Running With Scissors: Evolutionary Conflicts Between Viral Proteases and the Host Immune System

Brian V. Tsu†, Elizabeth J. Fay†, Katelyn T. Nguyen, Miles R. Corley, Bindhu Hosuru, Viviana A. Dominguez and Matthew D. Daugherty*

Division of Biological Sciences, University of California, San Diego, CA, United States

Many pathogens encode proteases that serve to antagonize the host immune system. In particular, viruses with a positive-sense single-stranded RNA genome [(+)ssRNA], including picornaviruses, flaviviruses, and coronaviruses, encode proteases that are not only required for processing viral polyproteins into functional units but also manipulate crucial host cellular processes through their proteolytic activity. Because these proteases must cleave numerous polyprotein sites as well as diverse host targets, evolution of these viral proteases is expected to be highly constrained. However, despite this strong evolutionary constraint, mounting evidence suggests that viral proteases such as picornavirus 3C, flavivirus NS3, and coronavirus 3CL, are engaged in molecular ‘arms races’ with their targeted host factors, resulting in host- and virus-specific determinants of protease cleavage. In cases where protease-mediated cleavage results in host immune inactivation, recurrent host gene evolution can result in avoidance of cleavage by viral proteases. In other cases, such as recently described examples in NLRP1 and CARD8, hosts have evolved ‘tripwire’ sequences that mimic protease cleavage sites and activate an immune response upon cleavage. In both cases, host evolution may be responsible for driving viral protease evolution, helping explain why viral proteases and polyprotein sites are divergent among related viruses despite such strong evolutionary constraint. Importantly, these evolutionary conflicts result in diverse protease-host interactions even within closely related host and viral species, thereby contributing to host range, zoonotic potential, and pathogenicity of viral infection. Such examples highlight the importance of examining viral protease-host interactions through an evolutionary lens.

Keywords: viral proteases, host-virus evolution, innate antiviral immunity, molecular arms races, effector-triggered immunity, inflammasome

INTRODUCTION

Positive-sense single-stranded RNA [(+)ssRNA, see Table 1 for glossary of abbreviations] viruses represent the largest group of RNA viruses, spanning 30 divergent viral families that include important human pathogens in Flaviviridae, Picornaviridae, and Coronaviridae such as dengue virus, poliovirus, and SARS-CoV-2 (1). Despite their diversity, many viruses in this group share a common replication strategy: their (+)ssRNA viral genomes are delivered to host cells as a
translation-ready mRNA that encodes a multidomain viral polyprotein. Following translation of the viral polyprotein by host ribosomes, one or more embedded viral proteases cleave the polyprotein into individual, functional proteins at numerous sequence-specific positions (Figure 1A). Polyprotein cleavage at these specific sites is necessary for sustained virus replication and propagation, making viral proteases an attractive target for development of antiviral therapeutics (2, 3).

In addition to their essential role in the viral life cycle, (+) ssRNA viral proteases also cleave host proteins to manipulate host processes, including the host innate antiviral immune response (4). Importantly, host targets are cleaved with the same sequence specificity as sites within the viral polyprotein (Figure 1B). These dual roles place viral proteases at the intersection of two opposing selective pressures. On one side, the virus and its polyprotein site targets are under strong
pressure to be conserved, as any changes to the protease sequence specificity or protease sites without concomitant changes to the other would be deleterious for viral fitness. On the other side, viral fitness may be expected to benefit from a protease’s ability to adapt to and cleave new host targets, newly evolved sequences in the same host, or divergent sequences in a different host to facilitate cross-species transmission. This type of direct engagement between viral proteases and host factors thus generates an evolutionary conflict where both sides may be driven to adapt in a type of escalating molecular ‘arms race’ (5–8).

Molecular arms races exist as a result of the competing evolutionary interests of viruses and their hosts. Such competing interests establish an evolutionary equilibrium that is characterized by cyclical adaptations that exemplify a so-called ‘Red Queen’ genetic conflict (9). In these cases, viral adaptations that allow for successful infection of a host will provide a temporary advantage to the virus. However, host adaptations may restore the advantage to the host, applying selection pressure back to the virus (Figure 2A). Thus, molecular interactions between viruses and their hosts, particularly those interactions that contribute to potentiation or inhibition of virus...
replication, are shaped by immense evolutionary pressure on both parties: hosts are driven to both maintain interactions that activate or carry out antiviral defenses and evade virus interactions that prevent these responses, and viruses are driven to do the opposite. The result is recurrent adaptation of both virus and host to promote either virus replication or host antiviral mechanisms, respectively (Figure 2A) (5–8). Due to the fact that the direct molecular host-virus interfaces are those that are being remodeled during such molecular arms races, single amino acid changes can change the outcome of these conflicts (5, 7). Indeed, traces of these host-virus conflicts can be detected in host genomes by identifying gene codons that show evolutionary signatures of recurrent diversifying (positive) selection (5, 7, 10). Similarly, viruses are known to adapt during or following cross-species transmission to a novel host, and such adaptations can also be characterized by signatures of positive selection in viral genomes (11–14). Importantly, whether the host has evolved to the virus or the virus has evolved to the host, the resulting genetic and molecular changes determine the host range and pathogenesis of viruses, including influencing the ability of viruses to zoonotically transmit into the human population (7, 8, 13).

Due to the importance of sequence specificity to protease-host interactions, evolutionary arms races at the interfaces of proteases and their targets would be expected to exist. For instance, a single amino acid change in a targeted host protein at a position that is important for sequence-specific protease cleavage could completely reverse cleavage susceptibility (Figure 2B). As a result, single lineage-specific changes at any number of positions in the cleavage motif would be expected to alter cleavage susceptibility even among closely related hosts, establishing species-specific host-virus interactions that could drive viral host range (Figure 2C). Indeed, while a great deal of research on viral proteases has focused on conserved elements of protease function, emerging evidence suggests that both hosts and viruses are evolving in ways that can impact the host- and virus-specificity of cleavage. Here, we review the host-viral molecular conflicts engaged by the main proteases of flaviviruses, picornaviruses, and coronaviruses to emphasize how proteases of (+)ssRNA viruses act as evolutionary drivers of host innate immunity, and how viral proteases are being shaped by these same molecular conflicts. This evolutionary perspective highlights the importance of viral proteases and their host targets as being an important determinant of viral host range, tissue tropism and pathogenesis, and zoonotic potential of (+)ssRNA viruses.

