages and order lead to aberrant assembly and production of non-infective immature virions [**Wiegers et al., 1998**] [**Pettit et al., 2005**] [**Konvalinka et al., 2015**]. The significance of polyprotein processing lies not only in releasing the protein domains from the precursors, but also in serving as a functional control in time and space, as the precursor and mature forms of the proteins can differ in their functional characteristics, such as catalytic activity and the cleavage site sequence specificity. The cleavages are not limited to the viral proteins knowing that the host proteins are susceptible to inactivation by proteolysis and thereby contributing to viable replication. For example, viruses may evade the host immune reactions by manipulating cellular processes, either by proteolysis-dependent or -independent manner [**Tomasselli and Heinrikson, 1994**] [**Morazzani et al., 2019**] [**Tsu et al., 2021**] [**Chin and Saeed, 2022**] [**Kim and Shan, 2022**] [**Centazzo et al., 2023**] [**Ren**

et al., 2025]. For example, HIV-1 PR can downregulate the level of RIG-I pattern recognition receptor protein that can activate antiviral responses upon recognizing viral ssRNAs [Solis et al., 2011]. In addition, receptor-interacting protein kinases (RIPK1 and RIPK2), that are activated in response to viral infection, are also cleaved by HIV-1 PR, causing their inactivation [Wagner et al., 2015].

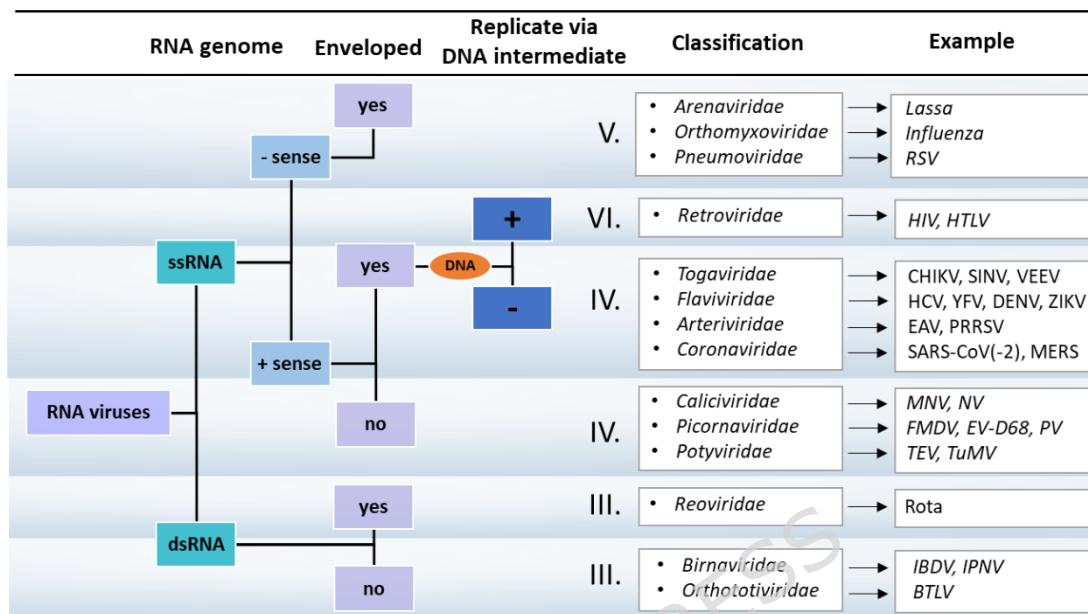


Figure 1. Classification of RNA viruses. The classification of RNA viruses is shown based on their genome and replication strategy (Baltimore classification), most important families and representative examples are shown.

1.2 Processing of polyproteins and activation of viral proteases

The initially synthesized polyproteins of the viruses whose genomes encode one or more PRs [Dougherty and Semler, 1993] [Majerová and Konvalinka, 2022] are considered as the full-length PR precursors. Self-cleavage is defined as autoprocessing, catalysed by the PR embedded within the polyprotein, resulting in the liberation of a fully active mature PR [Katoh et al., 1985] [Oroszlan and Luftig, 1990]. Despite belonging to different structural classes and utilizing different catalytic mechanisms, a common feature of the virus-encoded PRs is their high substrate specificity [Blair and Semler, 1991] [Babé and Craik, 1997] [Spall et al., 1997] [Majerová and Novotný, 2021] [Majerová and Konvalinka, 2022]. Autoprocessing of the polyproteins may be carried out by either unimolecular or bimolecular reactions, *i.e.*, intra- and intermolecular cleavages, respectively, to liberate the mature PR. The two reactions are distinguishable by their kinetic behaviour, irrespective to the involvement of more than one polypeptide chains, such as in the case of homodimerization (*e.g.* HIV-1 and SARS-CoV-2 main protease, MPro) or the contribution of a cofactor. An intramolecular reaction (*in cis* cleavage) displays first-order kinetics in protein concentration of the reacting protein species. The order of the reaction can be determined by observing a linear relationship between the initial rate of conversion and the protein concentration such as in the case of HIV-1 PR where the monomer lacks catalytic activity and the monomer-dimer equilibrium constant is well below the protein concentration in the model system used [Louis et al., 1994] or a constant percent conversion of the initial protein concentration to products in a defined amount of time at various initial protein concentrations, *i.e.* dilution-insensitive, as in the case of SARS-CoV-2 MPro [Aniana et al., 2023] where the

reaction takes place mainly from the monomer. In contrast, the intermolecular reaction (*in trans* cleavage) displays second-order kinetics in protein concentration, i.e. one molecule cleaves another molecule. In this case, the order of the reaction is determined by observing a linear relationship between the rate of cleavage and the concentration of the fully active enzyme, or between the initial rate of the reaction and the square of the protein concentration. It should be noted that observing a dilution-sensitive experiment would rule out only a unimolecular reaction but does not confirm a second-order reaction as the reaction may be of higher order [Dougherty and Semler, 1993] [Louis et al., 1994].

1.3 Methods for the investigation of protease autoprocessing

The full-length polyproteins are not suitable for studying precursor processing *in vitro* because of their large size and the presence of multiple natural cleavage sites. Autoprocessing reactions are typically studied using model polyproteins, which mimic protease precursors. Invariably, they contain the PR domain flanked by native cleavage sites at one or both termini, and in many cases, native protein domains and/or fusion tags. For example, model HIV-1 precursors contained the full-length or truncated trans-frame region (TFR) and/or p6* domain at the N-terminus and a truncated reverse transcriptase (RT) domain at the C-terminus of the PR [Sayer et al., 2012] [Park et al., 2016]. In some cases, fusion tags were also attached to HIV-1 PR precursor mimetics, such as maltose binding protein (MBP) [Louis et al., 1994] [Tien et al., 2018], mCherry and enhanced green fluorescent protein (EGFP) [Humpolíčková et al., 2018], or glutathione S-transferase (GST) and Flag [Huang and Chen, 2010] [Tien et al., 2018]. Similar model precursors have been designed for SARS-CoV-2 MPro [Nashed, Kneller et al., 2022] [Aniana et al., 2023] [Narwal et al., 2023] [Novotný et al., 2025] and papain-like protease (PLpro) [Patchett et al., 2021]. A model precursor of SARS-CoV-2 MPro encompassed MPro and its N- and C-terminal cleavage sites (nsp4/nsp5 and nsp5/nsp6, respectively) [Aniana et al., 2023], and similar mini-precursors were prepared for MPro of SARS-CoV and Middle East respiratory syndrome coronavirus (MERS) [Chen et al., 2010] [Muramatsu et al., 2013] [Huang et al., 2024]. The cleavage assays performed by using model proteins permit monitoring the kinetics of precursor processing *in vitro* or in cell culture-based experiments. Tagging the N- and/or C-termini of a model precursor enables following the reactions at both termini simultaneously as well as the kinetic order of each of these processes.

Besides investigation of model precursors by *in vitro* activity assays, experimental approaches such as X-ray crystallography, nuclear magnetic resonance (NMR) and cryogenic electron microscopy (cryo-EM) (for example HIV-1: 7SJX.pdb [Harrison et al., 2022] and SARS-CoV-2: 8EKE.pdb [Narwal et al., 2023]) have been used to gain insights into events leading to the processing reaction of polyproteins and PR precursors [Tabler and Tilton, 2024]. Structural studies can provide information about the conformations that are unique to the precursor, intermediates in its conversion and the mature forms. Although, polyprotein processing is common to many virus families, sufficiently detailed structural information is available only for a limited number of cases, such as HIV-1 PR's [Ishima et al., 2003] [Chatterjee et al., 2005] [Louis et al., 2007] [Tang et al., 2008] [Sadiq et al., 2012], SARS-CoV [Hsu et al., 2005] and SARS-CoV-2 MPro [Lee et al., 2020] [Noske et al., 2021] being the most thoroughly studied. Furthermore, structure determinations provided insights into the conformational characteristics of PR precursors or proteolytic events such as intra- and/or intermolecular interactions. In addition, direct or indirect lines of evidence have been reported for the PR of the flavivirus Zika virus (ZIKV) [Phoo et al., 2016] and for alphavirus PRs, such as

that of Sindbis virus (SINV) [Shin et al., 2012] and Venezuelan equine encephalitis virus (VEEV) [Compton et al., 2017] [Hoffka et al., 2023].

Attempts to suggest a mechanism of autoprocessing reaction from crystallographic X-ray structure can be misleading as X-ray structures are static conformations. The observation of an N- or a C-terminus bound to the active site of an X-ray structure is no direct evidence of a mode of reaction as both termini are substrates and products for the enzyme, and thus, expected to bind to the active site. Also, any conclusion about the processing reaction, *i.e.*, *in cis* or *in trans* cleavage, based on the X-structure of the mature enzyme ignore the possibility of conformation changes that may occur upon the autoprocessing reaction, see for example the discussion below for HIV-1 PR. Alternatively, molecular dynamics can be utilized to investigate the conformational characteristics of the precursor and mature forms and provide insights into the dynamics and conformational rearrangements of the self-activation reactions [Sadiq et al., 2012] [Kimura et al., 2014] [Hoffka et al., 2023]. The only direct evidence for the mechanism of autoprocessing reaction must come from a kinetic experiment as the order of the reaction reflects the composition of the transition state of the cleavage.

2. Protease precursor processing of RNA viruses

2.1 Enveloped ss(+)RNA viruses - Retroviral proteases - HIV-1 PR

HIV, the causative agent of acquired immunodeficiency syndrome (AIDS) belongs to *Retroviridae* family of enveloped ss(+)RNA viruses (Figure 1). Retroviruses are unique among RNA viruses as they replicate *via* a DNA intermediate. During their replication, their RNA genome is reverse transcribed to DNA that is subsequently integrated into the host cell's DNA by the viral integrase. The integrated provirus then serves as a template for the synthesis of the new viral RNA and proteins [Coffin et al., 1997]. HIV viruses utilize the polyprotein pathway for their replication; their genomes code for two large polyproteins, the Gag and the Gag-Pro-Pol, of which the larger Gag-Pro-Pol is synthesized *via* ribosomal frameshifting. Genomes of HIV-1 and HIV-2 also encode a single copy of an aspartic PR that is synthesized as a part of the Gag-Pro-Pol polyprotein (Figure 2A). HIV PRs are homodimeric having one active site per dimer in which each of the subunit contributes half of the catalytic residues. HIV PRs autocatalyze their own cleavage for activation from the polyproteins (Figure 2B), as described below.

The mechanism of HIV-1 PR's activation has been studied extensively using model polyproteins that comprised the PR domain and flanking sequences that limit the number of the PR cleavage sites at the N and/or C-termini. Model polyproteins containing the trans-frame region (TFR) peptide (TFP) and/or p6^{pol} attached to the N-terminus of PR were found to undergo cleavage at the p6^{pol}/PR site accompanied by a large increase in catalytic activity in a time-dependent manner. The reactions were monitored by following the appearance of the catalytic activity and the mature enzyme as well as the disappearance of the precursor. They display first-order kinetics in protein concentration indicating that the cleavage is an intramolecular reaction (Figure 2D). A polyprotein containing the octapeptide (TFP) followed by p6^{pol} and the PR domain having the natural cleavage sites TFP/p6^{pol} and p6^{pol}/PR (TFP-p6^{pol}-PR) undergoes two intramolecular cleavage steps sequentially below pH 5 and a single cleavage at the N-terminus above pH 5 accompanied by a large increase in catalytic activity [Louis et al., 1998] [Louis et al., 1999a]. Both TFP-p6^{pol}-PR and p6^{pol}-PR have low catalytic activities. Since the monomeric HIV-1 PR has half of the catalytic residues of the active site, dimer formation must occur prior to the intramolecular cleavage [Louis et al., 1994] [Louis et al., 1999a] [Louis et al., 1999b] [Louis et al., 2000] [Louis et al., 2007]. In contrast, the PR containing the C-terminal ΔPol (Δ denotes truncated) sequence has catalytic activity comparable to that of the mature enzyme, and the C-terminal cleavage displays

second-order kinetics in protein concentration, i.e. intermolecular reaction [Wondrak et al., 1996]. While the intramolecular cleavage at the N-terminus with native flanking sequence is a prerequisite to produce a fully active protease, other cleavages upstream from this site may occur intramolecularly in the full-length polyprotein in addition to the TFP/p6^{pol} cleavage [Louis et al., 1999a] [Pettit et al., 2003] [Pettit et al., 2004] [Pettit et al., 2005] [Majerová and Konvalinka, 2022].

The mechanism of activation of the homodimeric HIV-1 PR from the polyprotein precursor (**Figure 2D**) is nearly identical to that for the activation of the single polypeptide chain of the non-viral aspartic proteases. The main difference is the requirement of the viral PR to form a homodimer and generate a competent active site (see the structure of the mature enzyme **Figures 2B and 2C**). HIV-1 PR truncated at the N-terminus (residues 1-4) and the C-terminus (residues 96-99) exhibits a stable monomer fold spanning residues 10-90 like that of the single subunit of the wild-type dimer [Ishima et al., 2003] [Ishima et al., 2007]. An enzymatically inactive mini-precursor consisting of four residues of the trans-frame region fused to the N-terminus of the inactive PR domain is shown to form a small (3-5%) population of dimeric species by paramagnetic relaxation enhancement NMR [Tang et al., 2008]. Model polyprotein comprising the PR domain displays similar characteristics to the non-viral aspartic PR zymogen such as low enzymatic activity, binding inhibitors, and the intramolecular autocleavage at its N-terminus indicating that a fully dimeric polyprotein has a functional active site, see studies on TFP-p6^{pol}-PR model HIV-1 polyprotein in the previous paragraph [Louis et al., 1999a]. Another similarity is found in the amino acid sequence upstream from the viral PR domain: while the amino acid sequence upstream from the viral PR polyprotein, i.e. the 56 amino acid residues of TFP-p6^{pol}, has little sequence homology to the positively charged pro-sequences of the non-viral aspartic PRs (11-14 Lys and Arg residues), it comprises 8 Lys and Arg residues (**Figure 2E**). Given the facts that the liberation of the N-terminus of the mature PR is generated by autocatalytic reaction and the location of the N-terminus of the mature PR on the opposite side of the enzyme's active site participating in an extended β -sheet (**Figure 2C**), the cleavage at the N-terminus of HIV-1 PR must be accompanied by similar conformational changes to those observed in the non-viral enzymes to open the active site and make it available for substrates [Khan and James, 1998] [Khan et al., 1999] [Lee et al., 2020]. This conclusion is further supported by the study utilizing a model polyprotein having non-native sequences flanking the N-terminus of the PR [Wondrak and Louis, 1996]. In contrast to model polyproteins comprising native cleavage sites at the N-terminus, constructs having non-native cleavage sites (X₂₈-PR and X₂₈-TFR-Ala-PR-Ala-Pol, where X₂₈ is a random amino acid sequence) do not autoprocess at the N-terminus. They display enzymatic activities comparable to the mature enzyme and increased dimer dissociation constant. This indicates that a free terminal residue is not required to organize the active site for enzymatic activity, but it is required for stabilizing the dimer form of the enzyme. That raises the question: why does a polyprotein with native flanking sequence have low enzymatic activity? The only plausible answer is that the upstream sequence in the HIV-1 polyprotein is interacting with the active site residues and blocks the active site in a similar manner to that of the pro-sequence of the non-viral aspartic acid PRs [Khan et al., 1999]. This conclusion is further supported by the observed inhibition of the mature protease by the trans-frame peptide (TFP) and p6^{pol} containing its C-terminal residues [Louis et al., 1998] [Paulus et al., 1999]. Recent studies have also indicated that specific amino acid substitutions in the TFR significantly impair virus infectivity [Yu et al., 2022].

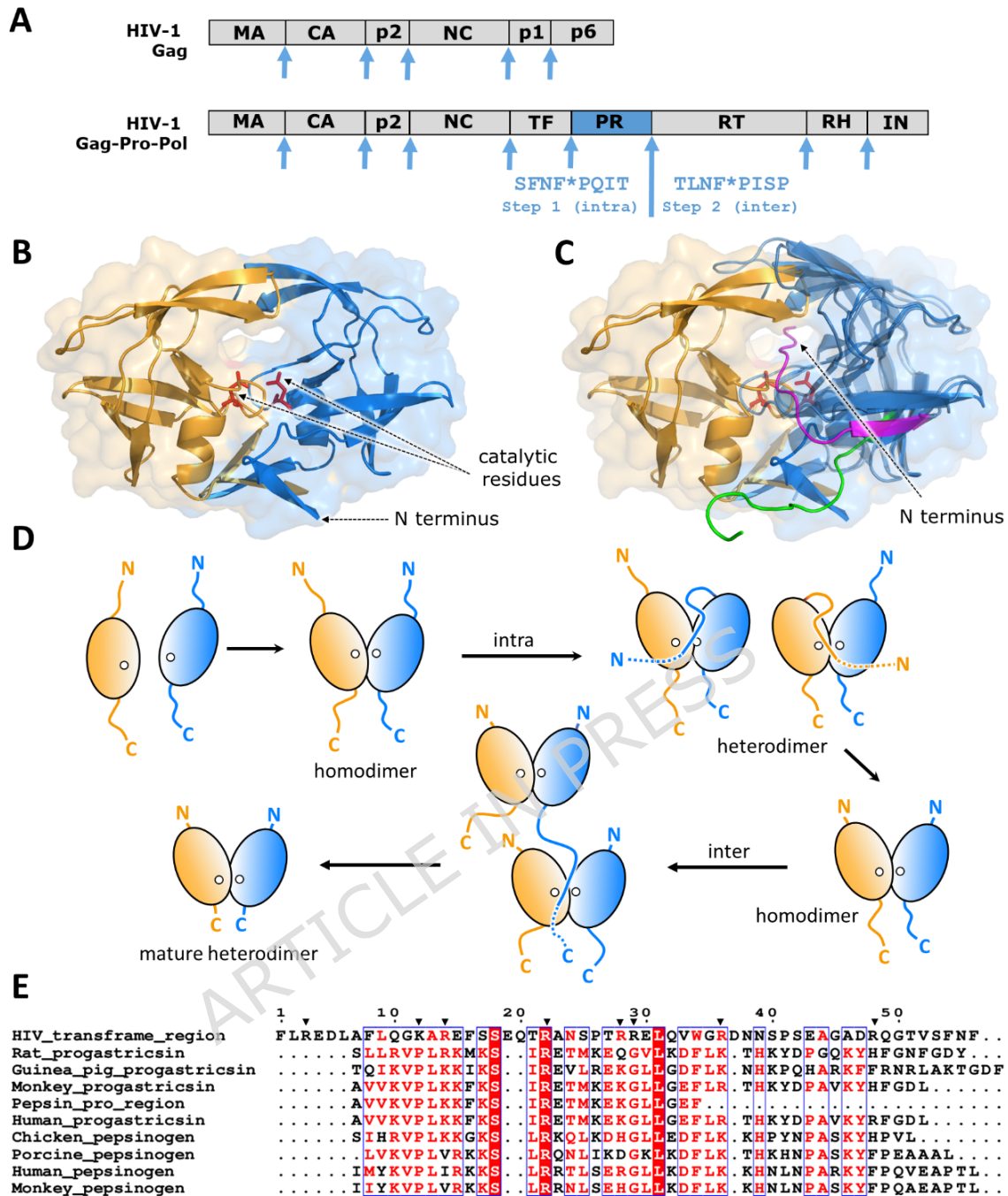


Figure 2. Structural representation of *in cis* N-terminal autoprocessing of HIV-1 PR. (A) Organization of HIV-1 Gag and Gag-Pro-Pol polyproteins. The PR is highlighted in blue along with the cleavage sites (arrows) and their sequences, the sequential order of cleavages is labeled. (B) Ribbon representation based of the crystal structure of mature HIV-1 PR (PDB ID: 5HVP) [Fitzgerald et al., 1990]. (C) The conformational states of the N-terminal cleavage site of the precursor are shown based on structures determined by NMR. N-terminal residues are shown in green and purple based on the solution NMR structure (PDB ID: 1Q9P) [Ishima et al., 2003] and a model prepared from structural coordinates provided by Chun Tang [Tang et al., 2008]. Ribbon/surface representations of the PR subunits are shown in light orange and blue and the catalytic residues are shown as red sticks. (D) Schematic representation of protease precursor autoprocessing *via* dimerization from the viral polyprotein. Polypeptide chains

corresponding to the PR precursor subunits are shown in blue and yellow. The aspartate residues at the active sites of the monomers are indicated by white circles, the catalytic conformation contributed by Asp-Thr-Gly triad is formed upon dimerization. Dashed lines represent the truncation of N and C-termini. (E) 56 residue long trans-frame region (TFR) upstream from HIV-1 PR aligned with the pro-sequences of several non-viral PRs [Khan et al., 1997]. Arrowheads show Lys and Arg residues of HIV-1 TFR. Sequence alignment was performed by using ESPript 3.0 online tool [Robert and Gouet, 2014].

