



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

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

Chapter 1

Fundamentals of Viruses and Their Proteases

Anjana Sharma*, Satya P. Gupta**

*Meerut Institute of Engineering and Technology, Meerut, Uttar Pradesh, India; **National Institute of Technical Teachers' Training and Research, Bhopal, Madhya Pradesh, India

1 INTRODUCTION

Viruses are small obligate intracellular parasites that may contain either an RNA or DNA genome encapsulated by a protective virus-coded protein coat. Both RNA and DNA viruses infect host organisms by entering and replicating in host cells. A virus makes multiple copies of itself in one cell, releases these copies to infect new host cells, and makes more copies. This process ultimately results in disease progression. Virus particles containing the viral genome are packaged in a protein coat called the capsid. This protein coat is surrounded by a lipid bilayer consisting of viral proteins that assist the virus in binding to the host cell membranes. The viral envelope plays a vital role in viral infection, including virus attachment to specific receptors cells. Then the penetration of enveloped viruses through fusion of the viral envelope with the host cell membrane takes place, which may or may not involve receptor-mediated endocytosis, while nonenveloped viruses penetrate through receptor-mediated endocytosis, releasing viral content into the cells and packaging newly formed viral particles. This process involves bringing together newly formed genomic nucleic acid and structural proteins to form the nucleocapsid of the virus.

Viral proteases reveal many new strategies for proteolysis, in addition to cleaving a peptide bond. Proteases have been identified in a wide range of viruses, without any correlation to capsid complexity, lipid envelope, or the nature of their genomes (Kräusslich and Wimmer, 1998). They are found to be present in both enveloped and nonenveloped viruses. The proteases present in various viruses specifically belong to the family of serine proteases as in hepatitis C virus (HCV), *Flavivirus*, and herpesviruses; the family of cysteine proteases as in adenoviruses (AdVs); or the family of aspartyl protease as in human immunodeficiency virus of type 1 (HIV1) (Marcin and Marcin, 2013). The present chapter presents introductory remarks to all kinds of viral proteases that may act simply as an appetizer to the readers, discussing in detail about the different

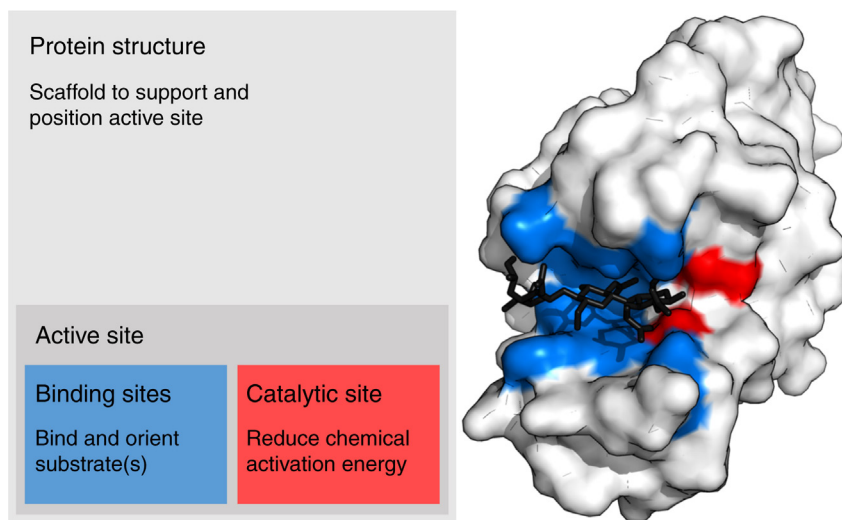


FIGURE 1.1 A representation of a protein structure showing its active site where the binding site is shown in blue, catalytic site in red, and the substrate (peptidoglycan) in black. (Image by Shafee, T. Available from Wikimedia Commons: <https://commons.wikimedia.org/w/index.php?curid=45801894>)

viruses, their structures, functions, and inhibition of their proteases. This chapter first presents the biochemistry of a protease and then introduces briefly about different viruses and their proteases.