DESPITE EVOLUTIONARY CONSTRAINTS, MAIN PROTEASES OF (+)SSRNA VIRUSES CONTINUE TO Evolve

Virus-encoded proteases are essential to the life cycle of numerous (+)ssRNA viruses. Newly synthesized viral polyproteins mature into individual, functional proteins via a series of cleavage events carried out by virus-encoded and host proteases. For Picornaviridae and Coronaviridae, the viral cysteine proteases 3C and 3C-Like (3CL) respectively, are responsible for the majority of polyprotein processing events (Figures 3A, B) (4, 16–19). Most picornaviruses have six or more 3C cleavage sites throughout the polyprotein (Figure 1A), and there is a preference to cleave between a glutamine (Q) in the P1 position and a small residue [e.g. glycine (G) or serine (S)] in the P1’ position (Figure 3A) (15, 16, 20). Likewise, coronaviruses (CoVs) have ten or more cleavage sites for the 3CL protease (also known as MPro or nsp5 in several CoVs including SARS-CoV-2) (Figure 3B) (17, 18, 21). Numerous other viral families, including members of Caliciviridae (e.g. norovirus) (22) and Potyviridae (e.g. tobacco etch virus) (23) encode a cysteine protease with a similar specificity for cleavage between a Q and a small residue, whereas members of Togaviridae (e.g. Chikungunya virus) use a cysteine protease with different cleavage specificity (24). Other viral families use a serine protease, including Flaviviridae, where the serine protease NS3 processes at least four polyprotein cleavage sites (Figure 3C) (25, 26). Here and in subsequent sections, we will predominantly discuss activities of the 3C, 3CL, and NS3 proteases of Picornaviridae, Coronaviridae, and Flaviviridae, respectively, due to their known roles in cleaving mammalian host factors. It is important to point out that because the polyprotein is sequentially processed, and not always to completion, protease activity may also be carried out when 3C, 3CL, or NS3 remains fused or associated with additional viral proteins (27, 28). This is especially true in the Flaviviridae, where the NS3 protease usually functions in association with NS2B (in the case of flaviviruses such as dengue and Zika viruses) or NS4A [in the case of hepatitis C virus (HCV)] (26, 29). However, for the sake of clarity, we will subsequently only refer to the protease domains of 3C, 3CL, or NS3. Moreover, many (+)ssRNA viruses encode additional proteases involved in both polyprotein processing and host antagonism, including the 2A protease in some picornaviruses and the papain-like protease (PLP) in coronaviruses (18, 30). Finally, viruses other than (+) ssRNA viruses can encode proteases that are important for polyprotein processing, most notably, the retroviroly-encoded aspartyl protease (31). While all of these additional proteases from (+)ssRNA viruses and retroviruses play important host antagonism roles, and likely shape host and viral evolution, they will not be extensively explored here.

The functions of the (+)ssRNA viral proteases described above are, by definition of being required for completion of the viral life cycle, well conserved. In addition to homology between the proteases themselves, the positions and sequences of the polyprotein cleavage motifs are often similar between members of the same viral family. Indeed, this conservation of polyprotein cleavage motifs has made it possible to compile sequences surrounding the cleavage site from genome sequences alone to generate a consensus motif for the viral protease that can be used to predict host and viral targets (15, 20, 32, 33) (Figure 3A). These consensus motifs are often generated using many diverse viruses, relying on the assumption that protease sequence
specifcity is well conserved among virus species. Interestingly, despite the evolutionary constraint to maintain cleavage across multiple sites in the polyprotein, virus-encoded proteases are substantially divergent across viruses (Figure 3). For instance, picornavirus 3C proteases can share less than 20% amino acid sequence identity, despite sharing an overall similar fold and many homologous cleavage sites (15). Similar evolutionary distances are observed with other families of proteases, including Coronaviridae 3CL and Flaviviridae NS3 (Figure 3).

Even with the divergence of protease sequences, protease sequence specificity is expected to be well conserved within closely related viruses given the essentiality of cleaving multiple site-specific polyprotein sites. Surprisingly, there is mounting evidence that this is not the case. For instance, among closely related serotypes of dengue virus (DENV), biochemical substrate profiling has revealed a subtle but clear shift in the NS3 protease cleavage sequence specificity profile (34). This type of in-depth comparative biochemical analysis of other (+) ssRNA proteases has not been conducted, but assays on model substrates have revealed differences in cleavage specificity even among 3C proteases within the Enterovirus genus of Picornaviridae (35). Some of the best evidence that protease sequence specificity is changing between related viruses has come from studies using chimeric viruses in which the protease of one virus species is inserted into the backbone of another virus. If such protease swaps result in insufficient or improper cleavage of the polyprotein and reduced viral replication, it would suggest divergence in protease sequence specificity between the parental viruses. For example, among enteroviruses (Figure 3A), replacing the poliovirus (PV) 3C protease with 3C proteases from human rhinovirus 14 or coxsackievirus B3 (CVB3) resulted in reduced, changed, or loss of cleavage products (36). Likewise, within the Flavivirus genus of Flaviviridae (Figure 3C), swapping the protease domain of DENV NS3 for the protease domain of yellow fever virus (YFV) ablates processing of polyproteins containing DENV cleavage sites (37). Additionally, West Nile virus (WNV) NS3 can cleave a polyprotein site in only one of two closely related DENV2 strains, where the only difference is in the residue in the P1’ position (38). While some of these differences may be attributed