2.2 Enveloped ss(+)RNA viruses - Coronavirus proteases - SARS-CoV and SARS-CoV-2

SARS-CoV, SARS-CoV-2 and MERS are enveloped viruses having ss(+)RNA genome. They belong to Group IV of viruses and are classified into the *Betacoronavirus* genus of the *Coronaviridae* family (Figure 1). The genomic organization of SARS-CoV and SARS-CoV-2 are similar and code for multiple non-structural proteins (nsps), two of them are cysteine PRs (Figure 3A). Nsp3 is a papain-like PR (PLpro), while nsp5 is a 3-chymotrypsin-like PR (3CLpro) and referred to as the main protease (MPro). Both PRs play an important role in processing the viral polyproteins. MPro of SARS-CoV and SARS-CoV-2 share 96% sequence identity and differ only by 12 residues [Lee et al., 2022]. They are expressed embedded in polyproteins from which MPro autocatalyzes its own cleavage to form homodimer (Figure 3B, 3C). MPro consists of three domains: the chymotrypsin-like domain I and II contain the active site of the enzyme while domain III (helical domain) is involved in enhancing dimer stability. The two subunits form an asymmetric interface through their N-terminal residues ("N-finger") and domain III [Nashed et al., 2022b]. Upon dimerization of MPro, the active site oxyanion loop is transitioned from a predominantly unwound inactive conformation (E) in the monomer to a wound active conformation (E*), which has been clearly demonstrated both for SARS-CoV and SARS-CoV-2 MPro [Xia and Kang, 2011] [Kovalevsky et al., 2022] [Nashed et al., 2022a].

Several studies attempted to examine the mechanism of self-maturation by crystallography; N- and C-terminal mutations and X-ray structure analyses resulted in differing conclusions for both SARS-CoV [Hsu et al., 2005] [Li et al., 2010] [Chen et al., 2010] [Muramatsu et al., 2013] and SARS-CoV-2 [Noske et al., 2021]. Recently, model polyproteins of SARS-CoV-2 containing native cleavage site at both termini has revealed sequential order of the cleavages during autoactivation. The first cleavage is at the N-terminus of MPro and its percent cleavage in a window of time is constant and independent from the initial protein concentration of model polyproteins, which indicated that the N-terminal cleavage displays first-order kinetics and the reaction is intramolecular (Figure 3C) [Aniana et al., 2023].

While the mechanism of activation of MPro from the polyprotein differs from the mechanism of activation of the eukaryotic chymotrypsin family of enzymes, the properties of the model polyprotein and the eukaryotic zymogens as well as the structural changes that are observed after cleaving a peptide bond are strikingly similar. The zymogens of the eukaryotic chymotrypsin family of enzymes are single polypeptide chains having low enzymatic activities, bind inhibitors and small molecules indicating that there are two conformers of the zymogen, an active conformation E* and inactive conformer E [Chakraborty et al., 2018]. Similar observations were made in several model dimer polyproteins. Model polyprotein precursors in which key residues such as E290 and R298 in the dimer interface are mutated or the helical domain is completely deleted and containing native flanking sequences at the N- and C-termini have been studied. In all cases, the model polyproteins were fully processed at their N-termini both *in vivo* and *in vitro*, indicating that the PR is folded and having enzymatic activity, and bind inhibitors that are designed for the mature MPro [Nashed et al., 2022a] [Nashed

et al., 2022b] [Kovalevsky et al., 2022] [Kovalevsky et al., 2024]. Model protein of MPro carrying the E290A and R298A (MPro^M) mutations and containing a native flanking sequence at the N-terminus binds the inhibitor GC373 and accelerate the N-terminal cleavage *in vitro* [Aniana et al., 2023]. Such inhibitor-induced acceleration of activity was also observed for the MPro precursor in enzyme kinetic assays [Novotny et al., 2025]. Mature MPro^M displays low enzymatic activity from a small population of the dimer, but the catalytic activity increases significantly upon the addition of an optimal concentration of a known inhibitor of MPro^{WT}. Inhibitor binding to MPro^M stabilizes the dimer form and results in the formation of an increased population of dimer molecules in which one of the active sites is occupied with the inhibitor, and the other is in the proper conformation and available for catalytic activity and thereby resulting in an overall increased catalytic activity [Nashed et al., 2022b]. The mature MPro lacking the dimerization domain III (MPro¹⁻¹⁹⁹) displays very low enzymatic activity only from the monomeric form [Nashed et al., 2022a]. In contrast, the product of the N-terminal cleavage of a model polyprotein containing the C-terminal native flanking sequence or purification tag has comparable enzymatic activity to the mature enzyme. The C-terminal flanking sequence is processed *via* intermolecular reaction mediated by the action of a mature wild-type MPro [Aniana et al., 2023].

The mechanism of activation of the chymotrypsin family of enzymes and the structural changes associated with the generation of a new N-terminus have been reviewed [Khan and James, 1998]. The eukaryotic chymotrypsin family of enzyme are activated by intermolecular cleavage. For example, the activation of chymotrypsinogen is initiated by cleaving the peptide bond Arg15-Ile16 catalysed by trypsin or enterokinase. Comparing the structure of chymotrypsinogen and α -chymotrypsin shows that the newly generate amino terminus Ile16 forms an ion pair with Asp194 triggering the formation of the active enzyme by reorganization of a disordered loop (residues 186-194) in the zymogen to form the oxyanion hole and makes the active site accessible to the substrate. Similarly, comparing the room temperature X-ray structures of monomeric MPro¹⁻¹⁹⁹ with and without the inhibitor GC373 show the protein has a native like fold similar to that of the wild type with the active site dyad Cys145-His41. However, unlike the native mature MPro dimer, the structure of MPro¹⁻¹⁹⁹ shows that the oxyanion loop in an unwound inactive conformation E*, whereas the structure of MPro¹⁻¹⁹⁹-GC373 complex display an oxyanion conformation in the active conformation (E) like that of the native dimer [Nashed et al., 2022a]. The newly generated N-terminal Ser1 of MPro plays a critical role in stabilizing the active E*-state like that observed in eukaryotic chymotrypsin family of enzymes. High resolution X-ray and neutron crystal structures of the native dimer show the protonated α -amino group of the N-terminal Ser1 of one subunit forms an ion pair with the carboxylate side chain of Glu166 of the other subunit of the homodimer, *i.e.* an inter-subunit interactions which stabilizes the active conformation of the active site [Cheng et al., 2010] [Kneller et al., 2020] [Lee et al., 2020]. That intermolecular ionic pairing requires dimer formation for stabilizing the active conformation of the oxyanion loop and to unmask the wild-type activity.

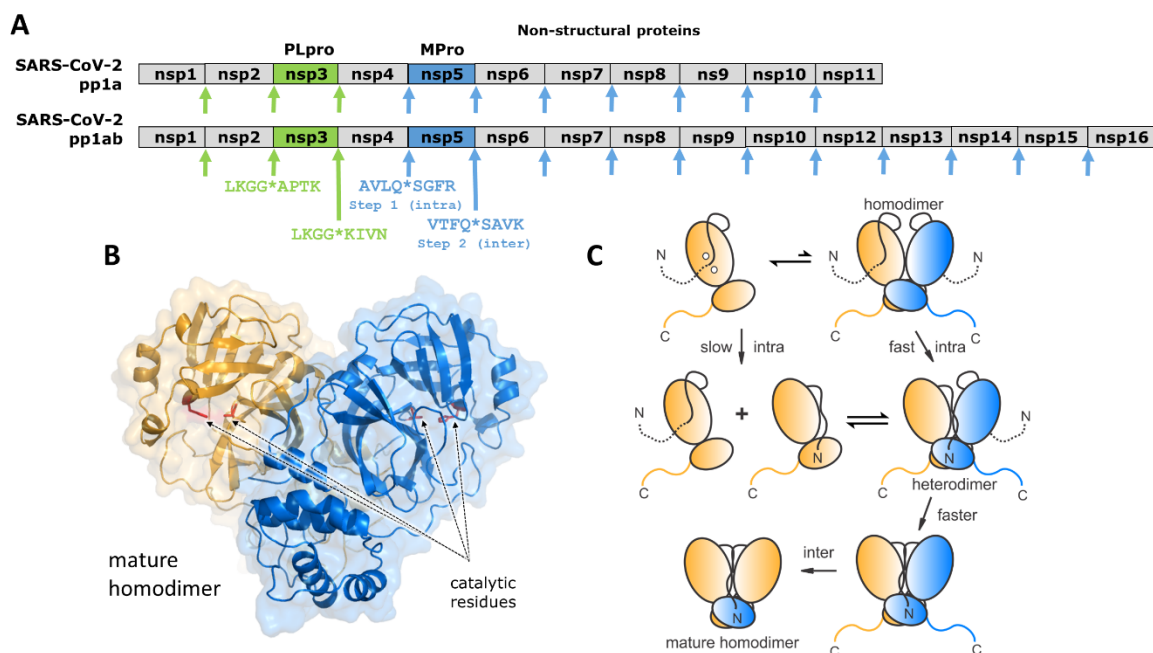


Figure 3. Steps in the autoprocessing of SARS-CoV-2 MPro. (A) Organization of pp1a and pp1ab SARS-CoV-2 polyproteins. PLpro (nsp3) and MPro (nsp5) are highlighted green and blue, respectively. Arrows pointing to the cleavage sites, and their sequences are colored accordingly. (B) Mature MPro crystal structure (PDB ID: 7JOY) [Lee et al., 2020]. The subunits are shown as surface/ribbon representations in light orange and blue. (C) Proposed mechanism of MPro autoprocessing. This figure part was reproduced from Figure 7 of Aniana et al. [Aniana et al., 2023] (<http://creativecommons.org/licenses/by/4.0/>) with recoloring of the MPro domains and additional labels in the legend. Catalytic (domains I and II) and helical (domain III) regions are shown as orange and blue ovals, respectively. Solid black lines denote the N-terminal residues (N-finger) of MPro. Dashed black and solid blue lines represent the truncated nsp4 and nsp6 regions flanking the N and C-termini of MPro, respectively. Catalytic dyad H41 and C145 residues are indicated as white circles in the top left monomer cartoon.

Besides MPro, genomes of coronaviruses code for one or two papain-like PR (PLpro), as well (Figure 3A). Members of *Alphacoronavirus* genus have two PLpro (PL1pro and PL2pro), while many of *Betacoronavirus* genus have only a single PLpro, such as SARS-CoV, SARS-CoV-2 and MERS. PLpro is a C-terminal domain of nsp3 that is a multifunctional cysteine PR. The characteristics of PLpro are well-reviewed in the literature [Báez-Santos et al., 2015] [Lei et al., 2018] [Klemm et al., 2020] [Wong and Saier, 2021] [Mahmoudvand and Shokri, 2021] [Yan and Wu, 2021] [Ullrich and Nitsche, 2022]. SARS-CoV PLpro is active in its monomeric form and one of its main functions is the proteolytic processing of pp1a and pp1ab viral polyproteins at three specific sites to release nsp1, nsp2, and nsp3 [Harcourt et al., 2004]. PLpro recognizes specific cleavage sites with the consensus sequence LXGG↓XX and cleaves the polyprotein to release itself and other non-structural proteins. This is supported by studies showing that PLpro can efficiently process its own N- and C-terminal cleavage sites during polyprotein processing [Barretto et al., 2005]. These sites are susceptible for cleavage *in trans* [Harcourt et al., 2004]. SARS-CoV-2 PLpro was suggested to undergo autoproteolysis [Patchett et al., 2021], based on the ~83% shared sequence identity with highly conserved catalytic triads and similar "thumb-palm-fingers" structures between SARS-CoV and SARS-CoV-2 PLpro [Osipiuk et al., 2021]. Precursor

processing of PLpro was studied recently *in vitro* by dose-dependent transfection and recombinant reporter substrates [Lei et al., 2025]. The studies revealed that the occurrence of the cleavage at nsp2/nsp3 cleavage site at low protein concentration, while the cleavage at nsp3/nsp4 junction was found to require high protein dosage, indicating concentration-dependence, i.e. processing of nsp3/nsp4 cleavage site *in trans*. Intermolecular nature of the cleavage at this site was found to be determined by membrane association of nsp3 as well as its domain rearrangements. In contrast, results suggest that cleavage of nsp2/nsp3 junction can occur both *in cis* and *in trans* [Lei et al., 2025].

2.3 Enveloped ss(+)RNA viruses - Flavivirus proteases

Dengue virus (DENV), Yellow fever virus (YFV), West Nile virus (WNV) and Zika virus (ZIKV) belong to the *Flavivirus* genus of the *Flaviviridae* family that have ss(+)RNA genome (Figure 1). They are mosquito-borne viruses that are known to have public health implications [Braack et al., 2018], for example, ZIKV infection might be associated with congenital foetal abnormalities such as microcephaly [Wahaab et al., 2024].

The RNA genomes of several viruses of *Flavivirus* genus are well characterized and have similar structure [Wahaab et al., 2021]. The genome of a virus encodes a single open reading frame for a large polyprotein comprising structural and non-structural proteins (NS1-5) (Figure 4A). Of the non-structural proteins only NS3 possesses proteolytic activity, a chymotrypsin-like serine PR at the N-terminus of NS3. A transmembrane protein, NS2B, is required for its enzymatic activity and for anchoring NS3 to the endoplasmic reticulum [Zhang et al., 2016]. NS2B and NS3 form a complex that is responsible for polyprotein processing, including cleavages at NS2A/NS2B and NS2B/NS3 sites. The *in vitro* processing of DENV polyprotein, comprising 101-218 residues of NS2A, all NS2B and the first 600 residues of NS3, showed that the cleavage at the NS2A/NS2B site preceded the cleavage at the NS2B/NS3 site [Preugschat et al., 1990]. This study provided indirect evidence that support *cis* cleavages at both sites. Similar results were observed for polyprotein constructs of YFV [Chambers et al., 1990]. NS2B consists of approximately 130 residues and contains three hydrophobic domains believed to be transmembrane and a central hydrophilic domain of only 40-50 residues required as a cofactor (NS2B_{CF}) for catalytic activity of NS3 protease [Wahaab et al., 2021]. Several studies have examined the cleavage at the NS2B/NS3 in detail by tethering NS2B_{CF} to NS3 or only to the protease domain of about 180 amino acid residues of NS3 using a flexible linker and mutation at NS2B/NS3 cleavage site [Bera et al., 2007] [von Hammerstein et al., 2019] [Phoo et al., 2016]. The results of these studies provided further indirect evidence for a *cis* cleavage at the NS2B/NS3 site and indicating that cleavage at the N-terminus of NS3 is not required to organize the active site into catalytically active conformation unlike the eukaryotic chymotrypsin family of enzymes and MPro of SARS-CoV and SARS-CoV2 (see above). This appears to be more like the bacterial chymotrypsin-like α -lytic protease where the active site of the proenzyme has identical conformation to that of the mature enzyme [Khan and James, 1998]. Investigation of DENV NS2B/NS3 showed that the catalytically inactive form of the protein lacks self-processing, suggesting that the NS2B/NS3 cleavage occurs exclusively *in cis*. Nevertheless, co-incubation of active and inactive forms of NS2B/NS3/NS4A revealed that the NS3/NS4A junction is also cleaved intramolecularly [Constant et al., 2018]. NS3 contains an unexpected internal cleavage site, as well [Teo and Wright, 1997], which is also cleaved strictly *in cis*. Interestingly, a mutation of this cleavage site (G459L) was found to abolish cleavage at internal NS3 junction, causing *trans*-dominant suppression of the wild-type virus' replication, an effect driven

by the precursor accumulation rather than by general loss of protease activity. Therefore, inhibiting precursor processing of susceptible viruses is considered to be a promising strategy because the accumulation of uncleaved precursors that act as *trans*-dominant inhibitors can block the replication [Constant et al., 2018]. The inhibition of protease precursor processing is discussed in section 3.

The X-ray structures for tethered and untethered NS2B_{CF}-NS3 complexes have been reported for ZIKV (Figure 4B) [Phoo et al., 2016] [Lei et al., 2016] [Zhang et al., 2016]. The structures are very similar. They show that the N-terminal 18 residues (49-67 of ZIKV) of NS2B_{CF} support the folding of NS3 by forming a β -strand inserted into the protease domain and C-terminal residues 68 to 96 forms a β -hairpin to generate the S2 and S3 pockets in the substrate binding site of NS3. The involvement of the N-terminal residues of NS2B in the folding of NS3, possibly as a chaperon, suggests that the cleavage at NS2A/NS2B which precedes the cleavage at NS2B/NS3 site is required to organize and stabilize the active site conformation. Based on the data discussed above, it is unclear whether the observed activity at the mutated NS2B/NS3 cleavage sites is an artifact of using the combination of NS2B_{CF} and the flexible linker. It is possible for the flexible linker to allow the small fragment NS2B_{CF} to assume the proper interaction with NS3 to produce a functional active site. In contrast, anchoring the full-length NS2B to the endoplasmic reticulum may result in a more rigid conformation that requires the cleavage at the N-terminus of NS3 to allow the proper interaction between NS2B and NS3 to produce a functional active site. In addition, the crystal structure of the helicase domain of YFV suggested that the cleavage at NS3/NS4A junction occurs *in trans*. The C-terminus appeared to be inaccessible for intramolecular cleavage and cannot bind to the active site without substantial structural rearrangement of either the PR or the helicase (PDB ID: 1YKS) [Wu et al., 2005].

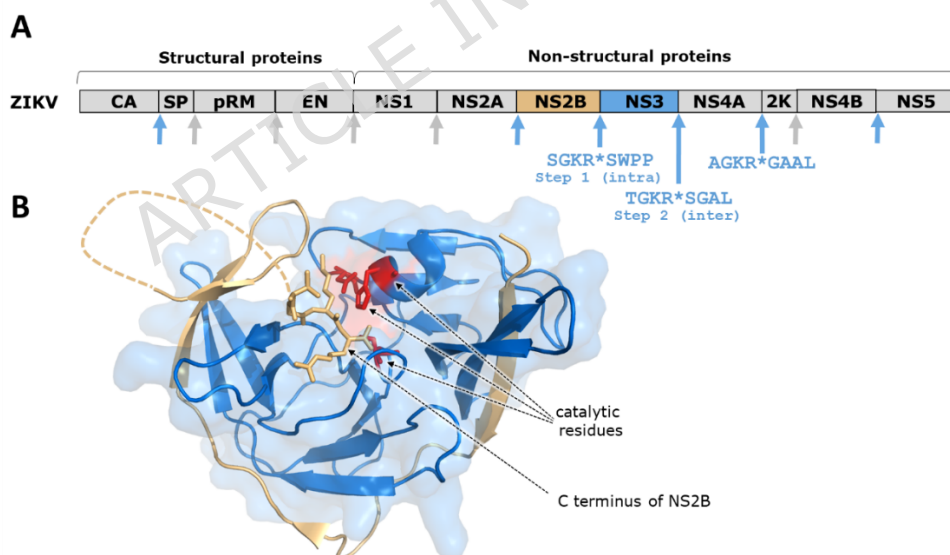


Figure 4. Structure of NS2B-NS3 protease of Zika virus. (A) Organization of ZIKV polyprotein. NS3 PR is highlighted in blue along with the cleavage sites (arrows) and their sequences. The cleavage sites of furin and other host PRs are indicated by grey arrows [Qin et al., 2022]. prM: pre-membrane, EN: envelope. (B) Ribbon representation of NS2B-NS3 (PDB ID: 5GJ4) [Phoo et al., 2016]. Surface representation of NS3 PR domain in light blue. The NS2B domain is shown in light orange and the residues of the catalytic triad (Ser135, His51 and Asp75) by red sticks. The C-terminal residues of NS2B bound to

the active site of NS3 PR are also shown as sticks. The dashed light orange line indicates a region that is not visible in the crystal structure.

Most Flaviviruses seem to use the same mechanism for processing of PR precursor, the NS2B-NS3 cleaves at the NS2B/NS3 junction *in cis* and at the NS3/NS4A site *in trans*, like as been described for YFV [Chambers et al., 1990], WNV [Bera et al., 2007] and DENV [Preugschat et al., 1990]. Interestingly, hepatitis C virus (HCV) that belongs to the *Hepacivirus* genus of *Flaviviridae* virus family exhibits different processing pathway than other flaviviruses such as ZIKV. Two HCV PRs are involved in polyprotein processing, the NS2/NS3 and the NS3 non-structural proteins that have distinct proteolytic activities. The C-terminal domain of NS2 contains a cysteine PR that constitutes a part of the NS2/NS3 PR which is responsible for the cleavage of the NS2/NS3 junction in a zinc-dependent manner. The autolytic cleavage at this site occurs intramolecularly (*in cis*) [Hijikata et al., 1993]. The PR domain of NS3 acts as a serine PR playing a crucial role in the proper folding of NS2/NS3 PR, although, its activity is not required for processing at the NS2/NS3 site [Grakoui et al., 1993]. Rather, NS3 catalyses the cleavages at NS3/NS4A, NS4A/NS4B, NS4B/NS5A, and NS5A/NS5B junctions (Figure 5A). The protease activity of NS3 is complemented *in trans* by the NS4A cofactor that is an amphipathic peptide having hydrophobic N- and hydrophilic C-terminus. A short N-terminal region of NS4A can act as an effector *in trans* even prior to cleavage from the C-terminus of NS3 (Figure 5B) [Failla et al., 1994]. The cleavage at NS3/NS4A junction by NS3 was found to occur only *in cis*, while NS4A/NS4B, NS4B/NS5A, and NS5A/NS5B cleavage sites are processed *in trans*. Accordingly, the NS3 PR is released from the polyprotein precursor *via* cleaving itself at its C-terminus intramolecularly [Tomei et al., 1993] [Failla et al., 1994] [Bartenschlager et al., 1995].