2 BIOCHEMISTRY OF PROTEASES

A protease, also called a peptidase or proteinase, is the enzyme that performs proteolysis, that is, protein catabolism by hydrolysis of peptide bonds. Different classes of protease can perform the same reaction by completely different catalytic mechanisms. For catalytic action, proteases or enzymes possess an active site that consists of a binding site and a catalytic site as shown in Fig. 1.1, both constituting residues. Residues in the binding site form temporary bonds with the substrate and those in the catalytic site catalyze the reaction. The active site in the enzyme is usually a groove or pocket located in a deep tunnel within the enzyme. The residues of the catalytic site are typically very close to the binding site, and some residues can have dual roles, in both binding and catalysis. Once the substrate is bound and oriented in the active site, catalysis can begin.

Usually proteases possess a catalytic triad. As shown in Fig. 1.2, a catalytic triad refers to a group of three amino acid residues that function together at the center of the active site. An acid–base–nucleophile triad is a common motif for generating a nucleophilic residue (usually serine or cysteine amino acid, but occasionally threonine) for covalent catalysis. In three-dimensional structures of the enzyme, the residues of a catalytic triad can be far from each other along the

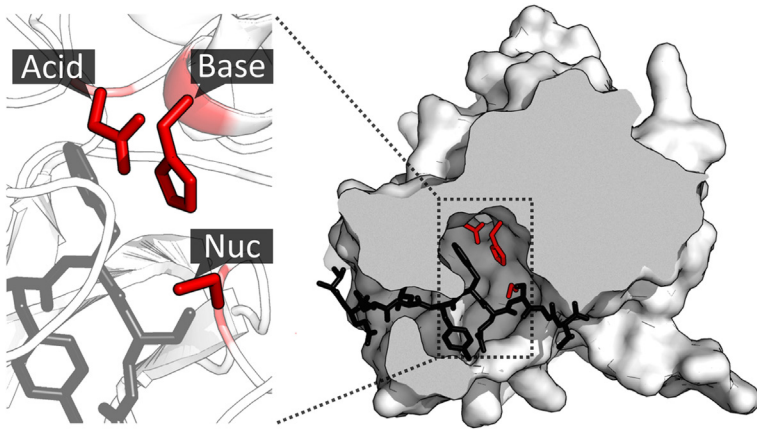


FIGURE 1.2 Example of a protease [tobacco etch virus (TEV) protease, TEV] with a catalytic triad consisting of an aspartate (acid), histidine (base), and serine (nucleophile). The substrate (black) is bound by the binding site to orient it next to the triad. (Image by Shafee, T. Available from Wikimedia Commons: <https://commons.wikimedia.org/w/index.php?curid=28958186>)

amino acid sequence in the primary structure. Some examples of catalytic triads are Ser-His-Asp, Cys-His-Asp, Ser-His-His, Ser-Glu-Asp, and Ser-*cis*Ser-Lys. Enzymes that contain a catalytic triad use it for one of two reaction types: either to split a substrate (hydrolases) or to transfer one portion of a substrate to a second substrate (transferases). Catalytic triads perform covalent catalysis using a residue as a nucleophile.

Proteases can be found in all members of the Animalia, Plantae, Fungi, Bacteria, Archaea, and Virus kingdoms. All proteases can be grouped on the basis of their catalytic residue, and thus they are classified into seven groups (Oda, 2012), namely, serine proteases (serine alcohol), cysteine proteases (cysteine thiol), threonine proteases (threonine secondary alcohol), aspartic proteases (aspartic carboxylic acid), glutamic proteases (glutamic carboxylic acid), metalloproteases (usually zinc), and asparagine peptide lyases (asparagine), where terms within parentheses refer to their catalytic residue. In the enzymes, the residues of the catalytic site are typically very close to the binding site, and some residues can have dual roles in both binding and catalysis.

Proteases are usually involved in breaking long protein chains into shorter fragments by splitting the peptide bonds that link amino acid residues. Some of them that detach the terminal amino acids from the protein chain, such as aminopeptidases and carboxypeptidase A, are called exopeptidases, and some that attack internal peptide bonds of a protein, such as trypsin, chymotrypsin, pepsin, papain, and elastase, are called endopeptidases. The catalytic mechanism of these proteases involves either the activation of a water molecule, which acts as a nucleophile to hydrolyze the peptide bond (Fig. 1.3, one-step catalysis), or the use of a nucleophilic residue in a catalytic triad to perform