FIGURE 3 | Main proteases in Picornaviridae, Coronaviridae, and Flaviviridae. (A) Phylogenetic tree of available RefSeq Picornaviridae 3C protease protein sequences (151 total, top). Names of viruses with human relevance or referenced throughout the text are listed next to their respective genus or singular node. The consensus enterovirus 3C cleavage motif (bottom) as was generated previously (15). The cleavage site is shown flanked by four amino acids upstream (labeled P4 through P1) and four amino acids downstream (labeled P1’ through P4’). (B) Schematic of the SARS-CoV-2 (family: Coronaviridae) nonstructural (ORF1ab) polyprotein, with the position of the 3CL protease and sites of 3CL-mediated cleavage shown. Phylogenetic tree of available RefSeq Coronaviridae 3CL protease protein sequences (64 total). Names of viruses with human relevance or referenced throughout the text are listed next to their respective genus. (C) Schematic of the dengue virus (DENV) (family: Flaviviridae) polyprotein, with the position of the NS3 protease and sites of NS3-mediated cleavage shown. Phylogenetic tree of available RefSeq Flaviviridae NS3 protease protein sequences (68 total). Names of viruses with human relevance or referenced throughout the text are listed next to their respective genus.
to the requirement for NS3 proteases to bind to lineage-specific activating cofactors to further augment cleavage specificity (39–47), it is also likely that these changes in cleavage specificity are dependent on non-conserved residues in the binding pocket of the NS3 protease (48). Similarly, within Coronaviridae, replication competent chimeric murine hepatitis virus (MHV) could not be recovered when the 3CL protease was replaced with one of many related alpha- or beta-coronavirus proteases including SARS-CoV, hCoV-229E and bat CoV-HKU4 (Figure 3B). Only when the MHV 3CL was replaced with the two most closely related beta-coronaviruses, hCoV-OC43 and hCoV-HKU1, could virus be recovered, but with a substantial fitness cost (49). Altogether, these biochemical and chimeric virus studies illustrate that (+)ssRNA viruses have undergone lineage-specific evolution in both their protease sequence specificity as well as their many polyprotein cleavage sites.

### (+)SSRNA VIRAL PROTEASES HAVE EVOLVED IN CONFLICT WITH THEIR HOSTS

The above-described changes in protease sequence specificity do not require invocation of adaptation. Indeed, evolutionary drift could result in changes to the viral protease and its cleavage sites, including those that result in loss of fitness for chimeric viruses. However, there is another selective pressure that likely shapes viral protease evolution: the advantage that viruses gain by cleaving host targets. Proteins in multiple cellular processes have been identified as targets of viral proteases, many of which are involved in the host antiviral immune response (4). Many of these host targets are divergent between species, potentially establishing molecular barriers to cross-species transmission. Although the ability to cleave the viral polyprotein is an invariant function of viral proteases, we posit that cleavage of specific host proteins may be selected for during viral evolution, especially during or following cross-species transmission. Indeed, pathogenicity of a mouse-adapted SARS coronavirus required two mutations in 3CL to facilitate rapid, robust virus replication (50). Although it has not been established whether these 3CL changes result in changes in host target cleavage, these data indicate that protease evolution may be required for successful adaptation to a novel host species.

Several excellent reviews have been written describing the diverse host targets that are cleaved by viral proteases (4, 16, 17, 19, 51). In many cases, the described host-virus interaction has focused on a single or a small number of related viral proteases and only a single host species, often humans. Thus, the importance of host and virus diversity in these interactions is often poorly understood. However, evidence is accumulating that viral proteases and their host targets are engaged in species-specific interactions. Below, we highlight such cases in which host and viral diversity alter the outcome of the interaction between host pathways and proteases of picornaviruses, flaviviruses and coronaviruses, illustrating this ongoing molecular arms race.

### Viral Proteases Target Essential Host Processes in a Virus-Specific Manner

Some of the best studied targets of viral proteases, especially from picornaviruses, are involved in well conserved processes such as translation initiation or translation control (52, 53). In many cases, the functional outcome is similar: viral proteases antagonize a host molecular function in a way that benefits the virus. However, the specific host protein or specific site within that host protein can be divergent between different viruses, highlighting differences in protease cleavage specificity between related viruses, as well as the convergence of viral protease cleavage on the same host pathways. Thus, even for host functions that are ‘well conserved’ targets of protease cleavage, there is surprising mechanistic diversity. Below we highlight two such examples in well described targets of picornavirus proteases, but likely many other similar examples exist.

Translation of picornavirus mRNAs occurs via an internal ribosome entry site (IRES) (54). This bypasses the need to engage with host cap-dependent translation machinery and offers the opportunity to induce a ‘host-shutoff’ of translation of host antiviral proteins while maintaining production of viral proteins. Many picornaviruses inhibit host translation in a protease-dependent manner via cleavage of subunits of the eIF4F cap-binding complex, which binds to host mRNA cap structures to establish the initiation complex, or poly-A binding protein (PABP), which binds the 3’ polyA tail of mRNAs and eIF4G to circularize mRNAs for optimal translation initiation (55) (Figure 4A). For instance, Foot-and-mouth disease virus (FMDV) 3C cleaves the eIF4G and eIF4A subunits of eIF4F (56). Interestingly, neither hepatitis A (HepA) virus nor encephalomyocarditis virus (EMCV) 3C target eIF4G for cleavage, but both target PABP (57, 58). Convergently, PV also targets PABP, but at a site that is ~100 residues away from the cleavage site of EMCV (59) and additionally uses its 2A protease to cleave eIF4G (60). Despite cleaving different host targets and/or host sites, these interactions all result in host translation shut-off. These data highlight functional conservation, rather than molecular conservation, of picornavirus 3C-mediated inhibition of host translation and suggests that even among related viruses, there are important differences in the viral specificity of host target cleavage.

A similar phenomenon is observed in another well-established target of picornavirus 3C proteases, the stress granule protein G3BP1 (Figure 4B). Numerous viruses manipulate stress granule formation for their benefit, as this is a major intersection point between translation control and cellular stress responses (53, 61). Among the mapped cleavage sites in G3BP1, PV 3C cleaves at Q326 (62) while FMDV 3C cleaves at E284 (63), but both of these cleavage events benefit the virus by manipulating stress granule formation. These findings further demonstrate the convergence of 3C cleavage onto the same host target, while highlighting how subtle differences in cleavage specificity can impact viral targeting of host factors. Of note, the P1 and P3’ positions of the PV cleavage site are altered in a way that would prevent cleavage of mouse G3BP1, which is otherwise >90% identical to the human protein (Figure 4B).
Whether host G3BP1 is cleaved by enteroviruses that infect rodents, and at what site, has yet to be determined.