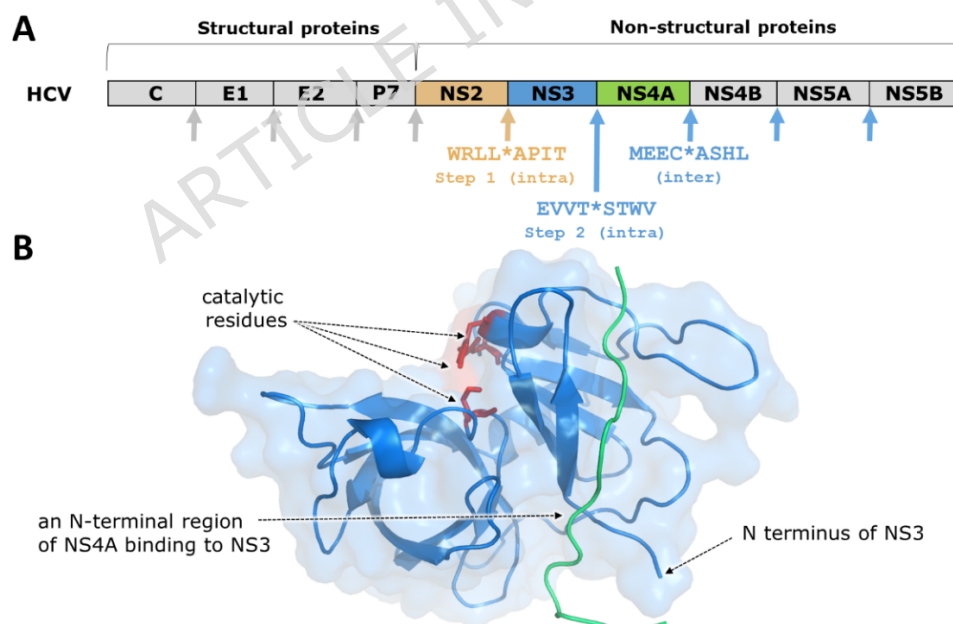


Figure 5. Structure of NS3 protease of Hepatitis C virus complexed with a short N-terminal peptide of its cofactor NS4A. (A) Organization of HCV polyprotein. The NS2 containing a cysteine PR, NS3 containing a serine protease domain and the NS4A cofactor peptide are shown in yellow, blue and green, respectively. Cleavage site sequences are colored accordingly. Cleavage sites of NS2 and NS3 are indicated by yellow and blue arrows, respectively. Cleavage sites of cellular signal peptidases are

indicated by grey arrows. Defined cleavage sites are based on the reference [Neufeldt et al., 2018]. C, E1 and E2 denote the core protein and the envelope glycoproteins E1 and E2, respectively. (B) X-ray structure of HCV NS3 PR complexed with NS4A cofactor peptide (PDB ID: 1A1R) [Kim et al., 1996]. The NS3 PR domain is shown by surface representation in light blue and the NS4A cofactor is green. The catalytic serine, histidine and aspartate residues are shown as red sticks.

2.4 Enveloped ss(+)RNA viruses - Togaviruses - Alphavirus proteases

Alphaviruses are Group IV viruses that have enveloped ss(+)RNA genome and belong to the *Togaviridae* family (Figure 1). They are classified into the group of Old- and New-World alphaviruses and are spread worldwide. The group of Old-World alphaviruses includes Chikungunya virus (CHIKV), Sindbis virus (SINV) and Semliki Forest virus (SFV), while the Venezuelan equine encephalitis virus (VEEV), Eastern Equine Encephalitis Virus (EEEV), and Western Equine Encephalitis Virus (WEEV) are New-World alphaviruses. Although, the two groups evolved independently, they share key features of their life cycles such as pathways of polyprotein processing [Garmashova et al., 2007].

Genomes of alphaviruses encode structural as well as non-structural polyproteins. The non-structural polyprotein (P1234) consisting of four proteins (nsp1, nsp2, nsp3, and nsp4) (Figure 6A) is processed by the C-terminal cysteine PR domain of nsp2, referred to as nsp2pro (Figure 6B) [Bozóki et al., 2020]. Studies on the processing of P1234 polyprotein of SINV [Hardy et al., 1989] and SFV [Vasiljeva et al., 2003] revealed that the cleavage at the nsp1/nsp2 junction liberating the N-terminus of nsp2 occurs exclusively *in cis*. The dilution-sensitive nature of the cleavage at the nsp2/nsp3 junction indicated that processing at this site occurs intermolecularly. Structural studies of a model precursor containing the nsp2 PR and the nsp3 central domains of SINV (PDB ID: 4GUA) revealed that the residues of the nsp2/nsp3 cleavage site are distant from the active site of the nsp2's PR domain (~40 Å), suggesting that processing at the C-terminus of nsp2 PR likely occurs *in trans* [Shin et al., 2012] (Figure 6). In accordance with this, structural analysis of VEEV nsp2pro also revealed that the nsp2/nsp3 cleavage site is not accessible for intramolecular cleavage, as it is distant from the active site (~42 Å) (PDB ID: 2HWK) [Russo et al., 2006]. These data are in agreement with those of *in vitro* analyses which revealed processing of nsp2/nsp3 cleavage site *in trans*, both in the case of SINV [Hardy et al., 1989] and SFV PRs [Vasiljeva et al., 2003].

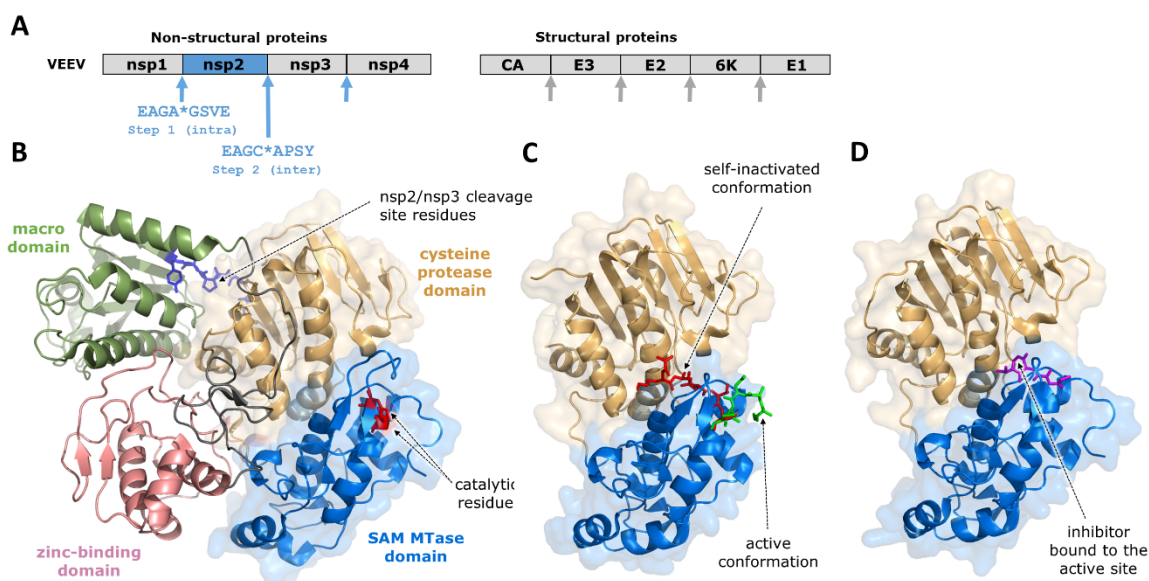


Figure 6. Structure of alphavirus nsP2pro. (A) Organization of VEEV polyproteins. The cleavage sites are indicated based on the report of Hernandez et al. [Hernandez et al., 2014]. The nsP2pro cleavage sites are shown by blue arrows along with their sequences. Other cleavage sites (capsid, furin, signal peptidase) are indicated by grey arrows. (B) Ribbon representation of SINV nsp2/nsp3 precursor (PDB ID: 4GUA) [Shin et al., 2012]. The N-terminal cysteine protease domain of nsP2pro is shown in light blue and the C-terminal S-adenosyl-L-methionine-dependent RNA methyltransferase domain (SAM MTase) in yellow. The macro domain (green), the zinc-binding domain (pink), and the linker between these domains (grey) are shown. The catalytic Cys and His residues are shown as red sticks, and the P4-P4' residues of the nsp2/nsp3 cleavage site (GVGA*APSY) as blue sticks. (C) The active and self-inactivated conformations of the N-terminus of nsP2pro are also shown (PDB ID: 8DUF) [Hoffka et al., 2023]. The N-terminus bound to the active site in the self-inactivated conformation is shown in red, while it is colored green in the active conformation being exposed to the surface. (D) Crystal structure of VEEV nsP2pro complexed with a peptide-like inhibitor (PDB ID: 5EZX) [Hu et al., 2016]. Bound inhibitor to the active site is shown in purple.

Crystallographic analyses of VEEV nsP2pro helped to visualize the reorganization of the N-terminal residues and their reorganization into the active site using a catalytically inactive mutant (N475A) (PDB ID: 6BCM) [Compton et al., 2017] and a surface mutant enzyme (K741A and K767A) (PDB ID: 8DUF) [Hoffka et al., 2023]. In this conformation the interaction of the N-terminus with the active site (Figure 6C) was found to mimic the binding of a natural substrate or a peptide-like inhibitor (Figure 6D). A continuous density observed for the N-terminus of active site- [Compton et al., 2017] and surface-mutant enzymes [Hoffka et al., 2023] implied that binding of these residues to the active site corresponds to a catalytically incompetent conformation. Although the structures representing a self-inactivated state of VEEV nsP2pro do not provide direct evidence for the mechanism of autoactivation *in cis* or *in trans*, the alternate conformations potentially contribute to the regulation of the PR activity [Yost and Marcotrigiano, 2013] [Compton et al., 2017], and the exact molecular mechanism remains to be elucidated.

Based on the available data, PR precursors of Old- and New-World alphaviruses likely utilize the same mechanism, *i.e.* processing of the N-terminus of nsP2 containing nsP2pro

in cis is followed by cleavage of nsp2's C-terminus *in trans*, resembling the cleavage events of HIV-1 PR and SARS-CoV-2 MPro precursor activation.

2.5 Enveloped ss(+)RNA viruses - Arterivirus proteases

The *Arteriviridae* family belongs to the group of enveloped ss(+)RNA viruses (**Figure 1**), employing a polyprotein expression pathway. The prototypic members of this family are the equine arteritis virus (EAV) as well as the porcine reproductive and respiratory syndrome virus (PRRSV), being economic concerns for horse- and swine-breeding industries, respectively. Genome of EAV encodes two large polyproteins, pp1a and pp1ab, that are produced by ribosomal frameshifting and are subsequently processed proteolytically into non-structural proteins by virus-encoded PRs [**Balasuriya et al., 2013**]. Of the non-structural proteins of EAV pp1ab (nsp1-nsp12) there are three distinct PR domains: the non-structural protein 1 and 2 (nsp1 and nsp2, respectively) are papain-like cysteine PRs, while nsp4 is a chymotrypsin-like serine PR being considered as the main PR of EAV [**Snijder et al., 1996**].

Major and minor polyprotein processing pathways have also been established for EAV. Cleavages of the major pathway correspond to highly efficient downstream processing events mediated by nsp4, which release the mature non-structural proteins [**Wassenaar et al., 1997**] [**Barrette-Ng et al., 2002**] [**Fang and Snijder, 2010**]. In this pathway, cleavage of nsp4 occurs first at its C-terminus, the resulting nsp3-4 protein is subsequently cleaved (at its N-terminus) to release the free nsp4 [**Snijder et al., 1994**]. The nsp2 was found to act as a cofactor of nsp4 PR in the major pathway, it is required for processing at nsp4/nsp5 junction, otherwise cleavage does not occur and further processing events occur *via* the minor pathway [**Wassenaar et al., 1997**]. In the minor pathway, nsp4 is cleaved first at its N-terminus (at the nsp3/nsp4 junction) with low efficiency producing a transient or partially processed intermediate, while the nsp4/nsp5 junction remains uncleaved [**Wassenaar et al., 1997**] [**Barrette-Ng et al., 2002**]. Crystallographic analysis implied that EAV nsp4 has alternate conformations which enable processing both N and C-termini *in cis* and contribute to regulating proteolytic activity, *e.g.* by preventing self-inactivation [**Barrette-Ng et al., 2002**]. Precursor of EAV nsp4 PR is processed *via* two alternative cleavage pathways, nsp4 is cleaved first at its C-terminus in the major proteolytic pathway, and cleavage does not occur at its C-terminus in the minor pathway. Accordingly, the mechanism of PR precursor activation of EAV nsp4 does not resemble those of HIV and SARS-CoV-2 PRs.

2.6 Non-enveloped ss(+)RNA viruses - Picornaviruses

Picornaviruses belong to the group of non-enveloped ss(+)RNA viruses (**Figure 1**). The members are classified into multiple genera, such as hepatitis A virus (*Haptovirus* genus), foot-and-mouth disease virus (FMDV) (*Aphthovirus* genus), as well as enteroviruses and rhinoviruses (*Enterovirus* genus) [**Ehrenfeld et al., 2010**]. These viruses also use polyprotein pathway and rely heavily on self-processing PRs to cleave their polyproteins [**Sun et al., 2016**], as exemplified by FMDV which infects cloven-hoofed animals and is the causative agent of foot-and-mouth disease [**Grubman and Baxt, 2004**]. The genome of FMDV contains one open reading frame that encodes a single polyprotein which consists of multiple domains including an autolytic N-terminal protease domain (leader protease, Lpro), capsid proteins (VP1-VP4), a transmembrane protein (3A), 3B peptides (3B1, 3B2 and 3B3), a protease being similar to chymotrypsin-like serine proteases (3Cpro) and an RNA-dependent RNA polymerase (3Dpol) (**Figure 7A**) [**Jamal and Belsham, 2013**].

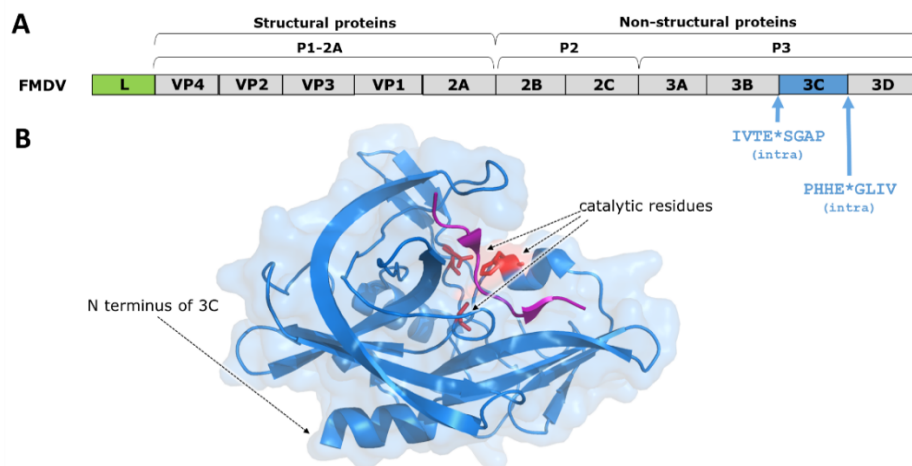


Figure 7. The FMDV polyprotein and the structure of 3C PR. (A) Organization of FMDV polyprotein. N- and C-terminal cleavage site sequences of 3C PR are indicated [Birtley et al., 2005]. (B) Ribbon representation of FMDV 3C PR (PDB ID: 2WV4) [Zunszain et al., 2010]. The enzyme is shown in light blue and the decapeptide that binds to the active site, corresponding to the VP1-2A cleavage site (APAKQ*LLNFD) in purple. Residues of the Cys-His-Asp catalytic triad are shown as red sticks (the catalytic Cys was mutated to Ala in this structure).

Not Lpro but 3Cpro is responsible for the processing of the 2BC-3ABCD polyprotein, it catalyses 10 of the 13 cleavages (Figure 7A) [Curry et al., 2007]. FMDV is considered to use alternative mechanism for processing of the P3 precursor, that can be differentially processed intra- and intermolecularly (i.e. *in cis* and *in trans*), which results in the release of alternative precursors (3ABC and 3CD). The two cleavage pathways are mutually exclusive [Herod et al., 2017] [Pierce et al., 2023]. Studies of picornavirus 3C PRs revealed that they are cleaved at their N-terminus more readily than at the C-terminus, indicating intramolecular cleavage at the N-terminal cleavage site (*in cis*) [Curry et al., 2007]. Similar to the cleavage of FMDV 3Cpro at its N-terminus (3B3/3C junction), processing of the C-terminal cleavage site (3C/3D junction) was also found to occur *in cis* [Herod et al., 2017]. A switch between the *cis* and *trans*-mediated processing was also reported for FMDV, the mature 3Cpro (or the 3Cpro-containing precursor) was found to mediate P3 processing by *in trans* cleavages, which is thought to contribute to regulation of replication *via* appearance of alternative precursors [Herod et al., 2017]. Accordingly, activation mechanism of FMDV 3Cpro precursor is likely different from those of HIV-1 and SARS-CoV-2 PRs.

2.7 Non-enveloped ss(+)RNA viruses - Potyviruses

The Tobacco etch virus (TEV) belongs to the *Potyviridae* family of non-enveloped ss(+)RNA viruses (Figure 1). Polyprotein of TEV consists of 10 protein domains: P1, HC-Pro, P3, 6K1, CI, 6K2, VPg, NIa-Pro, NIb, and coat protein [Adams et al., 2005]. Proteolytic processing of the viral polyprotein is mediated by three virus-encoded PRs: i) the P1-Pro trypsin-like serine PR, ii) the helper component papain-like cysteine PR (HC-Pro) and iii) the NIa-Pro serine-like cysteine PR [Urcuqui-Inchima et al., 2001]. The functional PR being responsible for polyprotein processing is NIa-Pro that is a 49 kDa protein containing a 27 kDa catalytic domain. It is important to note that NIa-Pro is an indispensable tool for recombinant protein technology, it is one of the most widely applied tool of enzymatic fusion tag removal [Kapust et al., 2002] [Waugh, 2011] [Mótyán et al., 2013] [Golda et al., 2024].

During processing of TEV polyprotein, both P1-Pro and HC-Pro cleave themselves at their C-termini *in cis* [Guo et al., 2011], while all other seven sites are processed by N1a-Pro. N1a-Pro was found to exhibit both intra- and intermolecular cleavage activities. Activity assays performed by using wild-type and cleavage site mutant proteins revealed that the release of N1a-Pro from the precursor is carried out by *in cis* cleavages, without the requirement for a specific order of cleavages at its N and C-termini, while the other sites of the polyprotein are processed *in trans* [Carrington and Dougherty, 1987] [Carrington et al., 1988]. Not only self-proteolysis, but also the cleavage of plant host proteins is carried out by N1a-Pro [Xiao et al., 2022]. N1a-Pro of pepper vein banding virus (PVBV) is also cleaved at its termini only *in cis* [Sabharwal et al., 2020], indicating that the mechanism of PR precursor processing of TEV and PVBV is different from those of HIV-1 and SARS-CoV-2.

2.8 Non-enveloped ss(+)RNA viruses - Caliciviridae - Noroviruses

Noroviruses are non-enveloped ss(+)RNA viruses and belong to the *Caliciviridae* family (Figure 1). Their genome encodes a chymotrypsin-like 3C protease, corresponding to non-structural protein 6 (NS6) that is responsible for the processing of the viral polyprotein [Hardy, 2005] [Sosnovtsev et al., 2006] [Emmott et al., 2015] [Campillay-Véliz et al., 2020]. The model of norovirus (NV) polyprotein processing has already been established based on *in vitro* assays and by studying cleavage site mutants [Belliot et al., 2003] [Scheffler et al., 2007]. The NS1/2-NS3-NS4-NS5-NS6-NS7 viral polyprotein (also referred to as Nterm-NTPase-p20-VPg-Pro-Pol) was found to be processed by “early” cleavages to release the NS1/2 domains. Then, the released NS3-7 precursor is subsequently cleaved by NS6 *via* more efficient “late” cleavages which yields the NS6-NS7 (*i.e.* Pro-Pol) precursor containing the PR domain. This cleavage occurs at the N-terminus of NS6 intramolecularly (*in cis*). In agreement with this, the extended conformation of the PR domain was assumed to enable the NS5/NS6 (*i.e.* VPg6Pro) junction to access the active site, making an intramolecular cleavage possible [Someya et al., 2002]. Autoproteolytic activity was not observed for the NS6-NS7 complex, this precursor is processed further only *in trans*, yielding the NS6 and NS7 proteins. Accordingly, the processing of NV PR precursor resembles that of HIV-1 PRs, namely the NV NS6 is also cleaved at its N and C-termini *via* intra- and intermolecular cleavages, respectively. In contrast to HIV-1 PR, the NS6-NS7 NV precursor has higher proteolytic activity than the free NS6 PR. In addition, the “late” cleavage events were found to occur in two possible pathways where a NS4-NS5-NS6 precursor is also formed, in this case the NS6 is cleaved first at its C-terminus [Belliot et al., 2003]. The existence of different cleavage pathways is thought to contribute to the regulation of the viral replication [Campillay-Véliz et al., 2020].

2.9 Non-enveloped dsRNA viruses - Birnaviruses

Birnaviruses belong to the group of non-enveloped ss(+)RNA viruses (Figure 1), the *Birnaviridae* family includes infectious bursal disease virus (IBDV), infectious pancreatic necrosis virus (IPNV), Blotched Snakehead virus (BSNV), Drosophila X and B virus (DXV and DBV, respectively), and Tellina virus 1 (TV-1). These viruses use polyprotein pathway for their replication, the viral genome encodes a non-canonical serine-lysine PR referred to as VP4 (Figure 8A). The investigation of TV-1 revealed that that VP4 PR also undergoes self-proteolysis and TV-1 polyprotein is processed at pVP2/VP4 and VP4/VP3 junctions to release VP4 PR [Chung and Paetzel, 2011] [Delmas et al., 2013].

Neither targeted structural studies nor biochemical assays were performed to study processing of VP4 at its N-terminus, therefore, there are only indirect evidences for the intramolecular cleavage at the VP2/VP4 junction [Delmas et al., 2013]. In contrast,

cleavage at the C-terminus of TV-1 VP4 has been studied in a more detailed manner. A crystallographic analysis of TV-1 VP4 showed the binding of the C-terminus to its own active site (**Figure 8B**), trapping the acyl-enzyme helped to visualize a cleavage event at VP4/VP3 junction that occurs intramolecularly (*in cis*) [**Chung and Paetzel, 2011**]. It is noteworthy that VP4 of IPNV was found to exhibit *in trans* activity, as well. The primary processing of the large IPNV polyprotein occurs in *in cis*, the yielded pre-VP2 capsid precursor (pVP2) is subsequently processed by VP4 at multiple sites *via* bimolecular reactions, *i.e.* conversion of pVP2 to VP2 occurs *in trans* [**Petit et al., 2000**]. In addition, IPNV VP4 was found to undergo cleavage at an internal site, as well, where cleavage also occurs *in trans*. The crystal structure of IPNV VP4 revealed the formation of an acyl-enzyme complex between the residues of the internal cleavage site and the active site of the adjacent enzyme, alluding to an intermolecular cleavage event (**Figure 8C**) [**Lee et al, 2007**].

Although, self-processing at the internal cleavage site *in trans* has been described for IPNV VP4, the activation mechanism of the PR precursor is not identical with those of HIV-1 and SARS-CoV-2, because self-cleavages at the N- and C-terminal boundaries of birnavirus VP4 likely occur only *in cis*.