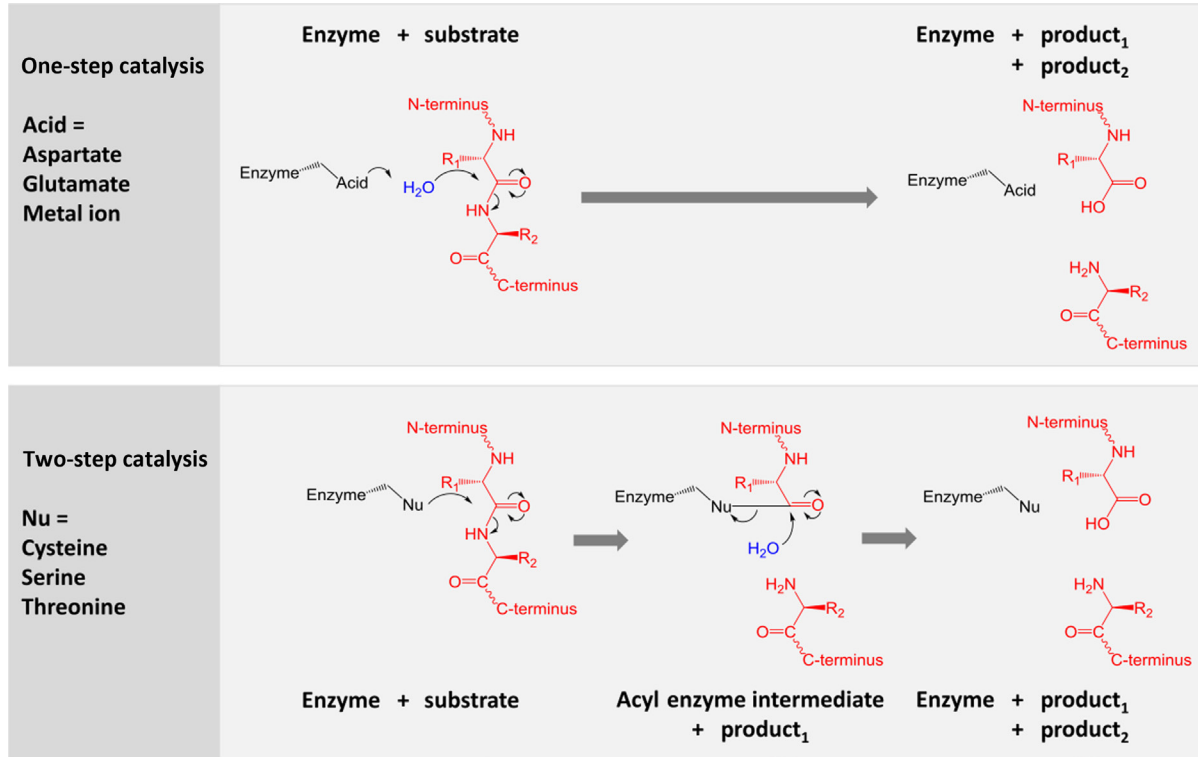


FIGURE 1.3 The two catalytic mechanisms adopted by proteases: one-step catalysis, showing the activation of a water molecule that performs a nucleophilic attack on the peptide bond to hydrolyze it, and a two-step catalysis, showing the use of a nucleophilic residue in a catalytic triad to perform a nucleophilic attack that covalently links the protease to the substrate protein, releasing the first half of the product. This covalent acyl enzyme intermediate is then hydrolyzed by activated water to complete the catalysis by releasing the second half of the product and regenerating the free enzyme. Enzyme is shown in black, substrate protein in red, and water in blue. (Image by Shafee, T. Available from Wikimedia Commons: <https://commons.wikimedia.org/w/index.php?curid=42551231>)

a nucleophilic attack that covalently links the protease to the substrate protein, releasing the first half of the product. This covalent acyl enzyme intermediate is then hydrolyzed by activated water to complete the catalysis through the release of the second half of the product and regeneration of the free enzyme (Fig. 1.3, two-step catalysis). Proteases that adopt the first mechanism are aspartic, glutamic, and metalloproteases and those that adopt second mechanism are serine, threonine, and cysteine proteases.

Regarding the specificity of proteases, it is said that proteolysis is highly promiscuous, such that a wide range of protein substrates are hydrolyzed. However, some proteases are highly specific and cleave only the substrate with a certain sequence, for example, tobacco etch virus (TEV) protease, which is specific for the sequence: ENLYFQ/S (where / = cleavage site).

3 VIRUSES

Viruses can be categorized into two major categories: enveloped and nonenveloped. Further classification is based upon the method of viral mRNA synthesis, that is, the Baltimore classification developed by Baltimore (1971), depending on their type of genome [DNA, RNA, single stranded (ss), or double stranded (ds), etc.] and their method of replication.