**Proteins in the Innate Antiviral Immune Response Are Common Targets of Viral Proteases**

In addition to essential cellular processes, proteins in the innate antiviral immune response are common targets of diverse viral proteases. The host antiviral response is initiated when cells detect viral products (Figure 5A). Following entry into a host cell, viral nucleic acids can be detected by host pattern recognition receptors (PRRs) such as RIG-I, MDA5, and cGAS. While RIG-I and MDA5 directly detect viral ssRNA or dsRNA as a product of (+)ssRNA virus replication (67–71), the cytosolic DNA sensor cGAS can be indirectly activated via virus-induced mitochondrial damage and subsequent release of mitochondrial DNA that can occur during (+)ssRNA viral infection (72). After ligand binding, PRRs recruit a series of adaptor proteins, ultimately resulting in the production and secretion of type I and III interferons (IFN-I and IFN-III) (73). IFN-I and IFN-III are antiviral cytokines that signal in an autocrine or paracrine manner to induce expression of interferon-stimulated genes (ISGs), which act to directly and indirectly inhibit virus replication and establish an antiviral state in the host (74) (Figure 5A).

Induction of IFN and subsequent upregulation of ISGs is critical to the host antiviral defense. Therefore, proteins involved in these pathways are common targets of viral antagonism (73), including several that are cleaved by (+)ssRNA viral proteases (Table 2 and Figure 5A). For instance, NS3 from DENV and other flaviviruses can cleave and inactivate STING to prevent sensing of cytoplasmic mitochondrial DNA (65, 75), whereas PV and possibly other 3C proteases cleave RIG-I during infection (76, 96) (Figure 5A). Tellingly, many proteases convergently cleave the same host targets. For instance, CVB3 3C and HCV NS3 are both able to cleave MAVS (77, 78), a critical innate immune adaptor for both MDA5 and RIG-I (Figure 5A). 3CL proteases from Porcine Epidemic Diarrhea Virus (PEDV), porcine deltacoronavirus (PDCoV), and feline infectious peritonitis virus (FIPV), as well as 3C proteases from FMDV and HepA can also inhibit RIG-I/MDA5 pathways by cleaving nuclear transcription factor kB (NF-kB) essential modulator (NEMO), a bridging adaptor protein involved in activating both NF-kB and interferon-regulatory factor signaling pathways (89–93, 97) (Figure 5A). Finally, STAT2, one of the critical transcription factors that transmits the signaling of IFN to ISG production (Figure 5A), is cleaved by the 3CL from PDCoV (85), although whether other 3CLs cleave this protein is unknown. Altogether these data show that viral protease-mediated cleavage of innate immune signaling proteins is a common strategy across (+)ssRNA viruses to prevent the antiviral response and promote virus replication.

Many proteins in the innate antiviral immune response are rapidly evolving within and between host populations (98–100). One potential consequence of these host changes is that a cleavage site for a viral protease may be present in one host but not another. If there is strong selection for the virus to restore antagonism of that host function, there would be selection for viral proteases that would change the sequence specificity of host target cleavage to either restore cleavage of the original site, cleave another site on the host protein, or cleave another protein in the host pathway (Figure 5B). Such an evolutionary model can be used to understand the genetic bases for host- and viral-specificity of protease cleavage. For many of the known interactions between host immunity proteins and viral proteases, there is little information on how host and viral evolution shapes the outcome. However, analyses on two host
FIGURE 5 | Protease antagonism of IFN induction and signaling pathways. (A) Examples of viral proteases that antagonize the innate antiviral immune response, including antagonism of IFN-induction (left) or signaling downstream of IFN (right). (B) Model for how protease sequence specificity may be driven to evolve by conflicts with host factors. Following host evolution, or cross-species transmission, viral proteases may no longer be able to antagonize a given host factor. To re-establish host antagonism, the protease can evolve to cut a different sequence at same host site (left) or may evolve to cut a new site elsewhere in the host protein (right). (C, D) Evolution of MAVS (64) and STING (65) across primates and other mammals confers resistance or susceptibility to flavivirus protease cleavage. Red asterisks mark the inferred branch in which an amino acid change occurred that alters cleavage susceptibility. (E) Human STING cleavage by flavivirus NS3 proteases is virus species-specific. Data adapted from (66).

TABLE 2 | Select list of IFN pathway-related targets of (+)ssRNA virus proteases.

| Host target | Viral protease | References |
|-------------|----------------|------------|
| STING       | NS3 (NS2B3) (ZIKV, JEV, WNV, YFV, DENV) | (65, 66, 75) |
| RIG-I       | 3C (PV)        | (76)       |
| MAVS        | NS3 (NS3-4A) (HCV, GBV-B) | (64, 77–82) |
|             | 3C (CVB3, SVV) |            |
| Riplet      | NS3 (NS3-4A) (HCV) | (83)     |
| MDA5        | 3C (FMDV)      | (84)       |
| STAT2       | 3CL (PDCoV)    | (85)       |
| TRIF        | NS3 (NS3-4A) (HCV) | (78, 81, 86–88) |
|             | 3C (CVB3, SVV, EV68) |            |
|             | 3CD (HAV)      | (89–93)    |
| NEMO        | 3CL (PEDV, FPV, PDCoV) | (89)       |
|             | 3C (FMDV, HAV) | (90)       |
| IRF7        | 3C (EV68)      | (91)       |
| IRF9        | 3C (EV71)      | (92)       |

Viral abbreviations are as follows: ZIKV, Zika virus; JEV, Japanese encephalitis virus; WNV, West Nile virus; YFV, Yellow fever virus; DENV, Dengue virus; PV, Poliovirus; HCV, Hepatitis C virus; GBV-B, GB virus B or Pegivirus B; CVB3, Coxsackievirus B3; SVV, Seneca Valley virus; FMDV, Foot and mouth disease virus; PDCoV, Porcine deltacoronavirus; EV68, Enterovirus D68; HAV, Hepatitis A virus; PEDV, Porcine epidemic diarrhea virus; FIPV, Feline infectious peritonitis virus.
In these cases, host evolutionary signatures reveal adaptations that are presumed to evade cleavage. However, another possibility exists, in which the host protein can sense the presence of the viral protease in the cytoplasm through an evolved sequence that mimics the viral polyprotein cleavage site. Sensing of pathogen-encoded activities such as toxins and effector enzymes, known as effector-triggered immunity (ETI), is well-described in plants but is also emerging as an important immune mechanism in animals (103–106). Three such signaling pathways, described below, are known to detect the main protease activity of human viruses.