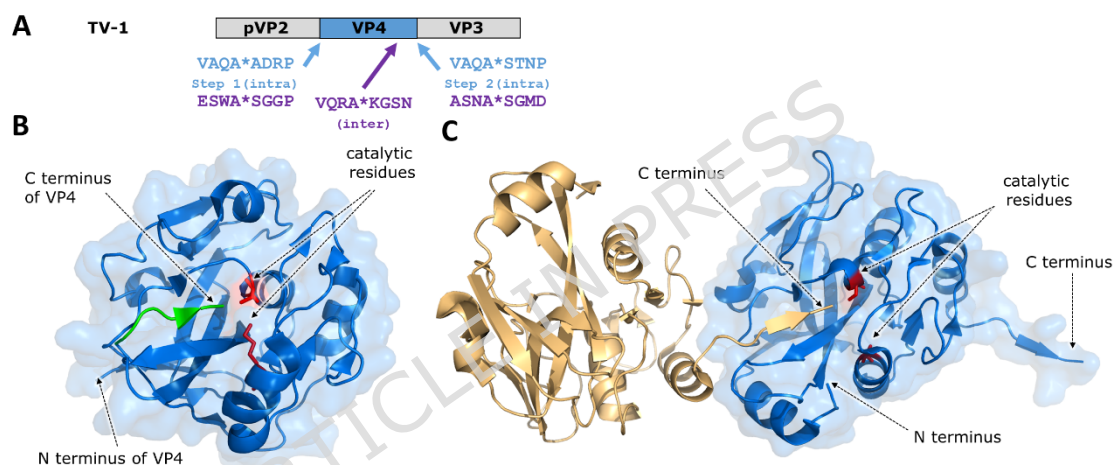


Figure 8. VP4 protease of birnavirus TV-1 and IPNV. (A) Organization of TV-1 polyprotein. The VP4 PR domain and sites of self-cleavage of TV-1 VP4 are shown in blue [**Chung and Paetzel, 2011**]. In addition to terminal sites, an internal cleavage site sequence is also shown for IPNV VP4 (VQRA*KGSN), IPNV sites are shown in purple [**Lee et al, 2007**]. (B) Ribbon model of TV-1 VP4 (PDB ID: 3P06) [**Chung and Paetzel, 2011**]. VP4 is blue, its C-terminal residues bound to the active site are green, while the catalytic residues (sticks) are red. (C) Intermolecular contacts between two (light orange and blue) active site mutant (Lys-to-Ala) VP4 PRs of IPNV (PDB ID: 2PNL) [**Lee et al, 2007**].

3. Inhibition of protease precursor processing

3.1 Autoprocessing as target for antiviral therapies

The proteolytic processing of the viral polyprotein is indispensable for the maturation of viral proteins, therefore, the PR(s) encoded by the viral genomes are considered as crucial targets of antiviral therapies. The identification and rational design of broad-spectrum PIs is of a special interest [**Kim et al., 2012**] [**Bafna et al., 2022**]. Since PIs of mature PRs are well-reviewed in the literature [**Gupta, 2018**] [**Zephyr et al., 2021**] [**Majerová and Konvalinka, 2022**] [**Borges et al., 2024**], therefore, in this review only the PIs targeting PR precursors are discussed.

Targeting the precursor with high affinity binders rather than mature PRs may offer considerable advantages for inhibiting the viral replication cycle. Inhibiting the early stages of the life-cycle can reduce the risk of resistance mutations, while the accumulation of uncleaved precursors may trans-dominantly inhibit the production of mature virions [Majerová and Novotný, 2021] [Majerová and Konvalinka, 2022] [Huang et al., 2024] [Novotný et al., 2025]. For example, PIs of the precursor forms of HIV-1 PR may potentially prevent the stabilization of PR dimers with mature trans-activity [Louis et al., 2011] [Park et al., 2016] [Huang et al., 2019]. No such PIs that can specifically target the PR precursors have yet been approved for retrovirus (such as HIV-1 PR) [Huang and Chen, 2013], coronavirus (such as SARS-CoV-2 MPro), or alphavirus PRs (*e.g.* VEEV nsP2pro).

From a therapeutic standpoint, the picornavirus PRs, especially 3Cpro, have been thoroughly explored as antiviral targets. Protease inhibitors (PIs) such as rupintrivir (a peptidomimetic inhibitor of rhinovirus 3Cpro) show potent activity against human rhinoviruses by blocking 3Cpro's cleavage activity, including its autocatalytic maturation [Dragovich et al., 2003]. However, due to rapid mutational selection and PR flexibility it has limited clinical use [Hayden et al., 2003] [Patick et al., 2005].

Clinical high affinity HIV-1 PIs (indinavir, ritonavir, amprenavir, nelfinavir, tipranavir, saquinavir, atazanavir, lopinavir, and darunavir) inhibit both HIV-1 and HIV-2 PR, with variable efficacies [Brower et al., 2008] [Mahdi et al., 2015]. These small molecular inhibitors bind tightly to the active sites of HIV PRs and can inhibit the mature, free enzymes. The inhibition of a mature PR is much more efficient (even ~10,000-fold higher) compared to the immature PR that is embedded within the polyprotein [Pettit et al., 2004] [Huang et al., 2011] [Louis et al., 2011] [Park et al., 2016] [Humpolíčková et al., 2018]. The clinical PIs are unable to inhibit HIV-1 precursor processing in a therapeutically relevant concentration, which is considered to contribute to their failure in the clinic. Darunavir is the most potent inhibitor of HIV-1 PR precursor, due to its ability to engage a secondary binding site within the enzyme, but it is effective only in the micromolar range, and its efficacy is drastically reduced by resistance mutations [Louis et al., 2011] [Park et al., 2016]. Interestingly, an experimental compound (brecanavir) was found to block the processing of the HIV-1 PR precursor more efficiently than the FDA-approved PIs [Humpolíčková et al., 2018], although further investigation is required.

HIV-1 PR is active as a homodimer, each monomer contributing an aspartate to the active site. The dimer upon its release at its termini is stabilized in part by the interactions of antiparallel N- and C-terminal β -strands at the dimer interface, the C-terminal strands forming the interior of the four-stranded β -sheet. The peptides mimicking the interface strands can replace the terminal strands of HIV PR, impairing dimer formation and catalytic activity. Such highly PR-specific peptides are considered as candidate PIs [Weber, 1990] and have already been designed for HIV-1 PR [Babé et al., 1992] [Zutshi et al., 1997] [Bowman and Chmielewski, 2002] [Shultz et al., 2004] [Davis et al., 2006]. For example, the HIV-1 PR dimerization inhibitor P27 peptide consists of 27 residues (PQITLRKKRRQRRRPPQVSFNFCTLNF) encompassing the five N- and five C-terminal residues of HIV-1 PR (PQITL and CTLNF, respectively) and a 13 residue-long sequence of HIV-1 Tat's cell permeable domain (RKKRRQRRRPPQV). These terminal residues are required for the dimerization of HIV-1 PR, which from the antiparallel β -strands of the dimer interface. The P27 peptide was found to inhibit both HIV-1 and HIV-2 PR with IC₅₀'s in the low micromolar range [Davis et al., 2006].

Inhibitors targeting other viral proteins can also influence PR activation. Efavirenz, a non-nucleoside reverse transcriptase inhibitor (NNRTIs) of HIV was found to enhance the dimerization of RT. Since RT and PR domains are part of the Gag-Pol polyprotein,

increased RT dimerization may promote polyprotein dimerization, which in turn facilitates PR dimerization and causes premature activation. The premature autoprocessing was found to decrease the production of viral particles [Jochmans et al., 2010] [Sudo et al., 2013]. It is important to note that this effect was observed in tissue culture experiments only at micromolar concentration range, but such a high concentration is not relevant from the viewpoint of *in vivo* therapy. Nevertheless, this highlights the potential of drugs that might cause premature activation of the PR.

HIV-1 PR remains inactive prior to budding, but some NNRTIs such as efavirenz and rilpivirine can cause its premature activation due to enhanced multimerization of Gag-Pol [Figueiredo et al., 2006], as described above. It has been reported that caspase recruitment domain-containing protein 8 (CARD8) can be cleaved by HIV-1 PR. Premature intracellular activity of HIV-1 PR can be sensed by CARD8 that mediates inflammasome activation, causing pyroptosis of HIV-1-infected cells and clearance of latent HIV-1 in patient CD4⁺ T cells [Wang et al., 2021]. Accordingly, the use of molecules that can promote Gag-Pol dimerization (and thus causing premature PR activation) is considered as a promising strategy to activate CARD8 inflammasome [Kim and Shan, 2022]. In addition, sensitization of CARD8 can improve the efficacy of NNRTI-induced inflammasome activation, enhancing the clearance of HIV-1 from CD4⁺ T cells [Clark et al., 2023]. The sensitization can be achieved by inhibiting dipeptidyl peptidase 9 (DPP9) that is a negative regulator of CARD8, its binding to CARD8 inhibits inflammasome activation and pyroptosis. The DPP9 inhibitor Val-boroPro (VbP) was found to prevent inhibition of CARD8 by DPP9, increasing the intracellular concentration of CARD8. VbP was found to enhance the efficacy of NNRTIs and can boost CARD8 activity even in the absence of NNRTIs, which might be promising to overcome resistance against NNRTIs via sensitization of CARD8 inflammasome [Clark et al., 2023].

Similar to HIV-1 PR, the clinically approved SARS-CoV-2 MPro inhibitors nirmatrelvir [Lamb, 2022] and ensitrelvir [Syed, 2024] have well-documented efficacy against the mature enzyme. However, the precursor form was found to be less sensitive to inhibition [Huang et al., 2024] [Novotný et al., 2025]. Unlike HIV-1 PR, dimerization of MPro results in the active site loop equilibrium shifting from an inactive to an active state. It is evident from isothermal titration calorimetric experiments that potent clinical inhibitors [Nashed, Aniana et al., 2022] and GC373 [Nashed, Kneller et al., 2022] [Aniana et al., 2024] that were designed to bind to the active sites of MPro dimer exhibit very weak binding to the monomer form. Therefore, MPro and enzymes alike, which exhibit an inactive-to-active conformational equilibrium, provide an alternative strategy for drug design targeting the inactive state of the precursor monomer [Sun et al., 2022] [Yang et al., 2025]. However, potent inhibitor molecules that can selectively target the precursor SARS-CoV-2 MPro active site or that perturb its dimerization [Goyal and Goyal, 2020] [Novotný et al., 2025] are currently unavailable.

3.2 Targeting viral protease cofactors

Cofactor-dependent activation represents a unique vulnerability in several viral PRs, offering a promising strategy for selective antiviral intervention. Unlike catalytic sites, which often share structural homology with host enzymes and pose challenges for specificity, cofactors and their interaction interfaces tend to be virus-specific. This specificity reduces the risk of off-target effects and opens new possibilities for therapeutic targeting. As mentioned previously, disrupting the interaction between NS2B and NS3 in the case of flavivirus PRs (such as DENV and ZIKV) has already yielded potent inhibitors that either prevent complex formation or allosterically stabilize an inactive

conformation [Starvaggi et al., 2024]. These approaches underscore the feasibility of targeting cofactor dynamics as a viable drug strategy. Similarly, viral DNA and the pVIc peptide are both crucial cofactors for adenovirus PR activation. Targeting these interactions could inhibit virion maturation without affecting host PRs. For instance, molecules that mimic pVIc but fail to promote proper folding could act as decoy inhibitors, while intercalators or groove-binders could obstruct DNA-mediated activation [McGrath et al., 2013].

Other cofactor-dependent systems, such as HCV NS3 which requires NS4A (Figure 5B) [Failla et al., 1994], the NS3 catalytic domain of flaviviruses that utilize NS2B peptide as cofactor (Figure 4B) [Falgout et al., 1991], and herpes viral PRs, which often depend on scaffold proteins or dimerization [Shimba et al., 2004], further highlight the diversity of viral cofactor mechanisms. Each presents a potential site for pharmacological intervention, particularly where the cofactor contributes to structural stabilization or substrate orientation. As our understanding of these protease-cofactor interactions deepen, rational drug design can increasingly exploit these dependencies, thereby offering a pathway to develop broad-spectrum or virus-specific antivirals with an improved drug resistance safety profile.

4. Functional characteristics of proteases

4.1 Cleavage site sequences

The PRs (or the non-structural proteins encompassing the PR domain) that release themselves from the polyprotein contain consensus cleavage site sequences at their N and C-termini, being specific for the given PR. For example, SARS-CoV-2 MPro shows high specificity for the cleavage sites containing leucine in P2, glutamine in P1, and Ser/Ala/Gly residue in P1' position [Koudelka et al., 2021]. Both its N- and C-terminal self-cleavage sites correspond to the consensus sequence to facilitate MPro autoprocessing. Likewise, the self-processing sites of SARS-CoV-2 PLpro also represent its own consensus cleavage site sequence (Figure 3A). Interestingly, unlike MPro and PLpro, natural cleavage sites of HIV-1 PR at its termini do not have a consensus sequence motif (Figure 2A). Rather, their sequences are diverse, the recognition is strongly sequence context-dependent [Tózsér, 1997] [Tózsér, 2010] and determined by the shape and space filled by the substrate [Prabu-Jeyabalan et al., 2002]. This is referred to as substrate envelope model or hypothesis that has been defined for SARS-CoV-2 MPro, as well [Shaqra et al., 2022]. The N- and C-terminal cleavage sites of ZIKV NS3 protease (NS2B/NS3 and NS3/NS4A, respectively) are also highly similar and represent the consensus recognition sequence of the PR (Figure 4A) [Shiryaev et al., 2017] [Wahaab et al., 2021], similar to the nsp1/nsp2 and nsp2/nsp3 cleavage sites of VEEV (Figure 6A) [Morazzani et al., 2019]. Accordingly, the cleavage sites that are processed *in cis* do not have strikingly different sequence as compared to the sites that are cleaved *in trans*; and the presence of the consensus cleavage sites at the termini of a given PR ensures its self-release from the polyprotein by *in cis* and/or *in trans* cleavages.

The cleavage sites at N and C-termini of the PR domain might be considerably different if more than one PR facilitates polyprotein processing. In these cases, different cleavage site sequences are recognized by these PRs. For example, genome of HCV encodes two different PRs which release the NS3 PR from the polyprotein: the NS3 protein is cleaved at its N-terminus by the NS2/NS3 zinc-dependent protease domain at the NS2B/NS3 cleavage site, while the NS3/NS4 junction at the C-terminus is cleaved by NS3 [Failla et al., 1994]. Accordingly, the cleavage site sequences at the termini of HCV NS3 non-structural protein represent a different consensus sequence (Figure 5A).

The cleavage site sequences do not provide direct information about the order of self-cleavages, and the sequential order cannot be estimated based on cleavage efficiencies of substrates representing different autolytic sites because the enzyme kinetic measurements and specificity studies are performed almost exclusively in the context of the mature enzymes, and the oligopeptide substrates that represent the intramolecular cleavage sites can be easily processed in the *in vitro* assays by the mature enzymes *in trans*.

4.2 Evolution and functional significance of proteases

The ability of viral PRs catalysing their release from their polyprotein precursors is a conserved pathway across numerous viral families, serving multiple evolutionary and functional purposes. The long polyproteins, which are post-translationally processed into individual functional units, reduce the need for complex gene regulation strategies such as multiple promoters or splicing events. Particularly in ss(+)RNA viruses like picornaviruses, flaviviruses, and coronaviruses, self-cleaving PRs not only initiate their own release but also coordinate the maturation of viral proteins in precise amounts. Functionally, self-proteolysis enables temporal and spatial regulation of viral protein activation, ensuring that PRs are activated when meeting the requirements of transient correctly folded (*in cis* cleavage), and subsequently capable of initiating further processing steps (*via* delayed *in trans* cleavage), thereby avoiding premature cleavage events that could disrupt replication or assembly. Additionally, this autonomous activation pathway reduces reliance on host PRs, and especially advantageous in diverse cellular environments. For example, SARS-CoV-2 MPro self-activates within the infected cell to coordinate the replication cycle [Romano et al., 2020], while in retroviruses like HIV, the activation of the PR is delayed until after particle assembly, ensuring infectivity only after proper assembly [Tabler et al., 2022]. The order of Gag polyprotein cleavage is also modulated through conformational dynamics within the flaps of HIV PR leading to transient, sparsely populated production complex between PR and Gag substrates [Deshmukh et al., 2017]. The modular nature of self-cleaving PRs also supports evolutionary plasticity, allowing cleavage sites and domains to adapt independently to new substrates or host conditions, as observed in 3Cpro and NS3 PRs [Pathak et al., 2020] [Matthews et al., 1994] [Tsu et al., 2021]. Moreover, many viral PRs acquire multi-functionality after self-processing, engaging in immune evasion by targeting host signaling proteins such as mitochondrial antiviral-signaling protein (MAVS) or Toll-like receptor 3 adaptor protein (TRIF) [Chin et al., 2022]. Thus, self-proteolysis not only enhances viral replication efficiency, but also underscores the protease's broader role in host-virus interactions and evolutionary adaptation.

5. Summary

Most of the RNA viruses utilize the polyprotein synthesis pathway for replication, and the large polyprotein(s) translated from the viral genome are cleaved subsequently into functional proteins. In contrast, most ss(-)RNA viruses do not use this pathway as they have a segmented RNA genome. The dsRNA viruses exhibit differences in their replication. The enveloped dsRNA viruses (such as *reoviridae*) have segmented genome, while the genome of non-enveloped dsRNA viruses is non-segmented and they use polyprotein s pathway (such as birna- and orthototiviruses). Accordingly, the replication of the non-enveloped dsRNA viruses is more similar to those of ss(+)RNA viruses.

The viral polyproteins can be processed by virus-encoded and/or by host PRs. The ss(-)RNA viruses having segmented genome do not have their own PR, while the genomes of most ss(+)RNA viruses using polyprotein pathway encode at least one PR.

Most of the dsRNA viruses also lack a viral PR, nevertheless, birnaviruses have their own (VP4 PR). The majority of ss(+)RNA virus PRs are serine or cysteine PRs, while only retroviruses encode an aspartic PR. Some of the RNA viruses have a non-canonical PR, for example the VP4 protease of dsRNA birnaviruses contains a Ser-Lys catalytic dyad (**Figure 9**). Downregulation of activity through self-proteolysis at specific sites within the PR upon depletion of natural substrates is also a feature observed in HIV PRs [**Mildner et al., 1994**] [**Louis et al., 2007**].

Viruses that encode their own PR(s) exhibit differences in polyprotein processing, especially in the mechanism of PR precursor activation. This includes the *in cis* or *in trans* pathways and sequential order of the cleavages that vary between these viruses (**Figure 9**). Despite these differences, all the mature PRs retain their trans-activity upon their release. Polyproteins of HIV-1 and SARS-CoV-2 are processed exclusively by the viral PR. Not only viral but host PRs might also contribute to polyprotein processing, such as in the case of ZIKV and HCV. Some of the viruses use only intramolecular cleavages to release the PR from the precursor, such as ss(+)RNA potyvirus TEV and picornavirus FMDV, as well as dsRNA birnavirus VP4 PR.

Envelope Virus		Protease	Cleavage N-term. / C-term.			
ssRNA	ss(-)RNA	yes	• Lassa			
			• Influenza			
			• RSV			
	ss(+)RNA	yes	• HIV			
			• CHIKV, SINV, VEEV	nsP2		
			• HCV	NS3		
			• YFV, JEV, DENV, WNV, ZIKV	NS3		
			• EAV, PRRSV	nsp4		
			• SARS-CoV(-2)	Mpro		
	no	• MNV, NV	NS6			
		• FMDV, EV-D68, PV	3C pro			
		• TEV, TuMV	NIa-Pro			
dsRNA	yes	• Rota				
	no	• IBDV, IPNV	VP4			
		• BTLV				

Protease

Aspartic protease

Serine protease

Cysteine protease

No viral protease

Cleavage

Intramolecular (*in cis*) cleavage

Intermolecular (*in trans*) cleavage

Figure 9. Protease precursor activation of RNA viruses. The RNA viruses are grouped based on their genome and replication strategy (Baltimore classification). Salient families of the groups and medically important viruses are listed. The viral PRs

are colored based on their catalytic mechanism. The symbols indicate the type of cleavage (intra- or intermolecular) at the N and C-terminus of the protease during PR precursor processing. Question mark indicates yet unknown mechanism.

The PRs of enveloped ss(+)RNA viruses such as coronaviruses (*e.g.* SARS-CoV-2), flaviviruses (*e.g.* ZIKV), and alphaviruses (*e.g.* SINV) share the PR precursor activation mechanism of HIV-1 PR. In these viruses, the intramolecular (*in cis*) cleavage at the N-terminus of the PR is followed by intermolecular (*in trans*) processing of its C-terminus (**Figure 9**). This self-activation mechanism is utilized by most enveloped ss(+)RNA viruses either replicating with (like retroviruses) or without DNA intermediate (**Figure 9**). It is important to note that the number of viruses whose PR precursor activation has been explored in detail is only limited, accordingly, it is hard to establish general activation mechanism for all RNA virus groups and families. For example, the molecular mechanism of PR precursor processing has not been determined for all retroviruses but only for HIV-1 PR. Consequently, it can be only hypothesized that similar to HIV-1 PR, PRs of other retroviruses (*e.g.* human T-lymphotropic virus, HTLV) are also cleaved first at their N-terminus intramolecularly, followed by intermolecular processing at their C-terminus.

In order to determine possible determinants of PR precursor activation mechanism, we correlated some characteristics of the viral life-cycles with those of the activation mechanism. Utilization of only intramolecular cleavages (*e.g.* flavivirus HCV, arterivirus EAV, potyvirus TEV, and picornavirus FMDV PRs), was found to be characteristic for both enveloped and non-enveloped RNA viruses (**Figure 9**). Only monomeric enzymes are processed at their N- and C-termini *in cis*, but this is not a unique feature because other monomeric PRs are also processed *in trans*, such as calicivirus NV and togavirus CHIKV PRs. The catalytic mechanism can not be considered as a sole determinant of the utilized mechanism because *cis*-only activation was reported for both serine and cysteine PRs. In addition, the use of a cofactor is not a prerequisite for *cis*-only activation, as *e.g.* TEV PR does not use any cofactor.