3.1 Enveloped/Nonenveloped Viruses

Enveloped viruses have a coat of lipids and proteins over which there are projection of spikes. They are able to induce antibody- and cell-mediated immunity. They are less virulent and are often released by budding and rarely cause cell lysis. They are highly sensitive to heat and light, while infectivity gets lost upon drying. Generally, they cannot survive inside the gastrointestinal tract. Their transmission is through blood, secretions, or organ transplantation. On the other hand, nonenveloped viruses have only a capsid protein coat and are more virulent. They have more lytic actions as compared to enveloped viruses, survive longer in host cells, and are more resistant to harsh environmental conditions, and thus survive in the gastrointestinal tract for a long time. Their mode of transmission is via fomites, dust, or intake of contaminated fecal or oral matter. They induce antibodies and can retain their infectivity even after drying. Some of the examples of nonenveloped viruses are *Rotavirus*, hepatitis A virus (HAV), adenovirus (AdV), etc., and that of enveloped viruses are *Ebolavirus*, HIV, hepatitis B virus, *Togavirus*, *Flavivirus*, influenza virus, etc.

3.2 DNA and RNA Viruses

A DNA virus is a virus that has DNA as its genetic material and replicates using a DNA-dependent DNA polymerase. They can be either dsDNA or ssDNA viruses. ssDNA is usually expanded to double stranded in infected cells. These two classes of DNA viruses are also called as Group I and Group II viruses,

respectively. Group I (dsDNA) viruses are highly dependent on the cell cycle of host cell. Before they are able to replicate, they must enter the host cell, where they forcefully induce the cell to undergo cell division. They require host cell polymerases to replicate the viral genome. Examples of Group I viruses are Herpesviridae, Adenoviridae, and Papoviridae. The examples of Group II (ss-DNA) viruses are Anelloviridae, Circoviridae, and Parvoviridae, which infect vertebrates; Geminiviridae and Nanoviridae, which infect plants; and the Microviridae, which infects prokaryotes. Most of Group II viruses contain circular genomes, except paroviruses.

RNA viruses are the viruses that have RNA as their genetic material. This nucleic acid is usually ssRNA, but may also be dsRNA. The example of RNA viruses are *Influenzavirus*, HCV, severe acute respiratory syndrome (SARS) *Coronavirus*, and polio and measles viruses. RNA viruses also are dsRNA or ssRNA viruses, which are put in Group III and Group IV or Group V, respectively, according to the Baltimore classification system. Viruses with RNA as their genetic material, but that include DNA intermediates in their replication cycle, are called retroviruses and comprise Group VI of the Baltimore classification. Notable human retroviruses include HIV1 and HIV2 that cause AIDS. Unlike DNA viruses, dsRNA viruses do not require host replication polymerases, rather they replicate in the core “capsid” that is in the cytoplasm. The ssRNA viruses are grouped into positive sense or negative sense according to sense of polarity of RNA. The replication of these viruses occurs in the cytoplasm and is not dependent on the cell cycle, such as that of DNA viruses. Well-known examples of positive-sense ssRNA viruses include picornaviruses [human rhinoviruses and foot-and-mouth disease virus (FMDV)], togaviruses (alphaviruses), and flaviviruses [HCV, West Nile virus (WNV), and yellow fever virus (YFV)]. The examples of negative-sense ssRNA viruses include viruses, such as measles virus, mumps virus, rabies virus, etc.

4 VIRUSES CONTAINING PROTEASES

The most important viruses that contain proteases and whose proteases have been exploited to design drugs of high therapeutic values have been mainly alphaviruses, HCV, HIV, picornaviruses, herpesviruses, AdVs, and flaviviruses. We are presenting here a brief discussion of these viruses and their proteases.

4.1 Adenoviruses

AdVs are common pathogens of humans, other mammals, and birds. They were first isolated in 1950s while studying the growth of poliovirus in the adenoidal tissue, hence their name (Rowe et al., 1953). It belongs to the family Adenoviridae, which includes viruses affecting other mammals and birds, and is divided into five genera: *Mastadenovirus*, *Aviadenovirus*, *Atadenovirus*, *Siadenovirus*, and *Ichtoadenovirus* (Martin et al., 2007).