NLRP1 Mimics Diverse Picornaviral 3C Cleavage to Trigger Inflammation

One of the best described cases of mammalian ETI involves NLRP1 (NACHT, LRR, and PYD domains-containing protein 1; Figure 6A), a critical sensor for the innate immune complex known as the inflammasome. Mouse NLRP1B was identified in a genetic screen as a determinant of differential susceptibility between mouse strains to Lethal Toxin, a virulence factor responsible for the major pathologies seen during infection by the bacterial pathogen *Bacillus anthracis* (107). Further research identified that NLRP1B was a target of cleavage by the secreted bacterial protease component of Lethal Toxin, termed Lethal Factor (LF). Interestingly, mice with a cleavage-resistant variant of NLRP1B were protected from *B. anthracis* challenge, indicating that cleavage of NLRP1B was immunologically protective (108, 109). The mechanism by which this occurs, termed ‘functional degradation’ (110, 111), depends on the FIIND domain encoded within NLRP1B, which undergoes a constitutive self-cleavage event (known as ‘auto-processing’) such that the N-terminal domains and C-terminal CARD-containing fragment of NLRP1B exist as two distinct, noncovalently associated polypeptides (112, 113) (Figure 6B). Once LF cleaves upstream of the FIIND domain in NLRP1B, the released product has a new N-terminus that is recognized by the N-end rule cellular machinery and targets it for proteasome-mediated degradation. However, as a result of the break in the polyprotein backbone within the FIIND domain, proteasome-mediated degradation of NLRP1B ceases after degrading the N-terminal domains, leaving the bioactive C-terminal fragment intact and able to assemble into an active inflammasome (Figure 6B) (110, 111). The unusual domain architecture of NLRP1B thus facilitates the mounting of the inflammasome response upon proteolytic cleavage of the N-terminus.

LF cleaves within the rapidly evolving ‘tripwire’ region of mouse NLRP1B but fails to cleave or activate human NLRP1. Interestingly, human NLRP1 has an analogous rapidly-evolving ‘tripwire’ region (Figure 6A), and cleavage of human NLRP1 via an engineered tobacco etch virus (TEV) protease cleavage site can activate the inflammasome (114). These data suggested that human NLRP1 may also detect pathogen-encoded proteases and activate the inflammasome via a functional degradation mechanism. Indeed, we and others recently identified that human NLRP1 recognizes picornavirus 3C protease activity and serves as a tripwire for inflammatory cell death and downstream...
During enterovirus infection, 3C cleavage of NLRP1 results in assembly of the active inflammasome and subsequent pro-inflammatory cytokine release (15, 115), including in human primary airway epithelial cells (115). Interestingly, based on phylogenetic analyses, the 3C-protease site mimic in this specific region of NLRP1 only evolved in the primate lineage, and is only cleavable in some primates (15). Differences across simian primates and a SNP within the human population prevent cleavage and inflammasome activation (15). Although mice lack this human-aligned cleavage site, we discovered a similar phenomenon where picornavirus 3C proteases cleave NLRP1 at different sites to activate the inflammasome in a virus- and mouse-strain-specific manner (15).

In addition to host diversity, viral diversity also determines NLRP1 cleavage. While all enteroviruses cleave the same site within NLRP1 and activate the inflammasome, other picornaviruses cleave NLRP1 at different sites within the N-terminal domain or do not cleave NLRP1 (15). For instance, the 3C protease of EMCV does not cleave NLRP1, and resulting no activation of the NLRP1 inflammasome was observed upon EMCV infection (15). As numerous sites in the protease-sensing N-terminal region of NLRP1 are evolving under positive selection (114), other independently evolved tripwire sites within NLRP1 may sense divergent 3C or other viral proteases.

**Intracellular HIV-1 Protease Activity Triggers Inflammation via CARD8**

Another inflammasome mediator, CARD8, is known to share the unusual C-terminal domain structure critical for the sensing mechanism of NLRP1 – the FIIND domain followed by a CARD domain (Figure 6A) (116, 117). In addition to these domain similarities, CARD8 inflammasome assembly can also be activated by the same small molecules as NLRP1 (117). Such similarities initially suggested that CARD8 could also be activated using a functional degradation model to act as a tripwire sensor of pathogen-encoded activities (116). Indeed, the protease of human immunodeficiency virus 1 (HIV-1) can cleave and activate the CARD8 inflammasome in an activation mechanism that resembles NLRP1 (118). While HIV-1 protease is normally important for cleaving viral polyproteins in the maturing capsid, treatment with specific non-nucleoside reverse transcriptase inhibitors (NNRTIs) can result in protease activity in the cytoplasm (119). Under these NNRTI treatment conditions, HIV-1 proteases from four prevalent HIV-1 subtypes cleave CARD8 and activate the inflammasome, resulting in pro-inflammatory cytokine release and influencing clearance of latent HIV-1 in primary CD4+ T cells (118). While the extent to which host evolution or evolution of other viruses influences the activation of the CARD8 system remains unknown, these findings reveal a broader role for host encoded proteases.
tripwires for viral proteases that can activate a robust immune response using mimicry of viral protease cleavage sites.