Besides their distinct kinetic behaviours (the *in cis* cleavages are dilution-insensitive), the intra- and intermolecular cleavage reactions exhibit other considerable differences. The *in cis* cleavages are unimolecular and therefore, such cleavages are independent from the stoichiometry of the enzyme and the "substrate" (*i.e.* the PR being a part of the polyprotein). The *in cis* cleavages can be carried out efficiently even at low precursor concentration and do not require maturation or oligomerization of the PR (except HIV PRs that are active as obligate homodimers). In addition, the precursor forms are less susceptible for active site-directed inhibition than the mature enzymes. Based on our comparison, using both *in cis* and *in trans* cleavages is a more common activation mechanism among the RNA viruses (**Figure 9**). This activation mechanism is considered to enable a faster activation of PR precursors, due to the amplification. Namely, a single PR can activate multiple other PRs through *in trans* cleavages, making the precursor activation more rapid (as a positive feedback), while the intramolecular activation is limited to the precursor itself and no amplification occurs. In addition, the *in cis*-processing is considered to be a rate-limiting cleavage step of precursor activation that occurs much slower than the subsequent *in trans* cleavages (such as the activation of HIV and SARS-CoV-2 PRs).

Processing of viral PR precursors is a rate-limiting step of virus maturation, therefore, inhibitors that can block the precursor forms of the proteases are considered to be promising targets of drug discovery, and targeting immutable components of the cells such as the inflammasome, causing premature PR activation, is also a promising approach. In addition, targeting the interactions between the PRs and their cofactors is

also a possible approach to inhibit PR maturation, emphasizing the significance of understanding precursor processing. The presence of a complete active site in SARS-CoV-2 MPro and enzymes alike, unlike that of HIV PR in which the active site is formed upon dimerization, underscores the importance of identifying and characterizing compounds that tightly bind to the inactive state of the PR precursor and thereby providing an alternative avenue to arrest polyprotein maturation. The latter strategy may be advantageous in curbing drug resistance invariably observed upon treatment with drugs designed to bind to mature PRs.

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of data and materials: Not applicable.

Competing interests: The authors declare that they have no competing interests

Author Contributions:

All authors have read and agreed to the published version of the manuscript. Conceptualization, J.A.M., M.G., M.M., N.T.N., J.M.L., J.T.; Funding acquisition, J.A.M., J.M.L., J.T.; Visualization, J.A.M., M.G., J.M.L.; Writing - original draft, J.A.M., M.G., M.M., N.T.N., J.M.L., J.T.; Writing - review & editing, J.A.M., M.G., N.T.N., J.M.L., J.T.

Funding:

Project no. TKP2021-EGA-20 has been implemented with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the TKP2021-EGA funding scheme. This project was supported by the ADVANCED_150532 project of National Research, Development and Innovation Fund (NKFIH) (to József Tózsér). János A. Mótyán is supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences (BO/00110/23/5). Supported by the University of Debrecen (UD) Program for Scientific Publication. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Acknowledgments:

The authors are grateful to the colleagues at the Department of Biochemistry and Molecular Biology (Faculty of Medicine, University of Debrecen), especially to Gyula Hoffka (member of Laboratory of Retroviral Biochemistry) for structural representation of HIV-1 PR and Károly Jambrovics for his critical comments and suggestions. The authors are grateful for Chun Tang (College of Chemistry and Molecular Engineering, Beijing National Laboratory for Molecular Sciences & Center for Quantitative Biology, Peking University) for providing coordinate files of HIV-1 PR [**Tang et al., 2008**]. This research was supported by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK; Project number DK075166-01) within the National Institutes of Health (NIH). The contributions of the NIH author(s) are considered Works of the United States Government. The findings and conclusions presented in this paper are those of the author(s) and do not necessarily reflect the views of the NIH or the U.S. Department of Health and Human Services.

Conflicts of Interest:

The authors declare that they have no conflicts of interest with the contents of this article.

List of abbreviations:

AIDS	Acquired immunodeficiency syndrome
BSNV	Blotched Snakehead virus
BTLV	Birch toti-like virus
CHIKV	Chikungunya virus
3CLpro	3-chymotrypsin-like protease
DBV	Drosophila B virus
DENV	Dengue virus
DXV	Drosophila X virus
EAV	Equine arteritis virus
EEEV	Eastern Equine Encephalitis Virus
EV	Enterovirus
FMDV	Foot-and-mouth disease virus
HCV	Hepatitis C virus
HIV-1	Human immunodeficiency virus type 1
HIV-2	Human immunodeficiency virus type 2
HTLV	Human T-lymphotropic virus
IBDV	Infectious bursal disease virus
IPNV	Infectious pancreatic necrosis virus
MERS	Middle East respiratory syndrome coronavirus
MNV	Murine norovirus
MPro	Main protease
NNRTI	Non-nucleoside reverse transcriptase inhibitor
nsp	Non-structural protein
NV	Norovirus
PI	Protease inhibitor
PLpro	Papain-like protease
PRRSV	Porcine reproductive and respiratory syndrome virus
PV	Picornavirus
PVBV	Pepper vein banding virus
PR	Protease
RSV	Respiratory syncytial virus
RT	Reverse transcriptase
SARS-CoV	Severe acute respiratory syndrome coronavirus
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SFV	Semliki Forest virus

SINV	Sindbis virus
TEV	Tobacco etch virus
TFP	trans-frame region (TFR) peptide
TFR	Trans-frame region
TV-1	Tellina virus 1
VEEV	Venezuelan equine encephalitis virus
WEEV	Western equine encephalitis virus
WNV	West Nile virus
YFV	Yellow fever virus
ZIKV	Zika virus

References

- Adams, M. J., Antoniw, J. F., & Beaudoin, F. (2005). Overview and analysis of the polyprotein cleavage sites in the family Potyviridae. *Molecular plant pathology*, 6(4), 471–487.
- Aniana, A., Nashed, N. T., Ghirlando, R., Coates, L., Kneller, D. W., Kovalevsky, A., Louis, J. M. (2023). Insights into the mechanism of SARS-CoV-2 main protease autocatalytic maturation from model precursors. *Commun Biol*, 6(1), 1159.
- Aniana, A., Nashed, N. T., Ghirlando, R., Drago, V. N., Kovalevsky, A., & Louis, J. M. (2024). Characterization of alternate encounter assemblies of SARS-CoV-2 main protease. *The Journal of biological chemistry*, 300(9), 107675.
- Báez-Santos, Y. M., St John, S. E., & Mesecar, A. D. (2015). The SARS-coronavirus papain-like protease: structure, function and inhibition by designed antiviral compounds. *Antiviral research*, 115, 21–38.
- Babé, L. M., Rosé, J., Craik, C. S. (1992). Synthetic "interface" peptides alter dimeric assembly of the HIV 1 and 2 proteases. *Protein science: a publication of the Protein Society*, 1(10), 1244–1253.
- Babé, L. M., & Craik, C. S. (1997). Viral proteases: evolution of diverse structural motifs to optimize function. *Cell*, 91(4), 427–430.
- Bafna, K., Cioffi, C. L., Krug, R. M., & Montelione, G. T. (2022). Structural similarities between SARS-CoV2 3CL^{pro} and other viral proteases suggest potential lead molecules for developing broad spectrum antivirals. *Frontiers in chemistry*, 10, 948553.
- Balasuriya, U. B., Go, Y. Y., & MacLachlan, N. J. (2013). Equine arteritis virus. *Veterinary microbiology*, 167(1-2), 93–122.
- Barrette-Ng, I. H., Ng, K. K., Mark, B. L., Van Aken, D., Cherney, M. M., Garen, C., Kolodenco, Y., Gorbalenya, A. E., Snijder, E. J., & James, M. N. (2002). Structure of arterivirus nsp4. The smallest chymotrypsin-like proteinase with an alpha/beta C-terminal extension and alternate conformations of the oxyanion hole. *The Journal of biological chemistry*, 277(42), 39960–39966.
- 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. *Journal of virology*, 79(24), 15189–15198.
- Bartenschlager, R., Ahlborn-Laake, L., Yasargil, K., Mous, J., & Jacobsen, H. (1995). Substrate determinants for cleavage in cis and in trans by the hepatitis C virus NS3 proteinase. *Journal of virology*, 69(1), 198–205.
- Belliot, G., Sosnovtsev, S. V., Mitra, T., Hammer, C., Garfield, M., & Green, K. Y. (2003). In vitro proteolytic processing of the MD145 norovirus ORF1 nonstructural polyprotein yields stable precursors and products similar to those detected in calicivirus-infected cells. *Journal of virology*, 77(20), 10957–10974.

- Bera, A. K., Kuhn, R. J., & Smith, J. L. (2007). Functional characterization of cis and trans activity of the Flavivirus NS2B-NS3 protease. *The Journal of biological chemistry*, *282*(17), 12883–12892.
- Birtley, J. R., Knox, S. R., Jaulent, A. M., Brick, P., Leatherbarrow, R. J., & Curry, S. (2005). Crystal structure of foot-and-mouth disease virus 3C protease. New insights into catalytic mechanism and cleavage specificity. *The Journal of biological chemistry*, *280*(12), 11520–11527.
- Blair, W. S., Semler, B. L. (1991). Self-cleaving proteases. *Current opinion in cell biology*, *3*(6), 1039–1045.
- Borges, P. H. O., Ferreira, S. B., & Silva, F. P., Jr (2024). Recent Advances on Targeting Proteases for Antiviral Development. *Viruses*, *16*(3), 366.
- Bowman, M. J., & Chmielewski, J. (2002). Novel strategies for targeting the dimerization interface of HIV protease with cross-linked interfacial peptides. *Biopolymers*, *66*(2), 126–133.
- Bozóki, B., Mótyán, J. A., Hoffka, G., Waugh, D. S., & Tózsér, J. (2020). Specificity Studies of the Venezuelan Equine Encephalitis Virus Non-Structural Protein 2 Protease Using Recombinant Fluorescent Substrates. *International journal of molecular sciences*, *21*(20), 7686.
- Braack, L., Gouveia de Almeida, A. P., Cornel, A. J., Swanepoel, R., & de Jager, C. (2018). Mosquito-borne arboviruses of African origin: review of key viruses and vectors. *Parasites & vectors*, *11*(1), 29.
- Brower, E. T., Bacha, U. M., Kawasaki, Y., Freire, E. (2008). Inhibition of HIV-2 protease by HIV-1 protease inhibitors in clinical use. *Chemical biology & drug design*, *71*(4), 298–305.
- Campillay-Véliz, C. P., Carvajal, J. J., Avellaneda, A. M., Escobar, D., Covián, C., Kalergis, A. M., & Lay, M. K. (2020). Human Norovirus Proteins: Implications in the Replicative Cycle, Pathogenesis, and the Host Immune Response. *Frontiers in immunology*, *11*, 961.
- Carrington, J. C., & Dougherty, W. G. (1987). Processing of the tobacco etch virus 49K protease requires autoproteolysis. *Virology*, *160*(2), 355–362.
- Carrington, J. C., Cary, S. M., & Dougherty, W. G. (1988). Mutational analysis of tobacco etch virus polyprotein processing: cis and trans proteolytic activities of polyproteins containing the 49-kilodalton proteinase. *Journal of virology*, *62*(7), 2313–2320.
- Centazzo, M., Manganaro, L., & Alvisi, G. (2023). Cellular Targets of HIV-1 Protease: Just the Tip of the Iceberg? *Viruses*, *15*(3), 712.
- Chakraborty, P., Acquasaliente, L., Pelc, L. A., & Di Cera, E. (2018). Interplay between conformational selection and zymogen activation. *Scientific reports*, *8*(1), 4080.
- Chambers, T. J., Weir, R. C., Grakoui, A., McCourt, D. W., Bazan, J. F., Fletterick, R. J., & Rice, C. M. (1990). Evidence that the N-terminal domain of nonstructural protein NS3 from yellow fever virus is a serine protease responsible for site-specific cleavages in the viral polyprotein. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(22), 8898–8902.
- Chatterjee, A., Mridula, P., Mishra, R. K., Mittal, R., & Hosur, R. V. (2005). Folding regulates autoprocessing of HIV-1 protease precursor. *The Journal of biological chemistry*, *280*(12), 11369–11378.
- Chen, S., Jonas, F., Shen, C., & Hilgenfeld, R. (2010). Liberation of SARS-CoV main protease from the viral polyprotein: N-terminal autocleavage does not depend on the mature dimerization mode. *Protein & cell*, *1*(1), 59–74.
- Cheng, S. C., Chang, G. G., & Chou, C. Y. (2010). Mutation of Glu-166 blocks the substrate-induced dimerization of SARS coronavirus main protease. *Biophysical journal*, *98*(7), 1327–1336.
- Chin, C. V., & Saeed, M. (2022). Surgical Strikes on Host Defenses: Role of the Viral Protease Activity in Innate Immune Antagonism. *Pathogens (Basel, Switzerland)*, *11*(5), 522.
- Chung, I. Y., & Paetzel, M. (2011). Crystal structure of a viral protease intramolecular acyl-enzyme complex: insights into cis-cleavage at the VP4/VP3 junction of Tellina birnavirus. *The Journal of biological chemistry*, *286*(14), 12475–12482.

- Clark, K. M., Kim, J. G., Wang, Q., Gao, H., Presti, R. M., & Shan, L. (2023). Chemical inhibition of DPP9 sensitizes the CARD8 inflammasome in HIV-1-infected cells. *Nature chemical biology*, *19*(4), 431-439.
- Coffin, J.M., Hughes, S.H., Varmus, H.E. (1997) Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY)
- Compton, J. R., Mickey, M. J., Hu, X., Marugan, J. J., & Legler, P. M. (2017). Mutation of Asn-475 in the Venezuelan Equine Encephalitis Virus nsP2 Cysteine Protease Leads to a Self-Inhibited State. *Biochemistry*, *56*(47), 6221-6230.
- Constant, D.A., Mateo, R., Nagamine, C.M., & Kirkegaard, K. (2018). Targeting intramolecular proteinase NS2B/3 cleavages for *trans*-dominant inhibition of dengue virus. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(40), 10136-10141.
- Curry, S., Roqué-Rosell, N., Zunszain, P. A., & Leatherbarrow, R. J. (2007). Foot-and-mouth disease virus 3C protease: recent structural and functional insights into an antiviral target. *The international journal of biochemistry & cell biology*, *39*(1), 1-6.
- Davis, D. A., Brown, C. A., Singer, K. E., Wang, V., Kaufman, J., Stahl, S. J., Wingfield, P., Maeda, K., Harada, S., Yoshimura, K., Kosalaraksa, P., Mitsuya, H., Yarchoan, R. (2006). Inhibition of HIV-1 replication by a peptide dimerization inhibitor of HIV-1 protease. *Antiviral research*, *72*(2), 89-99.
- Delmas, B., Mundt, E., & Gorbalenya, A. E. (2013). Birnavirus VP4 Processing Endopeptidase. *Handbook of Proteolytic Enzymes*, 3517-3523.
- Deshmukh, L., Tugarinov, V., Louis, J. M., & Clore, G. M. (2017). Binding kinetics and substrate selectivity in HIV-1 protease-Gag interactions probed at atomic resolution by chemical exchange NMR. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(46), E9855-E9862.
- Dougherty, W. G., & Semler, B. L. (1993). Expression of virus-encoded proteinases: functional and structural similarities with cellular enzymes. *Microbiological reviews*, *57*(4), 781-822.
- Dragovich, P. S., Prins, T. J., Zhou, R., Johnson, T. O., Hua, Y., Luu, H. T., Sakata, S. K., Brown, E. L., Maldonado, F. C., Tuntland, T., Lee, C. A., Fuhrman, S. A., Zalman, L. S., Patick, A. K., Matthews, D. A., Wu, E. Y., Guo, M., Borer, B. C., Nayyar, N. K., Moran, T., ... Ferre, R. A. (2003). Structure-based design, synthesis, and biological evaluation of irreversible human rhinovirus 3C protease inhibitors. 8. Pharmacological optimization of orally bioavailable 2-pyridone-containing peptidomimetics. *Journal of medicinal chemistry*, *46*(21), 4572-4585.
- Ehrenfeld, E., Domingo, E., Roos, R.P. (2010) The Picornaviruses. *ASM Press*. DOI: 10.1128/9781555816698
- Emmott, E., Sweeney, T. R., & Goodfellow, I. (2015). A Cell-based Fluorescence Resonance Energy Transfer (FRET) Sensor Reveals Inter- and Intragenogroup Variations in Norovirus Protease Activity and Polyprotein Cleavage. *The Journal of biological chemistry*, *290*(46), 27841-27853.
- Failla, C., Tomei, L., & De Francesco, R. (1994). Both NS3 and NS4A are required for proteolytic processing of hepatitis C virus nonstructural proteins. *Journal of virology*, *68*(6), 3753-3760.
- Falgout, B., Pethel, M., Zhang, Y. M., & Lai, C. J. (1991). Both nonstructural proteins NS2B and NS3 are required for the proteolytic processing of dengue virus nonstructural proteins. *Journal of virology*, *65*(5), 2467-2475.
- Fang, Y., & Snijder, E. J. (2010). The PRRSV replicase: exploring the multifunctionality of an intriguing set of nonstructural proteins. *Virus research*, *154*(1-2), 61-76.
- Figueiredo, A., Moore, K. L., Mak, J., Sluis-Cremer, N., de Bethune, M. P., & Tachedjian, G. (2006). Potent nonnucleoside reverse transcriptase inhibitors target HIV-1 Gag-Pol. *PLoS pathogens*, *2*(11), e119.
- Fitzgerald, P. M., McKeever, B. M., VanMiddlesworth, J. F., Springer, J. P., Heimbach, J. C., Leu, C. T., Herber, W. K., Dixon, R. A., & Darke, P. L. (1990). Crystallographic analysis of a complex between human immunodeficiency virus type 1 protease and acetyl-pepstatin at 2.0-Å resolution. *The Journal of biological chemistry*, *265*(24), 14209-14219.

- Garmashova, N., Gorchakov, R., Volkova, E., Paessler, S., Frolova, E., & Frolov, I. (2007). The Old World and New World alphaviruses use different virus-specific proteins for induction of transcriptional shutoff. *Journal of virology*, *81*(5), 2472-2484.
- Gelderblom HR. Structure and Classification of Viruses. In: Baron S, editor. Medical Microbiology. 4th edition. Galveston (TX): University of Texas Medical Branch at Galveston; 1996. Chapter 41.
- Golda, M., Hoffka, G., Cherry, S., Tropea, J. E., Lountos, G. T., Waugh, D. S., Wlodawer, A., Tózsér, J., Mótýán, J. A. (2024). P1' specificity of the S219V/R203G mutant tobacco etch virus protease. *Proteins*, *92*(9), 1085-1096.
- Goyal, B., & Goyal, D. (2020). Targeting the Dimerization of the Main Protease of Coronaviruses: A Potential Broad-Spectrum Therapeutic Strategy. *ACS combinatorial science*, *22*(6), 297-305.
- Grakoui, A., McCourt, D. W., Wychowski, C., Feinstone, S. M., & Rice, C. M. (1993). A second hepatitis C virus-encoded proteinase. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(22), 10583-10587.
- Grubman, M. J., & Baxt, B. (2004). Foot-and-mouth disease. *Clinical microbiology reviews*, *17*(2), 465-493.
- Guo, B., Lin, J., & Ye, K. (2011). Structure of the autocatalytic cysteine protease domain of potyvirus helper-component proteinase. *The Journal of biological chemistry*, *286*(24), 21937-21943.
- Gupta, S.P. (2018) Viral Proteases and Their Inhibitors. *Academic Press*.
- Harcourt, B. H., Jukneliene, D., Kanjanahaluethai, A., Bechill, J., Severson, K. M., Smith, C. M., Rota, P. A., & Baker, S. C. (2004). Identification of severe acute respiratory syndrome coronavirus replicase products and characterization of papain-like protease activity. *Journal of virology*, *78*(24), 13600-13612.
- Hardy, W. R., & Strauss, J. H. (1989). Processing the nonstructural polyproteins of sindbis virus: nonstructural proteinase is in the C-terminal half of nsP2 and functions both in cis and in trans. *Journal of virology*, *63*(11), 4653-4664.
- Hardy M. E. (2005). Norovirus protein structure and function. *FEMS microbiology letters*, *253*(1), 1-8.
- Harrison, J. J. E. K., Passos, D. O., Bruhn, J. F., Bauman, J. D., Tuberty, L., DeStefano, J. J., Ruiz, F. X., Lyumkis, D., & Arnold, E. (2022). Cryo-EM structure of the HIV-1 Pol polyprotein provides insights into virion maturation. *Science advances*, *8*(27), eabn9874.
- Hayden, F. G., Turner, R. B., Gwaltney, J. M., Chi-Burris, K., Gersten, M., Hsyu, P., Patick, A. K., Smith, G. J., 3rd, & Zalman, L. S. (2003). Phase II, randomized, double-blind, placebo-controlled studies of rupintrivir nasal spray 2-percent suspension for prevention and treatment of experimentally induced rhinovirus colds in healthy volunteers. *Antimicrobial agents and chemotherapy*, *47*(12), 3907-3916.
- Hernandez, R., Brown, D. T., & Paredes, A. (2014). Structural differences observed in arboviruses of the alphavirus and flavivirus genera. *Advances in virology*, *2014*, 259382.
- Herod, M. R., Gold, S., Lasecka-Dykes, L., Wright, C., Ward, J. C., McLean, T. C., Forrest, S., Jackson, T., Tuthill, T. J., Rowlands, D. J., & Stonehouse, N. J. (2017). Genetic economy in picornaviruses: Foot-and-mouth disease virus replication exploits alternative precursor cleavage pathways. *PLoS pathogens*, *13*(10), e1006666.
- Hijikata, M., Mizushima, H., Akagi, T., Mori, S., Kakiuchi, N., Kato, N., Tanaka, T., Kimura, K., & Shimotohno, K. (1993). Two distinct proteinase activities required for the processing of a putative nonstructural precursor protein of hepatitis C virus. *Journal of virology*, *67*(8), 4665-4675.
- Hoffka, G., Lountos, G. T., Needle, D., Wlodawer, A., Waugh, D. S., Tózsér, J., & Mótýán, J. A. (2023). Self-inhibited State of Venezuelan Equine Encephalitis Virus (VEEV) nsP2 Cysteine Protease: A Crystallographic and Molecular Dynamics Analysis. *Journal of molecular biology*, *435*(6), 168012.
- Hsu, M. F., Kuo, C. J., Chang, K. T., Chang, H. C., Chou, C. C., Ko, T. P., Shr, H. L., Chang, G. G., Wang, A. H., & Liang, P. H. (2005). Mechanism of the maturation process of SARS-CoV 3CL protease. *The Journal of biological chemistry*, *280*(35), 31257-31266.