Besides adenoids, they can also infect other sites, such as the respiratory tract, eye, as well as gut, causing acute respiratory disease, pneumonia, gastroenteritis, and other disorders in humans. AdVs are icosahedral, nonenveloped viruses with a diameter of about 1000 Å, and a viral genome containing about 36,000 base pairs (Horwitz, 2001). The AdV genome is linear, nonsegmented dsDNA, which is approximately 30–38 kbp that allows the virus to carry 30–40 genes theoretically. In spite of its larger structure, in the Baltimore classification it is still a simple virus that largely relies on the host cell for survival and replication. The viral genome has an interesting feature that it has a terminal 55-kDa protein associated with each of the 5' ends of the linear dsDNA, which act as primers in DNA replication and ensure that the ends of virus linear genome are adequately replicated. They are highly diversified. In humans there are 51 immunologically distinct human AdV serotypes in 6 species (A–F). Different serotypes are associated with various conditions, such as:

- Respiratory disease: mainly species human adenovirus (HAdV) B and C
- Conjunctivitis (HAdV B and D)
- Gastroenteritis (HAdV F serotypes 40 and 41)

Segerman et al. (2003) subdivided the species B into B1 and B2. There could be some correlation between the species and their tissue tropism and clinical properties. Thus the species B1, C, and E are responsible for causing respiratory disease, whereas species B, D, and E can induce ocular disease. Russell (2005) described that species F is the main cause for gastroenteritis, and B2 viruses infect the kidneys and urinary tract. Entry of AdVs into the host cell involves two types of interactions. One of the interactions for HAdV serotypes is via CD46 type of receptor, while another one is by coxsackievirus adenovirus receptor (CAR) serotypes. It was also found that the major histocompatibility complex (MHC) molecules and sialic acid residues also function in the same way. This is followed by secondary interactions, where the penton base protein interacts with an integrin molecule. This is the coreceptor that stimulates the entry of AdV-stimulated cell signaling, that is, entry of virions into the host cell (Wu and Nemerow, 2004).

4.1.1 Adenovirus Protease

AdV protease is a cysteine protease. Its structure consists of a central mixed five-stranded β -sheet that is surrounded by helices on both sides. Despite of a unique fold, the arrangement of the catalytic residues of AdV protease is same as that of papain, including the catalytic triad (Cys122-His54-Glu71), as well as the oxyanion hole (Gln115) of the protease. The overall organization of structure is almost similar to that of papain and Picornavirus leader protease, while some positions of structural features are different in the primary sequences of the proteases. The structural conservation of active site suggests that AdV protease is likely to have the same catalytic mechanism as papain (Ding et al., 1996).

4.2 Alphaviruses

Alphaviruses are mosquito-borne small, enveloped viruses belonging to the family of *Togaviridae*. [Strauss and Strauss \(1994\)](#) classified the virus depending upon the geographical origin: as New World alphaviruses or Old World alphaviruses. Old World alphaviruses, such as Chikungunya virus (CHIKV), O’Nyong–Nyong virus (ONNV), Ross River virus (RRV), Semliki Forest virus (SFV), and Sindbis virus (SINV), cause a fever, rash, and arthralgia syndrome, while New World alphaviruses, such as Eastern equine encephalitis virus (EEEV), Venezuelan equine encephalitis virus (VEEV), and Western equine encephalitis virus (WEEV), typically cause encephalitis in humans and other mammals ([Ryman and Klimstra, 2008](#)).