**3C-Mediated Cleavage of a Regulator of NF-κB Triggers Apoptosis**

‘Tripwire’ mechanisms such as NLRP1 and CARD8 rely on a specific elegant, but rare, domain architecture that allows for coupling of a cleavage event to generation of a bioactive signaling molecule. An additional mechanism for sensing of viral proteases arises from the intricate ways that the innate immune response is negatively regulated. For instance, downstream of NLRP1 and CARD8, mature inflammatory cytokines are detected by the IL-1 receptor to activate the transcription factor NF-κB, which can amplify the inflammatory response (120). NF-κB is an essential transcription factor involved in many innate immune pathways and can mediate a variety of downstream responses depending on the input stimuli (121), including pro- or anti-apoptotic responses (122). Within the cytoplasm, the NF-κB heterodimer, composed of the Rel family proteins p65 and p50, remains bound and inactive by members of the inhibitor of NF-κB (IκB) family, including IκBα (Figure 5A). In response to cytokines such as IL-1β, IκB kinase (IKK) family proteins phosphorylate IκB proteins, releasing the active transcription factor to translocate into the nucleus (120). A previous study demonstrated that IκBα senses CVB3 3C protease activity (123). The 3C protease was shown to cleave IκBα, producing a fragment that stably complexes with p65 and translocates to the nucleus. This stable complex blocks NF-κB transcriptional activation, resulting in increased cell apoptosis and decreased viral replication (123). Thus, cleavage of IκBα may have evolved as another way to sense viral protease activity and induce cell death to prevent further virus propagation. Many viral proteases are known to cleave proteins in the NF-κB pathway (Table 2). Additional characterization of these virus-host interactions may reveal additional antiviral mechanisms associated with this critical immune pathway.

**Evolutionary Advantages of ETI**

In the continual evolutionary conflict between viruses and their hosts, cleavage mimicry encoded in NLRP1, CARD8 and NF-κB serve as examples of a successful strategy emerging in host organisms to exploit highly constrained pathogenic processes. Viruses are known to use molecular mimicry to antagonize or subvert the host immune response (124). In the cases of ETI described above, the host is turning the tables and using mimicry of viral protease cleavage sites to support the antiviral response. Rather than mimicry of entire proteins or protein domains, mimicry of these cleavage sites as ‘short linear motifs’ (SLIMs) require only a small number of amino acids to hijack the highly conserved protease activity (125, 126). In order to avoid these ‘tripwires’ and negative regulators of the immune response, these viruses must either evolve their respective main proteases along with all affiliate cleavage sites or antagonize the process some other way (Figure 6C). Supporting this idea, 3C proteases from some picornaviruses cleave NLRP1 but do not activate the NLRP1 inflammasome, suggesting that 3C proteases have evolved to evade detection by NLRP1 by antagonizing NLRP1 function elsewhere (15). We expect that this work may lead to the discovery that protease-driven ETI strategies may have evolved more broadly at other sites of host-pathogen conflicts.

**DISCUSSION**

The proteases of (+)ssRNA viruses have multiple roles in establishing and maintaining virus infection within a host. First and foremost, virally-encoded proteases cleave numerous sequence-specific sites within the viral polyprotein, which is essential for completion of the viral replication cycle. As a consequence of this essential activity, the ability of proteases to evolve novel sequence specificity is highly constrained. However, viral proteases also serve to manipulate numerous host processes in the infected cell through site-specific cleavage of host targets. In this context, changes in protease sequence specificity would allow the virus to cleave new host targets that might benefit the virus, or avoid cleaving host targets that are detrimental to the virus. It is at this intersection that viral proteases are engaged in evolutionary ‘arms races’ with the host, resulting in varied interactions across viral and host species and across evolutionary time. Several examples, including virus-specific cleavage of essential mRNA translation machinery and host-specific evasion of cleavage of innate antiviral immune components, highlight the consequences of these evolutionary conflicts. More recently, the discovery of host-encoded effector-triggered immunity (ETI) sensors such as NLRP1 and CARD8 suggest that host mimicry of viral protease cleavage sites is an efficient strategy to detect the cellular activity of viral proteases.

The extent to which viral protease evolution, and host target diversification, shape viral host range and pathogenesis remains unknown and is an exciting area of future research. The majority of characterized protease-host interactions have been described for a single virus against a single host, leaving open the opportunity for more detailed exploration of the evolutionary dynamics of these interactions. Indeed, examples such as cleavage of host proteins such as MAVS, STING, and NLRP1 highlight the insights that can be gained from additional analyses of host and viral diversity in these interactions. Likewise, future studies aiming to discover additional host targets of viral proteases, especially those that may be cleaved in a virus-specific manner, will advance our knowledge of the ways that protease-host interactions shape viral phenotypes. Finally, ETI sensors such as NLRP1 and CARD8 may represent just the start of host proteins that mimic viral protease cleavage sites to induce an immune response. Further studies aimed to identify ETI mechanisms against both viral and other pathogen-encoded proteases will likely continue to reveal novel mechanisms and evolutionary principles of the host innate immune response.

**AUTHOR CONTRIBUTIONS**

All authors discussed relevant literature. BT, EF, and MD wrote the first draft. All authors contributed to figure generation and
edits the manuscript. All authors contributed to the article and approved the submitted version.

**FUNDING**

This work was supported by grants from the National Institutes of Health (R35 GM133633), Pew Biomedical Scholars Program, Hellman Fellows Program, and Burroughs Wellcome Fund. Investigators in the Pathogenesis of Infectious Disease Program to MD and an NIH T32 grant (GM007240) to BT.

**ACKNOWLEDGMENTS**

We thank all members of the Daugherty laboratory for helpful discussions. We apologize to the colleagues whose work was not included in this review owing to space limitations.