- Hu, X., Compton, J. R., Leary, D. H., Olson, M. A., Lee, M. S., Cheung, J., Ye, W., Ferrer, M., Southall, N., Jadhav, A., Morazzani, E. M., Glass, P. J., Marugan, J., & Legler, P. M. (2016). Kinetic, Mutational, and Structural Studies of the Venezuelan Equine Encephalitis Virus Nonstructural Protein 2 Cysteine Protease. *Biochemistry*, 55(21), 3007-3019.
- Huang, L., & Chen, C. (2010). Autoprocessing of human immunodeficiency virus type 1 protease miniprecursor fusions in mammalian cells. *AIDS research and therapy*, 7, 27.
- Huang, L., Li, Y., Chen, C. (2011). Flexible catalytic site conformations implicated in modulation of HIV-1 protease autoprocessing reactions. *Retrovirology*, 8, 79.
- Huang, L., Chen, C. (2013). Understanding HIV-1 protease autoprocessing for novel therapeutic development. *Future medicinal chemistry*, 5(11), 1215-1229.
- Huang, L., Li, L., Tien, C., LaBarbera, D. V., Chen, C. (2019). Targeting HIV-1 Protease Autoprocessing for High-throughput Drug Discovery and Drug Resistance Assessment. *Scientific reports*, 9(1), 301.
- Huang, L., Gish, M., Boehlke, J., Jeep, R. H., & Chen, C. (2024). Assay Development and Validation for Innovative Antiviral Development Targeting the N-Terminal Autoprocessing of SARS-CoV-2 Main Protease Precursors. *Viruses*, 16(8), 1218.
- Humpolíčková, J., Weber, J., Starková, J., Mašíňová, E., Günterová, J., Flaisigová, I., Konvalinka, J., Majerová, T. (2018). Inhibition of the precursor and mature forms of HIV-1 protease as a tool for drug evaluation. *Scientific reports*, 8(1), 10438.
- Ishima, R., Torchia, D. A., Lynch, S. M., Gronenborn, A. M., & Louis, J. M. (2003). Solution structure of the mature HIV-1 protease monomer: insight into the tertiary fold and stability of a precursor. *The Journal of biological chemistry*, 278(44), 43311-43319.
- Ishima, R., Torchia, D. A., & Louis, J. M. (2007). Mutational and structural studies aimed at characterizing the monomer of HIV-1 protease and its precursor. *The Journal of biological chemistry*, 282(23), 17190-17199.
- Jamal, S. M., & Belsham, G. J. (2013). Foot-and-mouth disease: past, present and future. *Veterinary research*, 44(1), 116.
- Jochmans, D., Anders, M., Keuleers, I., Smeulders, L., Kräusslich, H. G., Kraus, G., & Müller, B. (2010). Selective killing of human immunodeficiency virus infected cells by non-nucleoside reverse transcriptase inhibitor-induced activation of HIV protease. *Retrovirology*, 7, 89.
- Kapust, R. B., Tözsér, J., Copeland, T. D., & Waugh, D. S. (2002). The P1' specificity of tobacco etch virus protease. *Biochemical and biophysical research communications*, 294(5), 949-955.
- Katoh, I., Yoshinaka, Y., Rein, A., Shibuya, M., Odaka, T., & Oroszlan, S. (1985). Murine leukemia virus maturation: protease region required for conversion from "immature" to "mature" core form and for virus infectivity. *Virology*, 145(2), 280-292.
- Khan, A. R., Cherney, M. M., Tarasova, N. I., & James, M. N. (1997). Structural characterization of activation 'intermediate 2' on the pathway to human gastricsin. *Nature structural biology*, 4(12), 1010-1015.
- Khan, A. R., & James, M. N. (1998). Molecular mechanisms for the conversion of zymogens to active proteolytic enzymes. *Protein science : a publication of the Protein Society*, 7(4), 815-836.
- Khan, A. R., Khazanovich-Bernstein, N., Bergmann, E. M., & James, M. N. (1999). Structural aspects of activation pathways of aspartic protease zymogens and viral 3C protease precursors. *Proceedings of the National Academy of Sciences of the United States of America*, 96(20), 10968-10975.
- Kim, J. L., Morgenstern, K. A., Lin, C., Fox, T., Dwyer, M. D., Landro, J. A., Chambers, S. P., Markland, W., Lepre, C. A., O'Malley, E. T., Harbeson, S. L., Rice, C. M., Murcko, M. A., Caron, P. R., & Thomson, J. A. (1996). Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. *Cell*, 87(2), 343-355.
- Kim, Y., Lovell, S., Tiew, K. C., Mandadapu, S. R., Alliston, K. R., Battaile, K. P., Groutas, W. C., & Chang, K. O. (2012). Broad-spectrum antivirals against 3C or 3C-like

- proteases of picornaviruses, noroviruses, and coronaviruses. *Journal of virology*, *86*(21), 11754–11762.
- Kim, J. G., & Shan, L. (2022). Beyond Inhibition: A Novel Strategy of Targeting HIV-1 Protease to Eliminate Viral Reservoirs. *Viruses*, *14*(6), 1179.
- Kimura, S., Caldarini, M., Broglia, R. A., Dokholyan, N. V., & Tiana, G. (2014). The maturation of HIV-1 protease precursor studied by discrete molecular dynamics. *Proteins*, *82*(4), 633–639.
- Klemm, T., Ebert, G., Calleja, D. J., Allison, C. C., Richardson, L. W., Bernardini, J. P., Lu, B. G., Kuchel, N. W., Grohmann, C., Shibata, Y., Gan, Z. Y., Cooney, J. P., Doerflinger, M., Au, A. E., Blackmore, T. R., van der Heden van Noort, G. J., Geurink, P. P., Ovaas, H., Newman, J., Riboldi-Tunnicliffe, A., ... Komander, D. (2020). Mechanism and inhibition of the papain-like protease, PLpro, of SARS-CoV-2. *The EMBO journal*, *39*(18), e106275.
- Kneller, D. W., Phillips, G., Weiss, K. L., Pant, S., Zhang, Q., O'Neill, H. M., Coates, L., & Kovalevsky, A. (2020). Unusual zwitterionic catalytic site of SARS-CoV-2 main protease revealed by neutron crystallography. *The Journal of biological chemistry*, *295*(50), 17365–17373.
- Konvalinka, J., Kräusslich, H. G., & Müller, B. (2015). Retroviral proteases and their roles in virion maturation. *Virology*, *479*–480, 403–417.
- Koudelka, T., Boger, J., Henkel, A., Schönherr, R., Krantz, S., Fuchs, S., Rodríguez, E., Redecke, L., & Tholey, A. (2021). N-Terminomics for the Identification of In Vitro Substrates and Cleavage Site Specificity of the SARS-CoV-2 Main Protease. *Proteomics*, *21*(2), e2000246.
- Kovalevsky, A., Coates, L., Kneller, D. W., Ghirlando, R., Aniana, A., Nashed, N. T., & Louis, J. M. (2022). Unmasking the Conformational Stability and Inhibitor Binding to SARS-CoV-2 Main Protease Active Site Mutants and Miniprecursor. *Journal of molecular biology*, *434*(24), 167876.
- Kovalevsky, A., Aniana, A., Coates, L., Ghirlando, R., Nashed, N. T., & Louis, J. M. (2024). Visualizing the Active Site Oxyanion Loop Transition Upon Ensitrelvir Binding and Transient Dimerization of SARS-CoV-2 Main Protease. *Journal of molecular biology*, *436*(13), 168616.
- Lamb, Y. N. (2022). Nirmatrelvir Plus Ritonavir: First Approval. *Drugs*, *82*(5), 585–591.
- Lee, J., Feldman, A. R., Delmas, B., & Paetzel, M. (2007). Crystal structure of the VP4 protease from infectious pancreatic necrosis virus reveals the acyl-enzyme complex for an intermolecular self-cleavage reaction. *The Journal of biological chemistry*, *282*(34), 24928–24937.
- Lee, J., Worrall, L. J., Vuckovic, M., Rosell, F. I., Gentile, F., Ton, A. T., Caveney, N. A., Ban, F., Cherkasov, A., Paetzel, M., & Strynadka, N. C. J. (2020). Crystallographic structure of wild-type SARS-CoV-2 main protease acyl-enzyme intermediate with physiological C-terminal autoprocessing site. *Nature communications*, *11*(1), 5877.
- Lee, J. T., Yang, Q., Gribenko, A., Perrin, B. S., Jr, Zhu, Y., Cardin, R., Liberator, P. A., Anderson, A. S., & Hao, L. (2022). Genetic Surveillance of SARS-CoV-2 M^{PRO} Reveals High Sequence and Structural Conservation Prior to the Introduction of Protease Inhibitor Paxlovid. *mBio*, *13*(4), e0086922.
- Lei, J., Hansen, G., Nitsche, C., Klein, C. D., Zhang, L., & Hilgenfeld, R. (2016). Crystal structure of Zika virus NS2B-NS3 protease in complex with a boronate inhibitor. *Science (New York, N.Y.)*, *353*(6298), 503–505.
- Lei, J., Kusov, Y., & Hilgenfeld, R. (2018). Nsp3 of coronaviruses: Structures and functions of a large multi-domain protein. *Antiviral research*, *149*, 58–74.
- Lei, F., Shu, J., Xia, C., Lei, Y., Zhang, Y., Zhang, Y., Yuan, Z., & Yi, Z. (2025). Viral determinants of cis- and trans-cleavage by SARS-CoV-2 Nsp3 and an on-off reporter for monitoring intracellular protease activity. *Antiviral research*, *242*, 106262.
- Li, C., Qi, Y., Teng, X., Yang, Z., Wei, P., Zhang, C., Tan, L., Zhou, L., Liu, Y., & Lai, L. (2010). Maturation mechanism of severe acute respiratory syndrome (SARS) coronavirus 3C-like proteinase. *The Journal of biological chemistry*, *285*(36), 28134–28140.
- Louis, J. M., Aniana, A., Weber, I. T., Sayer, J. M. (2011). Inhibition of autoprocessing of natural variants and multidrug resistant mutant precursors of HIV-1 protease by

- clinical inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 108(22), 9072–9077.
- Louis, J. M., Nashed, N. T., Parris, K. D., Kimmel, A. R., Jerina, D. M. (1994). Kinetics and mechanism of autoprocessing of human immunodeficiency virus type 1 protease from an analog of the Gag-Pol polyprotein. *Proceedings of the National Academy of Sciences of the United States of America*, 91(17), 7970–7974.
- Louis, J. M., Dyda, F., Nashed, N. T., Kimmel, A. R., & Davies, D. R. (1998). Hydrophilic peptides derived from the transframe region of Gag-Pol inhibit the HIV-1 protease. *Biochemistry*, 37(8), 2105–2110.
- Louis, J.M., Wondrak, E.M., Kimmel, A.R., Wingfield, P.T., Nashed, N.T. (1999a). Proteolytic processing of HIV-1 protease precursor, kinetics and mechanism. *J Biol Chem*; 274(33): 23437-42.
- Louis, J.M., Clore, G.M., Gronenborn, A.M. (1999b). Autoprocessing of HIV-1 protease is tightly coupled to protein folding. *Nat Struct Biol*. 6(9): 868-75.
- Louis, J. M., Weber, I. T., Tózsér, J., Clore, G. M., Gronenborn, A. M. (2000). HIV-1 protease: maturation, enzyme specificity, and drug resistance. *Advances in pharmacology (San Diego, Calif.)*, 49, 111–146.
- Louis, J. M., Ishima, R., Torchia, D. A., Weber, I. T. (2007). HIV-1 protease: structure, dynamics, and inhibition. *Advances in pharmacology (San Diego, Calif.)*, 55, 261–298.
- Mahdi, M., Szojka, Z., Mótyán, J. A., Tózsér, J. (2015). Inhibition Profiling of Retroviral Protease Inhibitors Using an HIV-2 Modular System. *Viruses*, 7(12), 6152–6162.
- Mahmoudvand, S., & Shokri, S. (2021). Interactions between SARS coronavirus 2 papain-like protease and immune system: A potential drug target for the treatment of COVID-19. *Scandinavian journal of immunology*, 94(4), e13044.
- Majerová, T., Konvalinka, J. (2022). Viral proteases as therapeutic targets. *Molecular aspects of medicine*, 88, 101159.
- Majerová, T., Novotný, P. (2021). Precursors of Viral Proteases as Distinct Drug Targets. *Viruses*, 13(10), 1981.
- Matthews, D. A., Smith, W. W., Ferre, R. A., Condon, B., Budahazi, G., Sisson, W., Villafranca, J. E., Janson, C. A., McElroy, H. E., & Gribskov, C. L. (1994). Structure of human rhinovirus 3C protease reveals a trypsin-like polypeptide fold, RNA-binding site, and means for cleaving precursor polyprotein. *Cell*, 77(5), 761–771.
- McGrath, W. J., Graziano, V., Zabrocka, K., & Mangel, W. F. (2013). First generation inhibitors of the adenovirus proteinase. *FEBS letters*, 587(15), 2332–2339.
- Mildner, A. M., Rothrock, D. J., Leone, J. W., Bannow, C. A., Lull, J. M., Reardon, I. M., Sarcich, J. L., Howe, W. J., Tomich, C. S., & Smith, C. W. (1994). The HIV-1 protease as enzyme and substrate: mutagenesis of autolysis sites and generation of a stable mutant with retained kinetic properties. *Biochemistry*, 33(32), 9405–9413.
- Morazzani, E. M., Compton, J. R., Leary, D. H., Berry, A. V., Hu, X., Marugan, J. J., Glass, P. J., & Legler, P. M. (2019). Proteolytic cleavage of host proteins by the Group IV viral proteases of Venezuelan equine encephalitis virus and Zika virus. *Antiviral research*, 164, 106–122.
- Mótyán, J. A., Tóth, F., & Tózsér, J. (2013). Research applications of proteolytic enzymes in molecular biology. *Biomolecules*, 3(4), 923–942.
- Muramatsu, T., Kim, Y. T., Nishii, W., Terada, T., Shirouzu, M., & Yokoyama, S. (2013). Autoprocessing mechanism of severe acute respiratory syndrome coronavirus 3C-like protease (SARS-CoV 3CLpro) from its polyproteins. *The FEBS journal*, 280(9), 2002–2013.
- Narwal, M., Armache, J. P., Edwards, T. J., & Murakami, K. S. (2023). SARS-CoV-2 polyprotein substrate regulates the stepwise M^{Pro} cleavage reaction. *The Journal of biological chemistry*, 299(5), 104697.
- Nashed, N. T., Aniana, A., Ghirlando, R., Chiliveri, S. C., & Louis, J. M. (2022). Modulation of the monomer-dimer equilibrium and catalytic activity of SARS-CoV-2 main protease by a transition-state analog inhibitor. *Communications biology*, 5(1), 160.
- Nashed, N. T., Kneller, D. W., Coates, L., Ghirlando, R., Aniana, A., Kovalevsky, A., & Louis, J. M. (2022a). Autoprocessing and oxyanion loop reorganization upon GC373 and

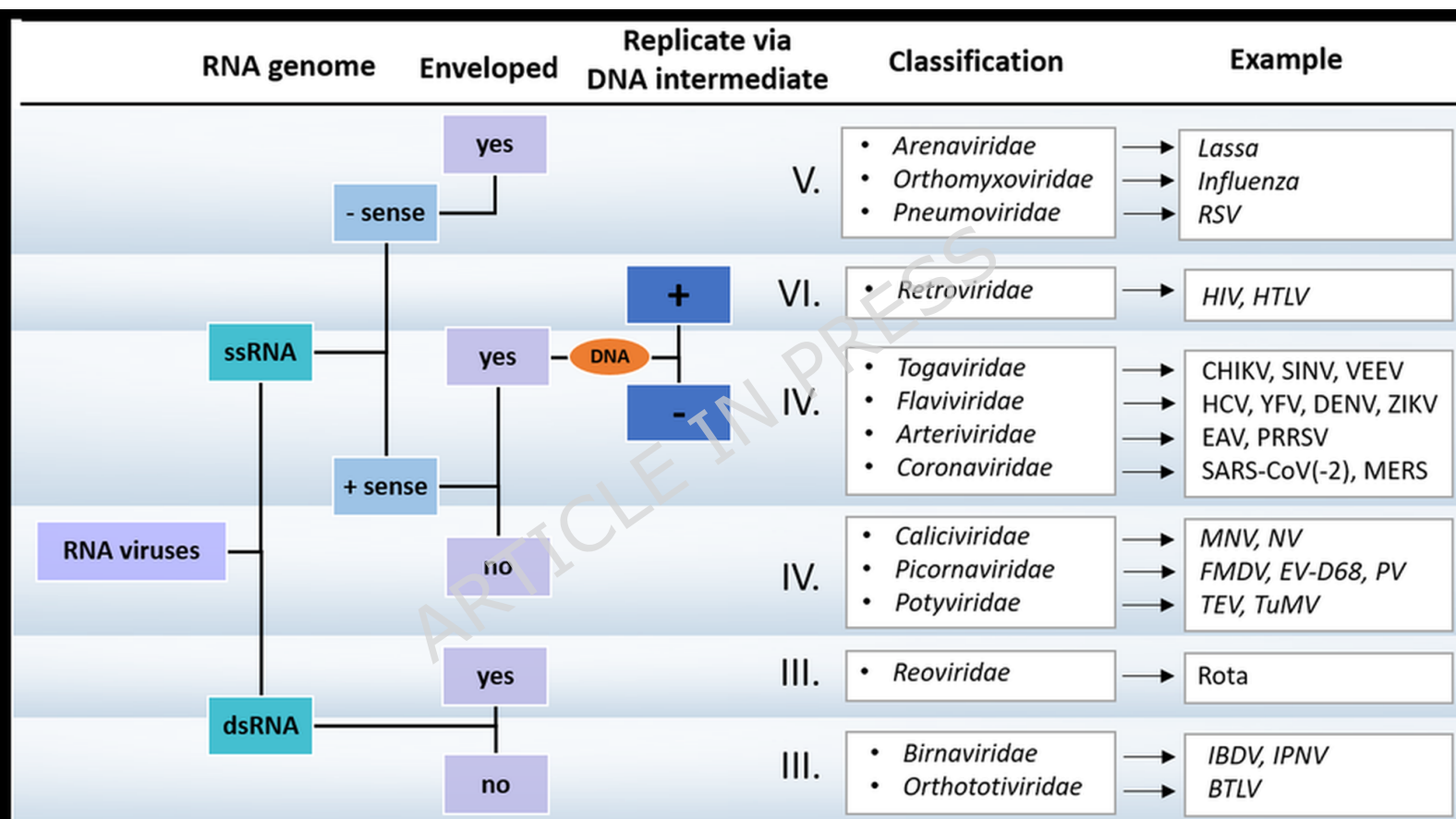
- nirmatrelvir binding of monomeric SARS-CoV-2 main protease catalytic domain. *Communications biology*, 5(1), 976.
- Nashed, N. T., Aniana, A., Ghirlando, R., Chiliveri, S. C., & Louis, J. M. (2022b). Modulation of the monomer-dimer equilibrium and catalytic activity of SARS-CoV-2 main protease by a transition-state analog inhibitor. *Communications biology*, 5(1), 160.
- Neufeldt, C. J., Cortese, M., Acosta, E. G., & Bartenschlager, R. (2018). Rewiring cellular networks by members of the Flaviviridae family. *Nature reviews. Microbiology*, 16(3), 125-142.
- Noske, G. D., Nakamura, A. M., Gawriljuk, V. O., Fernandes, R. S., Lima, G. M. A., Rosa, H. V. D., Pereira, H. D., Zeri, A. C. M., Nascimento, A. F. Z., Freire, M. C. L. C., Fearon, D., Douangamath, A., von Delft, F., Oliva, G., & Godoy, A. S. (2021). A Crystallographic Snapshot of SARS-CoV-2 Main Protease Maturation Process. *Journal of molecular biology*, 433(18), 167118.
- Novotný, P., Humpolíčková, J., Nováková, V., Stanchev, S., Stříšovský, K., Zgarbová, M., Weber, J., Kryštůfek, R., Starková, J., Hradilek, M., Moravcová, A., Güntherová, J., Bach, K., Majer, P., Konvalinka, J., Majerová, T. (2025). The zymogenic form of SARS-CoV-2 main protease: A discrete target for drug discovery. *The Journal of biological chemistry*, 301(1), 108079.
- Oroszlan, S., & Luftig, R. B. (1990). Retroviral proteinases. *Current topics in microbiology and immunology*, 157, 153-185.
- Osipiuk, J., Azizi, S. A., Dvorkin, S., Endres, M., Jedrzejczak, R., Jones, K. A., Kang, S., Kathayat, R. S., Kim, Y., Lisnyak, V. G., Maki, S. L., Nicolaescu, V., Taylor, C. A., Tesar, C., Zhang, Y. A., Zhou, Z., Randall, G., Michalska, K., Snyder, S. A., Dickinson, B. C., ... Joachimiak, A. (2021). Structure of papain-like protease from SARS-CoV-2 and its complexes with non-covalent inhibitors. *Nature communications*, 12(1), 743.
- Park, J. H., Sayer, J. M., Aniana, A., Yu, X., Weber, I. T., Harrison, R. W., & Louis, J. M. (2016). Binding of Clinical Inhibitors to a Model Precursor of a Rationally Selected Multidrug Resistant HIV-1 Protease Is Significantly Weaker Than That to the Released Mature Enzyme. *Biochemistry*, 55(16), 2390-2400.
- Patchett, S., Lv, Z., Rut, W., Békés, M., Drag, M., Olsen, S. K., & Huang, T. T. (2021). A molecular sensor determines the ubiquitin substrate specificity of SARS-CoV-2 papain-like protease. *Cell reports*, 36(13), 109754.
- Pathak, N., Kuo, Y. P., Chang, T. Y., Huang, C. T., Hung, H. C., Hsu, J. T., Yu, G. Y., & Yang, J. M. (2020). Zika Virus NS3 Protease Pharmacophore Anchor Model and Drug Discovery. *Scientific reports*, 10(1), 8929.
- Patick, A. K., Brothers, M. A., Maldonado, F., Binford, S., Maldonado, O., Fuhrman, S., Petersen, A., Smith, G. J., 3rd, Zalman, L. S., Burns-Naas, L. A., & Tran, J. Q. (2005). In vitro antiviral activity and single-dose pharmacokinetics in humans of a novel, orally bioavailable inhibitor of human rhinovirus 3C protease. *Antimicrobial agents and chemotherapy*, 49(6), 2267-2275.
- Paulus, C., Hellebrand, S., Tessmer, U., Wolf, H., Kräusslich, H. G., & Wagner, R. (1999). Competitive inhibition of human immunodeficiency virus type-1 protease by the Gag-Pol transframe protein. *The Journal of biological chemistry*, 274(31), 21539-21543.
- Petit, S., Lejal, N., Huet, J. C., & Delmas, B. (2000). Active residues and viral substrate cleavage sites of the protease of the birnavirus infectious pancreatic necrosis virus. *Journal of virology*, 74(5), 2057-2066.
- Pettit, S.C., Gulnik, S., Everitt, L., Kaplan, A.H. (2003) The dimer interfaces of protease and extra-protease domains influence the activation of protease and the specificity of GagPol cleavage. *J. Virol.* 77 (1), 366-374.
- Pettit, S.C., Everitt, L.E., Choudhury, S., Dunn, B.M., Kaplan, A.H. (2004) Initial cleavage of the human immunodeficiency virus type 1 GagPol precursor by its activated protease occurs by an intramolecular mechanism. *Journal of virology*, 78 (16), 8477-8485.
- Pettit, S.C., Clemente, J.C., Jeung, J.A., Dunn, B.M., Kaplan, A.H. (2005) Ordered processing of the human immunodeficiency virus type 1 GagPol precursor is