Alphaviruses are icosahedral in shape, small, enveloped, and about 70 nm in diameter ([Fuller, 1987](#); [Mancini et al., 2000](#); [Morgan et al., 1961](#)). It contains positive-sense ssRNA genome approximately 14 kb in length ([Simmons and Strauss, 1972](#); [Strauss et al., 1984](#)). Its subgenomic 26S RNA encodes the structural proteins, while nonstructural proteins are translated from genomic RNA ([Cancedda and Shatkin, 1979](#); [Simmons and Strauss, 1972](#); [Strauss et al., 1984](#)). The two polyproteins are cleaved posttranslationally by viral cysteine and host proteases. The structural proteins (C, E3, E2, 6K, and E1) and their cleavage intermediates are involved in budding and viral encapsidation, while RNA replication takes place by nonstructural proteins (nsP1–4) and their cleaved intermediates also ([Hardy and Strauss, 1989](#); [Melançon and Garoff, 1987](#); [Strauss et al., 1984](#)). After entry, the *Alphavirus* particles undergo disassembly and release genomic RNA into the cytoplasm of infected cells. From the viral genome, translation of nonstructural (P1234) and structural polyproteins takes place from open reading frames ([Glanville et al., 1976](#)). There is a *cis*-cleavage between nsP3 and nsP4 to yield P123 and nsP4, which forms an unstable replication complex (RC) and synthesizes negative-strand RNA. The nsP1 and P23 are *trans*-cleaved products of polyprotein, so nsP1, P23, and nsP4 form an RC within virus-induced cytopathic vacuoles that are active in negative-strand synthesis, as well as genomic RNA synthesis, but not in subgenomic RNA synthesis ([Dé et al., 1996](#); [Froshauer et al., 1988](#); [Kujala et al., 2001](#); [LaStarza et al., 1994](#); [Salonen et al., 2003](#)). The negative-strand synthesis is inactivated after complete cleavage to nsP1, nsP2, nsP3, and nsP4, and the synthesis of positive-strand genomic RNA is switched on by the stable RC. The *Alphavirus* nsP2 protein is a multifunctional enzyme with the N-terminus, having RNA helicase, nucleoside triphosphatase (NTPase), and RNA-dependent 5′ triphosphatase (RTPase) activities ([de Cedrón et al., 1999](#); [Karpe et al., 2011](#); [Pastorino et al., 2008](#)), while its C-terminus contains the protease domain ([Merits et al., 2001](#); [Strauss et al., 1992](#)). The nsP2 is considered to be an essential protein for viral replication and propagation due to its proteolytic activities ([Leung et al., 2011](#)). The antiviral gene expression is also inhibited after its translocation into the nucleus. The proteolytic activity of nsP2 has been

characterized in alphaviruses to be similar to that of papain-like cysteine protease, with a cysteine–histidine catalytic dyad in the active site.

4.3 Flaviviruses

Flaviviruses, belonging to the family of Flaviviridae, are a group of nearly 70 enveloped RNA viruses, which cause life-threatening diseases in both humans and animals. Mostly they are arthropodborne viruses and are transmitted to vertebrates via mosquito or tick bites (Gubler et al., 2007). The various members of this genus are Dengue virus (DENV), YFV, WNV, Japanese encephalitis virus (JEV), and St. Louis encephalitis virus (SLEV), all highly pathogenic to humans. According to WHO, DENV (serotypes DENV1–4) causes Dengue fever, which may progress to life-threatening shock syndrome and Dengue hemorrhagic fever. More than 200,000 cases suffering from YFV and 50,000 cases suffering from Japanese encephalitis virus are reported worldwide every year. Flaviviruses are enveloped viruses, having positive-sense ssRNA genome. They are introduced into the host cell by mosquito or tick during its blood meal. They enter into target cell via receptor-mediated endocytosis. The released RNA encodes a polyprotein precursor.

The *Flavivirus* genome is approximately 11 kb with a 5′ cap structure, but devoid of 3′ polyadenylation tail (Lindenbach et al., 2007). Both 5′ and 3′ ends play important roles in virus replication, viral protein translation, and virion assembly (Alvarez et al., 2006; Filomatori et al., 2006; Khromykh et al., 2003; Markoff, 2003; Polacek et al., 2009; Villordo and Gamarnik, 2008; Yu et al., 2008). The released RNA encodes a polyprotein precursor of approximately 3400 amino acids. This polyprotein is processed by host cell signalases and the viral NS2B–NS3 protease (NS2B–NS3pro) into three structural proteins, capsid (C), envelope (E), and membrane protein (M), and seven nonstructural proteins, NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5. The C-terminal region of NS3 contains an NTPase, a 5′ terminal RTPase, and an RNA helicase (Xu et al., 2005). NS5 possesses an RNA methyltransferase (Dong et al., 2012, 2014; Egloff et al., 2002; Ray et al., 2006) and RNA-dependent RNA polymerase activities (Malet et al., 2008; Yap et al., 2007). Together with the viral RNA, viral cofactors, and host cell cofactors, NS3 and NS5 form the virus RC assemble on the intracellular membrane to amplify the viral genome (Murray et al., 2008; Paul and Bartenschlager, 2013). Therefore, in principle, functional inhibition of the viral NS proteins and/or disruption of the RC underlie target-based anti-*Flavivirus* drug development (Bollati et al., 2010; Lim et al., 2013; Noble and Shi, 2012; Sampath and Padmanabhan, 2009). *Flavivirus* NS3 (69 kDa) is the second largest viral protein after NS5 in the *Flavivirus* genome, and plays several essential roles in the viral life cycle. NS3 has an N-terminal protease chymotrypsin-like domain that cleaves the viral polyprotein precursor to release individual NS proteins, and a C-terminal NTPase-dependent

RNA helicase [with an helicase superfamily 2 (SF2)-like fold] involved in genome replication and viral RNA synthesis (Lescar et al., 2008; Luo et al., 2010).