**REFERENCES**

1. Walker PJ, Siddell SG, LeKowitz EJ, Mushagian AR, Adriaenssens EM, Dempsey DM, et al. Changes to Virus Taxonomy and the Statutes Ratified by the International Committee on Taxonomy of Viruses (2020). *Arch Virol* (2020) 165:2737–48. doi: 10.1007/s00705-020-04752-x
2. Anirudhan V, Lee H, Cheng H, Cooper L, Rong L. Targeting SARS-CoV-2 Viral Proteases as a Therapeutic Strategy to Treat COVID-19. *J Med Virol* (2021) 93:2722–34. doi: 10.1002/jmv.26814
3. Steuber H, Hilgenfeld R. Recent Advances in Targeting Viral Proteases for the Discovery of Novel Antivirals. *Curr Top Med Chem* (2010) 10:323–45. doi: 10.2174/1568026021007902470
4. Lei J, Hilgenfeld R. RNA-Virus Proteases Counteracting Host Innate Immunity. *FEBS Lett* (2017) 591:3190–210. doi: 10.1002/1873-3468.12827
5. Daugherty MD, Malik HS. Rules of Engagement: Molecular Insights From Host-Virus Arms Races. *Annu Rev Genet* (2012) 46:677–700. doi: 10.1146/annurev-genet-110711-155522
6. Duggal NK, Emerman M. Evolutionary Conflicts Between Viruses and Restriction Factors Shape Immunity. *Nat Rev Immunol* (2012) 12:687–95. doi: 10.1038/nri3295
7. Meyerson NR, Sawyer SL. Two-Stepping Through Time: Mammals and Viruses. *Trends Microbiol* (2011) 19:286–94. doi: 10.1016/j.tim.2011.03.006
8. Rothenburg S, Brennan G. Species-Specific Host-Virus Interactions: Implications for Viral Host Range and Virulence. *Trends Microbiol* (2020) 28:46–56. doi: 10.1016/j.tim.2019.08.007
9. Van Valen L. A New Evolutionary Law. *Evol Theory* (1973) 1:1–30. doi: 10.20787/9788026115504
10. Sironi M, Cagliani R, Forni D, Clerici M. Evolutionary Insights Into Host-Pathogen Interactions From Mammalian Sequence Data. *Nat Rev Genet* (2015) 16:224–36. doi: 10.1038/nrg38905
11. Diehl WE, Lin AE, Grubaugh ND, Carvalho LM, Kim K, Kyew PY, et al. Ebola Virus Glycoprotein With Increased Infectivity Dominated the 2013-2016 Epidemic. *Cell* (2016) 167:1088–98.e6. doi: 10.1016/j.cell.2016.10.014
12. Martin DP, Weaver S, Tegally H, San JE, Shank SD, Wilkinson E, et al. The Emergence and Ongoing Convergent Evolution of the SARS-CoV-2 N501Y Lineages. *Cell* (2021) 184:5189–200.e7. doi: 10.1016/j.cell.2021.09.003
13. Sawyer SL, Eldnc NC. A Cross-Species View on Viruses. *Curr Opin Virol* (2012) 2:561–8. doi: 10.1016/j.coviro.2012.07.003
14. Wain LV, Bailes E, Bibollet-Ruche F, Decker JM, Kelle BF, Van Heussem WF, et al. Adaptation of HIV-1 to Its Human Host. *Mol Biol Evol* (2007) 24:1853–60. doi: 10.1093/molbev/msm110
15. Tsu BV, Beierschmidt C, Ryan AP, Agarwal R, Mitchell PS, Daugherty MD. Diverse Viral Proteases Activate the NLRP1 Inflammasome. *Elife* (2011) 10.7208/9780226115563_0043
16. Laitinen OH, Svedin E, Kapell S, Nurminen A, Hytonen VP, Flodstrom-Dempsey DM, et al. Changes to Virus Taxonomy and the Statutes Rati...
36. Dewalt PG, Lawson MA, Colombo RJ, Semler BL. Chimeric Picornavirus Polypeptides Demonstrate a Common 3C Proteinase Substrate Specificity. J Virol (1998) 63:3444–52. doi: 10.1128/jvi.63.8.3444-3452.1998

37. Preusschat F, Lenchus EM, Strauss JH. Flavivirus Enzyme-Substrate Interactions Studied With Chimeric Proteinases: Identification of an Intragenic Locus Important for Substrate Recognition. J Virol (1991) 65:4749–58. doi: 10.1128/jvi.65.9.4749-4758.1991

38. VanBlargan LA, Davis KA, Dowd KA, Akey DL, Smith JL, Pierson TC. Context-Dependent Cleavage of the Capsid Protein by the West Nile Virus Protease Modulates the Efficiency of Virus Assembly. J Virol (2015) 89:6832–42. doi: 10.1128/JVI.01253-15

39. Jan LR, Yang CS, Trent DW, Falgout B, Lai CJ. Processing of Japanese Encephalitis Virus Non-Structural Proteins: NS2B-NS3 Complex and Homologous Proteases. J Gen Virol (1995) 76(Pt 3):573–80. doi: 10.1099/0022-1317-63-3-573

40. Falgout B, Pethel M, Zhang YM, Lai CJ. Both Nonstructural Proteins NS2B and NS3 Are Required for the Proteolytic Processing of Dengue Virus Nonstructural Proteins. J Virol (1991) 65:2467–75. doi: 10.1128/jvi.65.5.2467-2475.1991

41. Falgout B, Miller RH, Lai CJ. Deletion Analysis of Dengue Virus Type 4 Nonstructural Protein NS2B-Identification of a Domain Required for NS2B-NS3 Protease Activity. J Virol (1993) 67:2034–42. doi: 10.1128/jvi.67.4.2034-2042.1993

42. Calboy A, Falgout B, Lai CJ. Cleavage of the Dengue Virus Polyprotein at the NS3/NS4A and NS4B/NS5 Junctions Is Mediated by Viral Protease NS2B-NS3, Whereas NS4A/NS4B may be Processed by a Cellular Protease. J Virol (1992) 66:1535–42. doi: 10.1128/jvi.66.3.1535-1542.1992

43. Chambers TJ, Grakoui A, Rice CM. Processing of the Yellow Fever Virus Nonstructural Polyprotein: A Catalytically Active NS3 Proteinase Domain. J Virol (1990) 65:4024–33. doi: 10.1128/jvi.65.7.4024-4033.1990