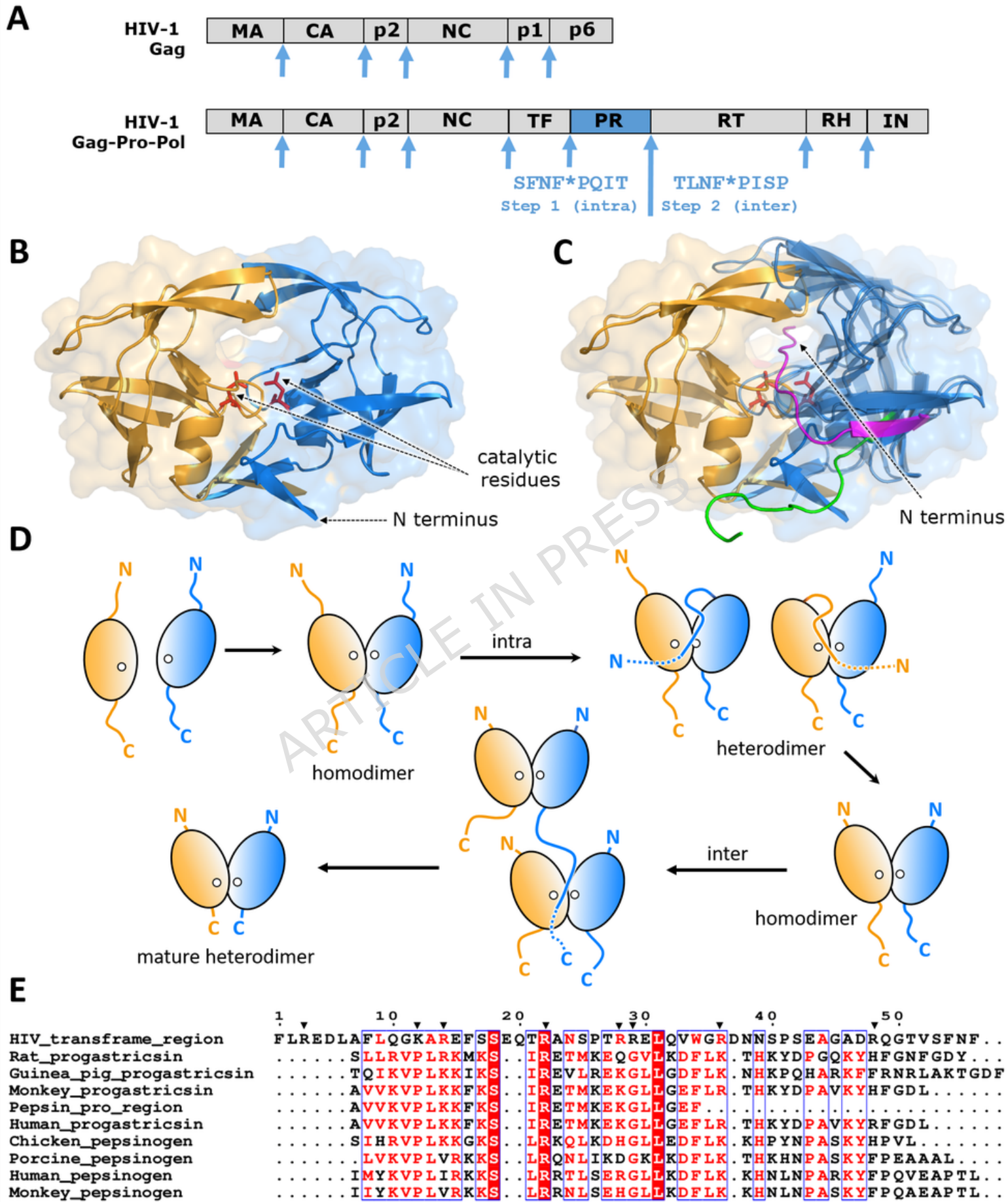
- influenced by the context of the embedded viral protease. *Journal of virology*, 79 (16), 10601–10607.
- Phoo, W. W., Li, Y., Zhang, Z., Lee, M. Y., Loh, Y. R., Tan, Y. B., Ng, E. Y., Lescar, J., Kang, C., & Luo, D. (2016). Structure of the NS2B-NS3 protease from Zika virus after self-cleavage. *Nature communications*, 7, 13410.
- Pierce, D. M., Hayward, C., Rowlands, D. J., Stonehouse, N. J., & Herod, M. R. (2023). Insights into Polyprotein Processing and RNA-Protein Interactions in Foot-and-Mouth Disease Virus Genome Replication. *Journal of virology*, 97(5), e0017123.
- Prabu-Jeyabalan, M., Nalivaika, E., & Schiffer, C. A. (2002). Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal structures of six substrate complexes. *Structure (London, England: 1993)*, 10(3), 369–381.
- Preugschat, F., Yao, C. W., & Strauss, J. H. (1990). In vitro processing of dengue virus type 2 nonstructural proteins NS2A, NS2B, and NS3. *Journal of virology*, 64(9), 4364–4374.
- Qin, Z. L., Yao, Q. F., Ren, H., Zhao, P., & Qi, Z. T. (2022). Lipid Droplets and Their Participation in Zika Virus Infection. *International journal of molecular sciences*, 23(20), 12584.
- Rampersad, S., & Tennant, P. (2018). Replication and Expression Strategies of Viruses. *Viruses*, 55–82.
- Ren, D., Ye, X., Chen, R., Jia, X., He, X., Tao, J., Jin, T., Wu, S., & Zhang, H. (2025). Activation and evasion of inflammasomes during viral and microbial infection. *Cellular and molecular life sciences : CMLS*, 82(1), 56.
- Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic acids research*, 42(Web Server issue), W320–W324.
- Romano, M., Ruggiero, A., Squeglia, F., Maga, G., & Berisio, R. (2020). A Structural View of SARS-CoV-2 RNA Replication Machinery: RNA Synthesis, Proofreading and Final Capping. *Cells*, 9(5), 1267.
- Russo, A. T., White, M. A., & Watowich, S. J. (2006). The crystal structure of the Venezuelan equine encephalitis alphavirus nsP2 protease. *Structure (London, England: 1993)*, 14(9), 1449–1458.
- Sabharwal, P., & Savithri, H. S. (2020). Functional Characterization of Pepper Vein Banding Virus-Encoded Proteins and Their Interactions: Implications in *Potyvirus* Infection. *Viruses*, 12(9), 1037.
- Sadiq, S. K., Noé, F., & De Fabritiis, G. (2012). Kinetic characterization of the critical step in HIV-1 protease maturation. *Proceedings of the National Academy of Sciences of the United States of America*, 109(50), 20449–20454.
- Sayer, J. M., Aniana, A., & Louis, J. M. (2012). Mechanism of dissociative inhibition of HIV protease and its autoprocessing from a precursor. *Journal of molecular biology*, 422(2), 230–244.
- Scheffler, U., Rudolph, W., Gebhardt, J., & Rohayem, J. (2007). Differential cleavage of the norovirus polyprotein precursor by two active forms of the viral protease. *The Journal of general virology*, 88(Pt 7), 2013–2018.
- Shultz, M. D., Ham, Y. W., Lee, S. G., Davis, D. A., Brown, C., & Chmielewski, J. (2004). Small-molecule dimerization inhibitors of wild-type and mutant HIV protease: a focused library approach. *Journal of the American Chemical Society*, 126(32), 9886–9887.
- Shaqra, A. M., Zvornicanin, S. N., Huang, Q. Y. J., Lockbaum, G. J., Knapp, M., Tandeske, L., Bakan, D. T., Flynn, J., Bolon, D. N. A., Moquin, S., Dovala, D., Kurt Yilmaz, N., & Schiffer, C. A. (2022). Defining the substrate envelope of SARS-CoV-2 main protease to predict and avoid drug resistance. *Nature communications*, 13(1), 3556.
- Shimba, N., Nomura, A. M., Marnett, A. B., & Craik, C. S. (2004). Herpesvirus protease inhibition by dimer disruption. *Journal of virology*, 78(12), 6657–6665.
- Shin, G., Yost, S. A., Miller, M. T., Elrod, E. J., Grakoui, A., & Marcotrigiano, J. (2012). Structural and functional insights into alphavirus polyprotein processing and pathogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 109(41), 16534–16539.

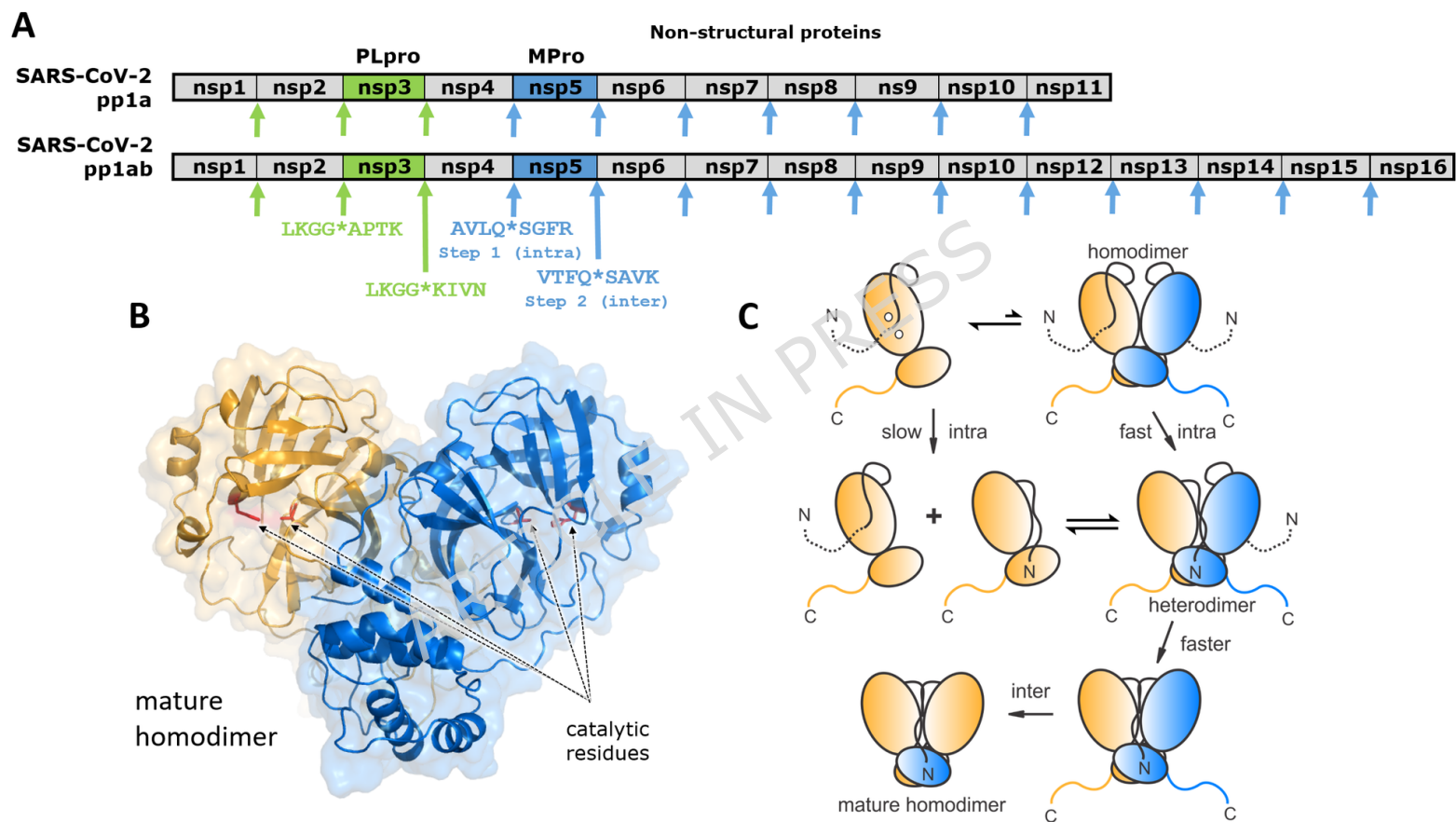
- Shiryaev, S. A., Farhy, C., Pinto, A., Huang, C. T., Simonetti, N., Elong Ngono, A., Dewing, A., Shresta, S., Pinkerton, A. B., Cieplak, P., Strongin, A. Y., & Terskikh, A. V. (2017). Characterization of the Zika virus two-component NS2B-NS3 protease and structure-assisted identification of allosteric small-molecule antagonists. *Antiviral research, 143*, 218-229.
- Snijder, E. J., Wassenaar, A. L., & Spaan, W. J. (1994). Proteolytic processing of the replicase ORF1a protein of equine arteritis virus. *Journal of virology, 68*(9), 5755-5764.
- Snijder, E. J., Wassenaar, A. L., van Dinten, L. C., Spaan, W. J., & Gorbalenya, A. E. (1996). The arterivirus nsp4 protease is the prototype of a novel group of chymotrypsin-like enzymes, the 3C-like serine proteases. *The Journal of biological chemistry, 271*(9), 4864-4871.
- Solis, M., Nakhaei, P., Jalalirad, M., Lacoste, J., Douville, R., Arguello, M., Zhao, T., Laughrea, M., Wainberg, M. A., & Hiscott, J. (2011). RIG-I-mediated antiviral signaling is inhibited in HIV-1 infection by a protease-mediated sequestration of RIG-I. *Journal of virology, 85*(3), 1224-1236.
- Someya, Y., Takeda, N., & Miyamura, T. (2002). Identification of active-site amino acid residues in the Chiba virus 3C-like protease. *Journal of virology, 76*(12), 5949-5958.
- Sosnovtsev, S. V., Belliot, G., Chang, K. O., Prikhodko, V. G., Thackray, L. B., Wobus, C. E., Karst, S. M., Virgin, H. W., & Green, K. Y. (2006). Cleavage map and proteolytic processing of the murine norovirus nonstructural polyprotein in infected cells. *Journal of virology, 80*(16), 7816-7831.
- Spall, V.E., Shanks, M., Lomonosoff, G.P. (1997) Polyprotein Processing as a Strategy for Gene Expression in RNA Viruses. *Seminars in Virology, 8*(1), 15-23.
- Starvaggi, J., Previti, S., Zappalà, M., & Ettari, R. (2024). The Inhibition of NS2B/NS3 Protease: A New Therapeutic Opportunity to Treat Dengue and Zika Virus Infection. *International journal of molecular sciences, 25*(8), 4376.
- Sun, D., Chen, S., Cheng, A., & Wang, M. (2016). Roles of the Picornaviral 3C Proteinase in the Viral Life Cycle and Host Cells. *Viruses, 8*(3), 82.
- Sun, Z., Wang, L., Li, X., Fan, C., Xu, J., Shi, Z., Qiao, H., Lan, Z., Zhang, X., Li, L., Zhou, X., & Geng, Y. (2022). An extended conformation of SARS-CoV-2 main protease reveals allosteric targets. *Proceedings of the National Academy of Sciences of the United States of America, 119*(15), e2120913119.
- Syed Y. Y. (2024). Ensitrelvir Fumaric Acid: First Approval. *Drugs, 84*(6), 721-728.
- Sudo, S., Haraguchi, H., Hirai, Y., Gatanaga, H., Sakuragi, J., Momose, F., & Morikawa, Y. (2013). Efavirenz enhances HIV-1 gag processing at the plasma membrane through Gag-Pol dimerization. *Journal of virology, 87*(6), 3348-3360.
- Tabler, C. O., Wegman, S. J., Chen, J., Shroff, H., Alhusaini, N., Tilton, J. C. (2022). The HIV-1 Viral Protease Is Activated during Assembly and Budding Prior to Particle Release. *Journal of virology, 96*(9), e0219821.
- Tabler, C. O., Tilton, J. C. (2024). Mechanism and Kinetics of HIV-1 Protease Activation. *Viruses, 16*(12), 1826.
- Tang, C., Louis, J. M., Aniana, A., Suh, J. Y., & Clore, G. M. (2008). Visualizing transient events in amino-terminal autoprocessing of HIV-1 protease. *Nature, 455*(7213), 693-696.
- Teo, K. F., & Wright, P. J. (1997). Internal proteolysis of the NS3 protein specified by dengue virus 2. *The Journal of general virology, 78* (Pt 2), 337-341.
- Tien, C., Huang, L., Watanabe, S. M., Speidel, J. T., Carter, C. A., & Chen, C. (2018). Context-dependent autoprocessing of human immunodeficiency virus type 1 protease precursors. *PloS one, 13*(1), e0191372.
- Tomasselli, A. G., Heinrikson, R. L. (1994). Specificity of retroviral proteases: an analysis of viral and nonviral protein substrates. *Methods in enzymology, 241*, 279-301.
- Tomei, L., Failla, C., Santolini, E., De Francesco, R., & La Monica, N. (1993). NS3 is a serine protease required for processing of hepatitis C virus polyprotein. *Journal of virology, 67*(7), 4017-4026.
- Tong L. (2002). Viral proteases. *Chemical reviews, 102*(12), 4609-4626.
- Tözsér J. (2010). Comparative studies on retroviral proteases: substrate specificity. *Viruses, 2*(1), 147-165.