4.3.1 NS2B–NS3 Protease

The N-terminal domain of NS3 is a chymotrypsin-like serine protease that cleaves the viral polyprotein both in *cis*- and *trans*-forms (Chambers et al., 1990a,b; Li et al., 2005a,b). To function as an active enzyme, the NS3 protease requires the NS2B cofactor (Falgout et al., 1991; Jan et al., 1995; Yusof et al., 2000; Zhang et al., 1992). NS2B is an integral membrane protein of 14 kDa that contains three domains: two transmembrane segments located at the N- and C-termini and a central region of 47 amino acids (spanning amino acids 49–96) that acts as an essential protein cofactor of the NS3 protease (Clum et al., 1997). The *Flavivirus* NS3 protein is neither soluble nor catalytically active as a protease *in vitro*, suggesting that it does not fold properly without the NS2B protein, that must be either provided in *cis*- (Xu et al., 2005) or in *trans*-form (Kim et al., 2013; Phong et al., 2011; Wu et al., 2003).

4.3.2 NS3 NTPase/RNA Helicase (NS3hel)

The C-terminal domain of the NS3 protein (aa 180–618) belongs to the SF2 family (Fairman-Williams et al., 2010; Gorbalenya and Koonin, 1993). The overall structure can be broken up into three subdomains. Subdomains 1 and 2 adopt the RecA-like fold (Rao and Rossmann, 1973; Story and Steitz, 1992) and contain eight conserved motifs essential for RNA binding, ATP hydrolysis, and communication between both binding sites (Fairman-Williams et al., 2010; Gorbalenya and Koonin, 1993; Pyle, 2008). The third subdomain forms the ss-RNA binding tunnel. There is also an evidence suggesting that subdomain 3 mediates the interaction between NS3 and NS5, and hence disrupting this interaction could constitute a strategy for the design of antiviral compounds (Brooks et al., 2002; Fang et al., 2013; Tay et al., 2015). NS3 also has RTPase activity, which shares the same active site for ATP binding and hydrolysis (Wang et al., 2009). RNA 5' triphosphate hydrolysis is the first step for viral RNA capping (Decroly et al., 2012). Viruses carrying a defective or impaired NS3 helicase gene cannot replicate properly, indicating an essential role for NS3 helicase/RTPase activity in virus replication (Matusan et al., 2001).

4.4 Hepatitis C Virus

HCV belongs to Flaviviridae family, which also includes other viruses, for example, St. Louis encephalitis virus and WNV, responsible for severe human diseases, such as yellow and Dengue fevers. HCV is an enveloped virus having a 9.6-kb plus-strand RNA genome that encodes a long polyprotein precursor of ~3000 amino acids (Rosenberg, 2001), which is proteolytically processed by viral and cellular proteases to produce structural (nucleocapsid, E1, and