44. Sbardellati A, Scarselli E, Amati V, Falcinelli S, Kekule AS, Traboni C. Processing of GB Virus B Nonstructural Protein 3. J Virol (1991) 65:2491–301. doi: 10.1128/JVI.65.4.2491-4301.2000

45. Xu J, Mendez E, Caron PR, Lin C, Murcko MA, Collett MS, et al. Bovine Viral Diarrhea Virus NS3 Serine Proteinase: Polyprotein Cleavage Sites, Cofactor Requirements, and Molecular Model of an Enzyme Essential for Pestivirus Replication. J Virol (1997) 71:5312–22. doi: 10.1128/jvi.71.5.5312-5322.1997

46. Wiskerchen M, Collett MS, Pestivirus Gene Expression: Protein P80 of Bovine Viral Diarrhea Virus Is a Proteinase Involved in Polyprotein Processing. Virology (1991) 184:341–50. doi: 10.1006/jviro.1991.9080-B

47. Junaid M, Chalayut C, Sehgelmeble Torrejon A, Angsuthanasombat C, Miyake K, Forster B, et al. Arterivirus and Nairovirus Ovarian Tumor Domain-Containing Deubiquitinases Target Activated RIG-I to Control Innate Immune Signaling. J Virol (2012) 86:7737–45. doi: 10.1128/JVI.00677-11

48. Luo R, Xiao S, Jiang Y, Jin H, Wang D, Liu M, et al. Porcine Reproductive and Respiratory Syndrome Virus (PRRSV) Suppresses Interferon-Beta Production by Interfering With the RIG-I Signaling Pathway. Mol Immunol (2008) 45:2839–46. doi: 10.1016/j.molimm.2008.01.028

49. Deddouche S, Goubau D, Rehwinkel J, Chakravarty P, Begum S, Maillard E, et al. Distinct RIG-I and MDA5 Signaling by RNA Viruses in Innate Immunity. J Virol (2008) 82:3335–45. doi: 10.1128/JVI.01080-07

50. Kato H, Takeuchi O, Sato S, Yoneyama M, Yamamoto M, Matsui K, et al. Differential Roles of MDA5 and RIG-I Helicases in the Recognition of RNA Viruses. Nature (2006) 441:101–5. doi: 10.1038/nature04734

51. West AP, Khoury-Hanold W, Staron M, Tal MC, Pineda CM, Lang SM, et al. Mitochondrial DNA Stress Primers the Antiviral Innate Immune Response. Nature (2015) 520:553–7. doi: 10.1038/nature14136

52. Hoffmann HH, Schneider WM, Rice CM. Interferonα and Interferonγ: An Evolutionary Arms Race of Molecular Interactions. Trends Immunol (2015) 36:124–38. doi: 10.1016/j.it.2015.01.004
115. Robinson KS, Teo DET, Tan KS, Toh GA, Ong HH, Lim CK, et al. Enteroviral 3C Protease Activates the Human NLRP1 Inflammasome in Airway Epithelia. Science (2020) 370:eay2002. doi: 10.1126/science.aay2002

116. Mitchell PS, Sandstrom A, Vance RE. The NLRP1 Inflammasome: New Mechanistic Insights and Unresolved Mysteries. Curr Opin Immunol (2019) 60:37–45. doi: 10.1016/j.coi.2019.04.015

117. Taabazuing CY, Griswold AR, Bachovchin DA. The NLRP1 and CARD8 Inflammasomes. Immunol Rev (2020) 297:13–25. doi: 10.1111/imr.12884

118. Wang Q, Gao H, Clark KM, Mugisha CS, Davis K, Tang JP, et al. CARD8 Is an Inflammasome Sensor for HIV-1 Protease Activity. Science (2021) 371: eabe1707. doi: 10.1126/science.abe1707

119. Figueiredo A, Moore KL, Mak J, Sluis-Cremer N, de Bethune MP, Tachedjian G. Potent Nonnucleoside Reverse Transcriptase Inhibitors Target HIV-1 Gag-Pol. PLoS Pathog (2006) 2:e119. doi: 10.1371/journal.ppat.0020119

120. Li X, Massa PE, Ihanida A, Peet GW, Aro P, Savitt A, et al. IKKalpha, IKKbeta, and NEMO/IKKgamma Are Each Required for the NF-Kappa B-Mediated Inflammatory Response Program. J Biol Chem (2002) 277:45129–40. doi: 10.1074/jbc.M205165200

121. Oeckinghaus A, Hayden MS, Ghosh S. Crosstalk in NF-kappaB Signaling Pathways. Nat Immunol (2011) 12:695–708. doi: 10.1038/ni.2065

122. Radhakrishnan SK, Kamalakaran S. Pro-Apoptotic Role of NF-Kappab: Implications for Cancer Therapy. Biochim Biophys Acta (2006) 1766:53–62. doi: 10.1016/j.bjba.2006.02.001

123. Zaragota C, Saura M, Padalko EY, Lopez-Rivera E, Lizarbe TR, Lamas S, et al. Viral Protease Cleavage of Inhibitor of Kappabalpha Triggers Host Cell Apoptosis. Proc Natl Acad Sci USA (2006) 103:19051–6. doi: 10.1073/pnas.0606019103

124. Elde NC, Malik HS. The Evolutionary Conundrum of Pathogen Mimicry. Nat Rev Microbiol (2009) 7:787–97. doi: 10.1038/nrmicro2222

125. Chemes LB, de Prat-Gay G, Sanchez IE. Convergent Evolution and Mimicry of Protein Linear Motifs in Host-Pathogen Interactions. Curr Opin Struct Biol (2015) 32:91–101. doi: 10.1016/j.csi.2015.03.004

126. Hagai T, Azia A, Babu MM, Andino R. Use of Host-Like Peptide Motifs in Viral Proteins Is a Prevalent Strategy in Host-Virus Interactions. Cell Rep (2014) 7:1729–39. doi: 10.1016/j.celrep.2014.04.052

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Tsu, Fay, Corley, Hosuru, Dominguez and Daugherty. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.