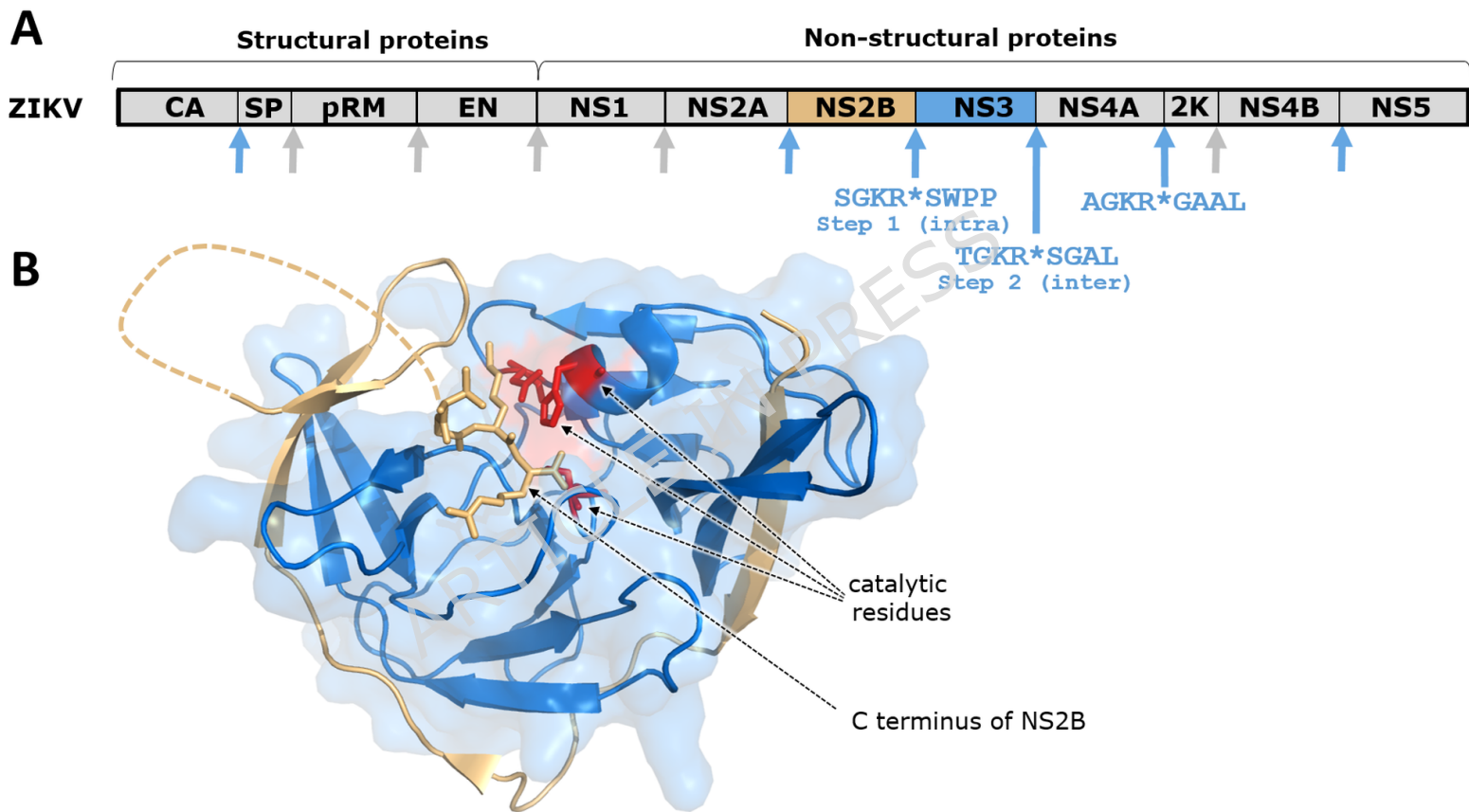
- Tözsér, J., Bagossi, P., Weber, I. T., Louis, J. M., Copeland, T. D., & Oroszlan, S. (1997). Studies on the symmetry and sequence context dependence of the HIV-1 proteinase specificity. *The Journal of biological chemistry*, *272*(27), 16807-16814.
- Tözsér, J., & Oroszlan, S. (2003). Proteolytic events of HIV-1 replication as targets for therapeutic intervention. *Current pharmaceutical design*, *9*(22), 1803-1815.
- Tsu, B. V., Fay, E. J., Nguyen, K. T., Corley, M. R., Hosuru, B., Dominguez, V. A., & Daugherty, M. D. (2021). Running With Scissors: Evolutionary Conflicts Between Viral Proteases and the Host Immune System. *Frontiers in immunology*, *12*, 769543.
- Ullrich, S., & Nitsche, C. (2022). SARS-CoV-2 Papain-Like Protease: Structure, Function and Inhibition. *Chembiochem : a European journal of chemical biology*, *23*(19), e202200327.
- Urcuqui-Inchima, S., Haenni, A. L., & Bernardi, F. (2001). Potyvirus proteins: a wealth of functions. *Virus research*, *74*(1-2), 157-175. [https://doi.org/10.1016/s0168-1702\(01\)00220-9](https://doi.org/10.1016/s0168-1702(01)00220-9)
- Vasiljeva, L., Merits, A., Golubtsov, A., Sizemskaja, V., Kääriäinen, L., & Ahola, T. (2003). Regulation of the sequential processing of Semliki Forest virus replicase polyprotein. *The Journal of biological chemistry*, *278*(43), 41636-41645.
- von Hammerstein, F., Lauth, L. M., Hammerschmidt, S., Wagner, A., Schirmeister, T., & Hellmich, U. A. (2019). Cis autocatalytic cleavage of glycine-linked Zika virus NS2B-NS3 protease constructs. *FEBS letters*, *593*(16), 2204-2213.
- Wagner, R. N., Reed, J. C., & Chanda, S. K. (2015). HIV-1 protease cleaves the serine-threonine kinases RIPK1 and RIPK2. *Retrovirology*, *12*, 74.
- Wahaab, A., Mustafa, B. E., Hameed, M., Stevenson, N. J., Anwar, M. N., Liu, K., Wei, J., Qiu, Y., & Ma, Z. (2021). Potential Role of Flavivirus NS2B-NS3 Proteases in Viral Pathogenesis and Anti-flavivirus Drug Discovery Employing Animal Cells and Models: A Review. *Viruses*, *14*(1), 44.
- Wahaab, A., Mustafa, B. E., Hameed, M., Batool, H., Tran Nguyen Minh, H., Tawaab, A., Shoaib, A., Wei, J., & Rasgon, J. L. (2024). An Overview of Zika Virus and Zika Virus Induced Neuropathies. *International journal of molecular sciences*, *26*(1), 47.
- Wang, Q., Gao, H., Clark, K. M., Mugisha, C. S., Davis, K., Tang, J. P., Harlan, G. H., DeSelm, C. J., Presti, R. M., Kutluay, S. B., & Shan, L. (2021). CARD8 is an inflammasome sensor for HIV-1 protease activity. *Science (New York, N.Y.)*, *371*(6535), eabe1707.
- Wassenaar, A. L., Spaan, W. J., Gorbalenya, A. E., & Snijder, E. J. (1997). Alternative proteolytic processing of the arterivirus replicase ORF1a polyprotein: evidence that NSP2 acts as a cofactor for the NSP4 serine protease. *Journal of virology*, *71*(12), 9313-9322.
- Waugh D. S. (2011). An overview of enzymatic reagents for the removal of affinity tags. *Protein expression and purification*, *80*(2), 283-293.
- Weber, I.T. (1990). Comparison of the crystal structures and intersubunit interactions of human immunodeficiency and Rous sarcoma virus proteases. *The Journal of biological chemistry*, *265*(18), 10492-10496.
- Wieggers, K., Rutter, G., Kottler, H., Tessmer, U., Hohenberg, H., & Kräusslich, H. G. (1998). Sequential steps in human immunodeficiency virus particle maturation revealed by alterations of individual Gag polyprotein cleavage sites. *Journal of virology*, *72*(4), 2846-2854.
- Wondrak, E. M., Nashed, N. T., Haber, M. T., Jerina, D. M., & Louis, J. M. (1996). A transient precursor of the HIV-1 protease. Isolation, characterization, and kinetics of maturation. *The Journal of biological chemistry*, *271*(8), 4477-4481.
- Wondrak, E. M., & Louis, J. M. (1996). Influence of flanking sequences on the dimer stability of human immunodeficiency virus type 1 protease. *Biochemistry*, *35*(39), 12957-12962.
- Wong, N. A., & Saier, M. H., Jr (2021). The SARS-Coronavirus Infection Cycle: A Survey of Viral Membrane Proteins, Their Functional Interactions and Pathogenesis. *International journal of molecular sciences*, *22*(3), 1308.
- Wu, J., Bera, A. K., Kuhn, R. J., & Smith, J. L. (2005). Structure of the Flavivirus helicase: implications for catalytic activity, protein interactions, and proteolytic processing. *Journal of virology*, *79*(16), 10268-10277.

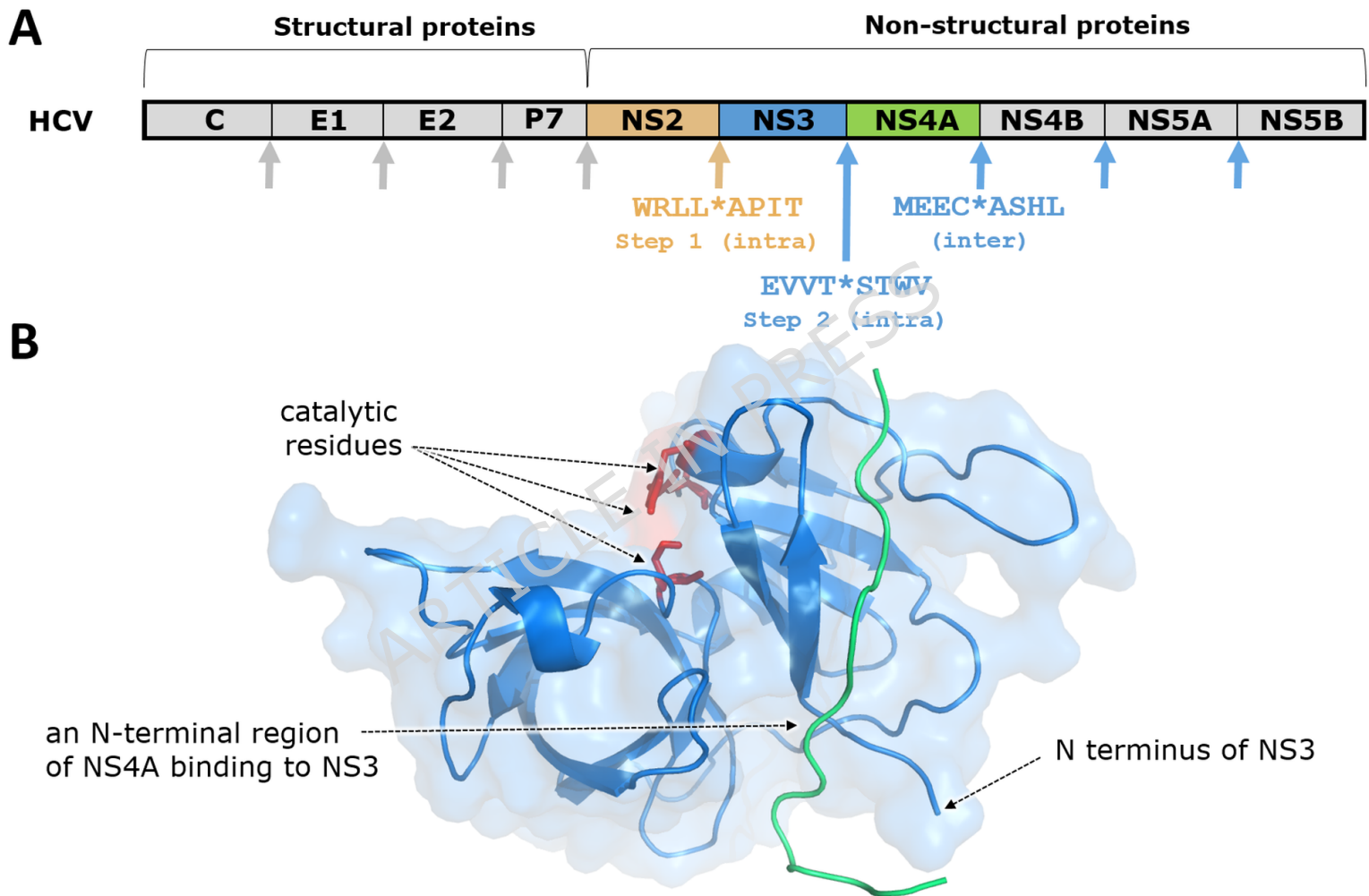
- Xia, B., & Kang, X. (2011). Activation and maturation of SARS-CoV main protease. *Protein & cell*, 2(4), 282-290.
- Xiao, H., Lord, E., & Sanfaçon, H. (2022). Proteolytic Processing of Plant Proteins by Potyvirus NIa Proteases. *Journal of virology*, 96(2), e0144421.
- Yan, S., & Wu, G. (2021). Spatial and temporal roles of SARS-CoV PL^{pro} -A snapshot. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, 35(1), e21197.
- Yang, Y., Zhao, Z., Li, X., Chen, Y., Liu, L., Zhang, S. L., & Yang, A. (2025). Design and Evaluation of Peptide Inhibitors Targeting the Dimerization of SARS-CoV-2 Main Protease. *Chembiochem: a European journal of chemical biology*, 26(1), e202400688.
- Yost, S. A., & Marcotrigiano, J. (2013). Viral precursor polyproteins: keys of regulation from replication to maturation. *Current opinion in virology*, 3(2), 137-142.
- Yu, F. H., Huang, K. J., & Wang, C. T. (2022). Amino acid substitutions at the HIV-1 transframe region significantly impair virus infectivity. *PloS one*, 17(1), e0262477.
- Zhang, Z., Li, Y., Loh, Y. R., Phoo, W. W., Hung, A. W., Kang, C., & Luo, D. (2016). Crystal structure of unlinked NS2B-NS3 protease from Zika virus. *Science (New York, N.Y.)*, 354(6319), 1597-1600.
- Zephyr, J., Kurt Yilmaz, N., & Schiffer, C. A. (2021). Viral proteases: Structure, mechanism and inhibition. *The Enzymes*, 50, 301-333.
- Zunzain, P. A., Knox, S. R., Sweeney, T. R., Yang, J., Roqué-Rosell, N., Belsham, G. J., Leatherbarrow, R. J., & Curry, S. (2010). Insights into cleavage specificity from the crystal structure of foot-and-mouth disease virus 3C protease complexed with a peptide substrate. *Journal of molecular biology*, 395(2), 375-389.
- Zutshi, R., Franciskovich, J., Shultz, M., Schweitzer, B., Bishop, P., Wilson, M., Chmielewski J. (1997). Targeting the Dimerization Interface of HIV-1 Protease: Inhibition with Cross-Linked Interfacial Peptides. *Journal of the American Chemical Society*, 119 (21), 4841-4845.

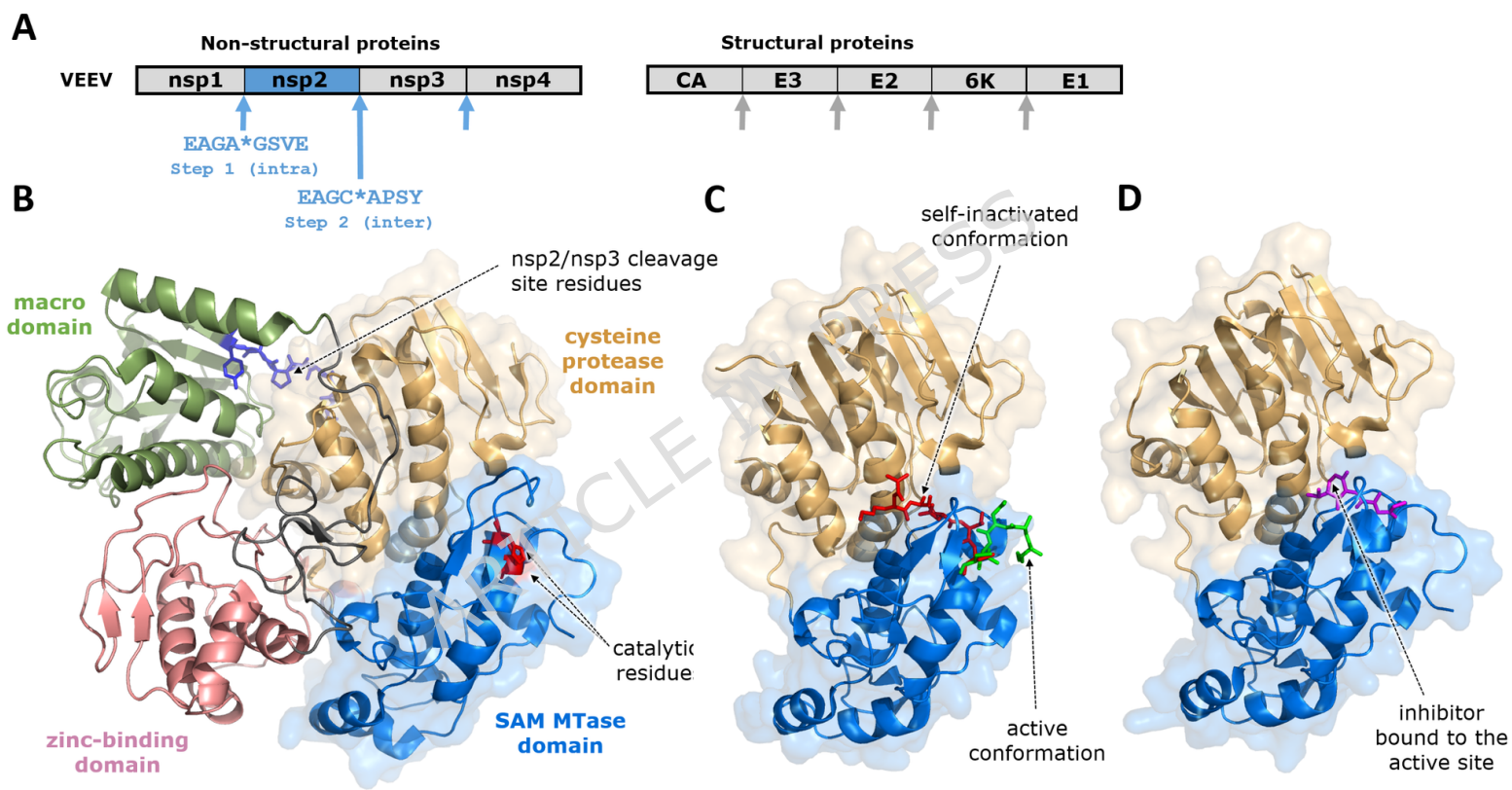


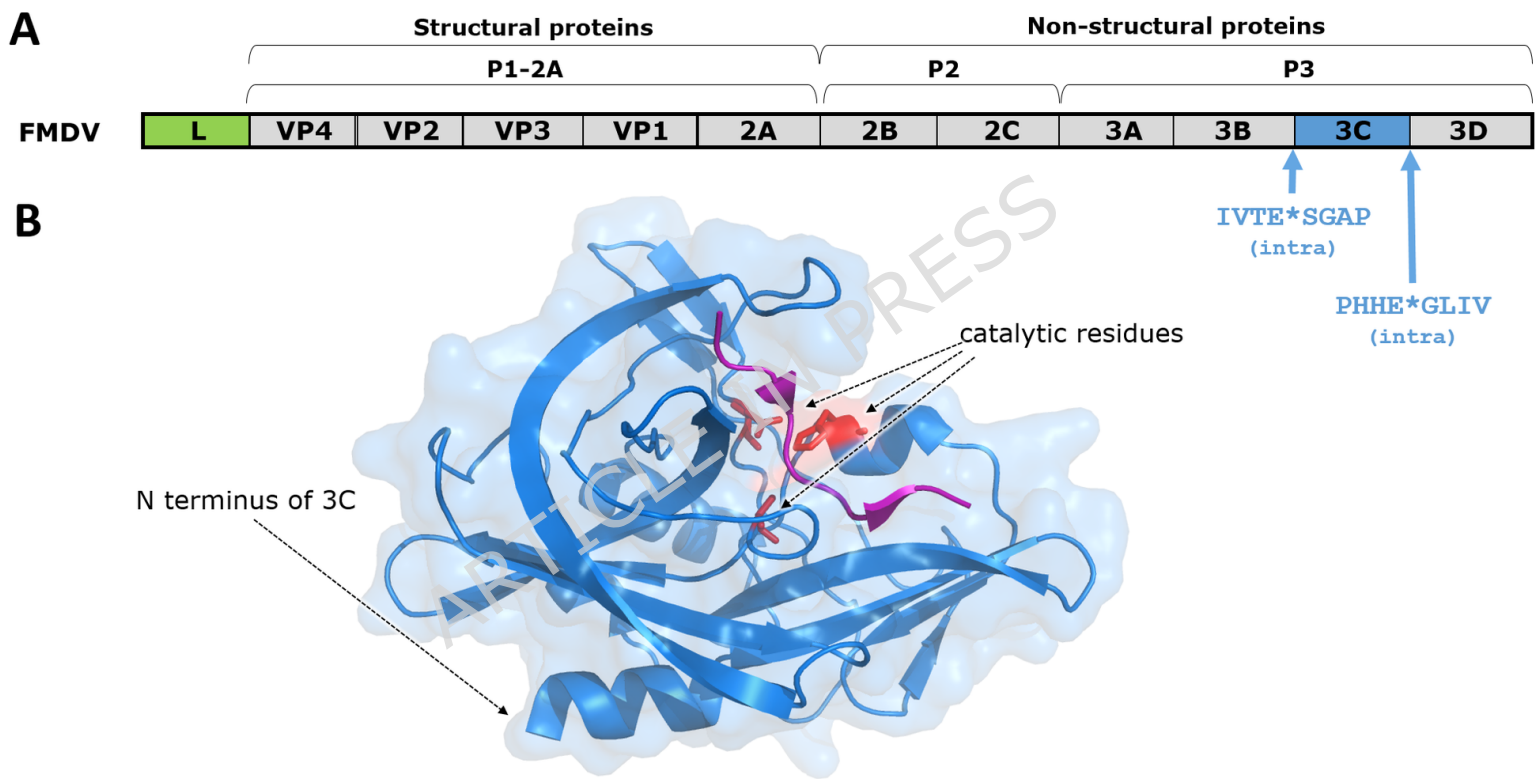


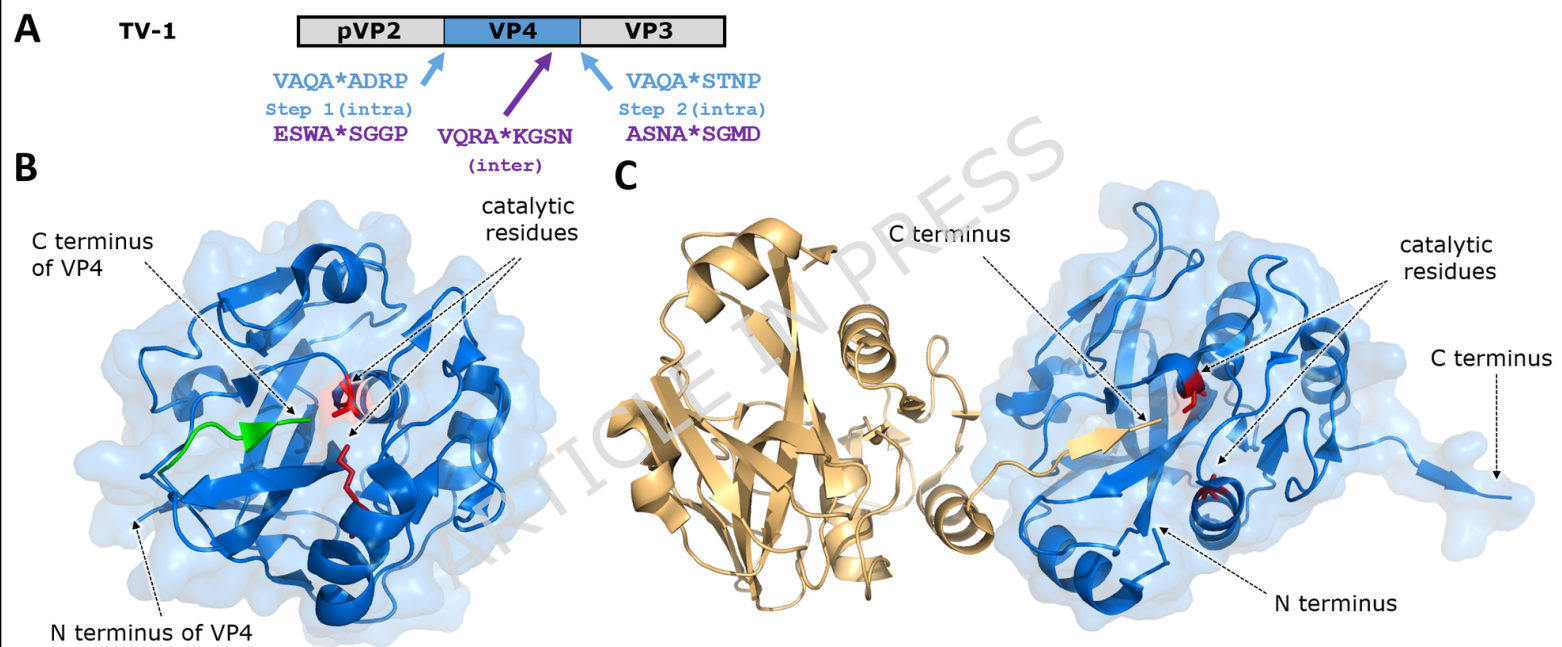












		Envelope	Virus	Protease	Cleavage N-term. / C-term.
ssRNA	ss(-)RNA	yes	• Lassa		
			• Influenza		
			• RSV		
	ss(+)RNA	yes	• HIV		
			• CHIKV, SINV, VEEV	nsP2	
			• HCV	NS3	
			• YFV, JEV, DENV, WNV, ZIKV	NS3	
			• EAV, PRRSV	nsp4	
			• SARS-CoV(-2)	Mpro	
			• MNV, NV	NS6	
	no	• FMDV, EV-D68, PV	3C pro		
		• TEV, TuMV	NIa-Pro		
• Rota					
dsRNA	yes	• Rota			
	no	• IBDV, IPNV	VP4		
		• BTLV			

Protease

Aspartic protease

Serine protease

Cysteine protease

No viral protease

Cleavage

Intramolecular
(*in cis*)
cleavage



Intermolecular
(*in trans*)
cleavage