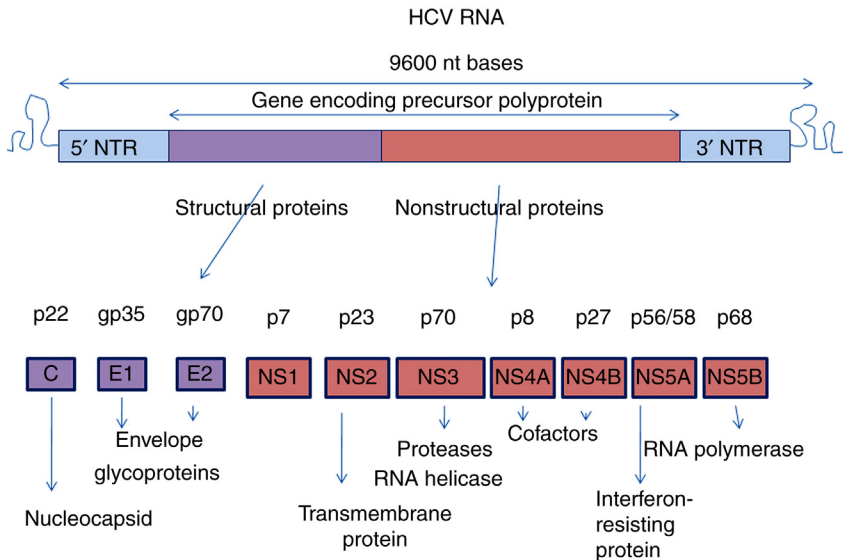


FIGURE 1.4 Diagrammatic representations of structural and nonstructural proteins in a hepatitis C virus (HCV) RNA genome. *NTR*, Nontranslated region. (Image by Colm, G. Available from Wikimedia Commons: https://commons.wikimedia.org/wiki/File:HCV_genome.png)

E2) (Santolini et al., 1994) and nonstructural proteins (NS1, NS2, NS3, NS4A, NS4B, NS5A, and NS5B) (Fig. 1.4). The structural glycosylated proteins, E1 and E2, act as the ligands for cellular receptors. For HCV entry, human CD81 receptors are required, which bind directly with E2 protein of HCV. This CD81 receptor is a widely distributed cell surface tetraspanin that participates in the formation of molecular complexes on various cell types, such as hepatocytes, natural killer cells, B lymphocytes, and T lymphocytes. Thus HCV not only exploits hepatocytes, but also modulates the host immune response. About 130 million people, estimating about 3% of the global population, are infected with HCV. Chronic HCV infection may lead to progressive liver injury, cirrhosis, and in some cases hepatocellular carcinoma (Sun et al., 2011).

4.4.1 HCV Protease

One of the HCV proteases that is responsible for viral replication in the cell cycle, is NS3-4A, a serine protease, which is a noncovalent heterodimer consisting of a catalytic subunit NS3 (the N-terminal one-third of NS3 protein) and an activating cofactor NS4A protein. NS3-4A has been considered as one of the most attractive targets for developing novel anti-HCV therapies. NS3 is also known as p70, a 70-kDa cleavage product of HCV polyprotein. A cleavage has to occur between NS2 and NS3 for the catalytic activity of NS3, and this is done by NS2-3 protease, an enzyme responsible for

catalytic cleavage. Structural studies have confirmed that NS3 protease is a chymotrypsin-like serine protease with a catalytic triad of Ser139-His57-Asp81 (Bazan and Fletterick, 1989; Chambers et al., 1990a,b; Gorbalenya et al., 1989; Miller and Purcell, 1990). Its crystal structure has been well characterized (De Francesco and Steinkuhler, 2000; Morikawa et al., 2011; Raney et al., 2010). The protease contains two β -barrels and a 30-residue extension at the N-terminus. The C-terminus, which is two-third of NS3, has helicase activity and is located in the active site of NS3 protease domain. The zinc ion also plays a vital role in the stabilization of the enzyme structure and is coordinated by three cysteine residues (Cys97, Cys99, and Cys145) and a histidine residue (His149).

4.5 Herpes Viruses

Herpes viruses are enveloped dsDNA viruses that cause diseases in animals, as well as humans. They belong to Herpesviridae family of viruses. The Herpesvirus Study Group of the International Committee on the Taxonomy of Viruses has divided the herpes viruses into three subfamilies, termed α -Herpesvirinae, β -Herpesvirinae, and γ -Herpesvirinae. The members of α -Herpesvirinae subfamily are characterized by a variable host range, a short replicative cycle in the host, rapid growth and spread in cell culture, and the establishment of latent infections in sensory ganglia (Roizman et al., 1981). They are often referred to as neurotropic herpesviruses. Among the human herpesviruses (HHVs), herpes simplex virus 1 (HSV1), herpes simplex virus 2 (HSV2), and varicella zoster virus (VZV) belong to the α -Herpesvirinae subfamily. The β -Herpesvirinae members are characterized by a restricted host range with a long reproductive cycle in cell culture and in the infected host, which often results in the development of a carrier state. Latency is established in lymphocytes, secretory glands, and cells of the kidney, as well as other cell types (Kimberlin, 1998; Roizman et al., 1981). The members of this subfamily are HHV6, HHV7, and cytomegalovirus (CMV). In another subfamily, γ -Herpesvirinae, which is characterized by a restricted host range with replication and latency occurring in lymphoid tissues, some members have demonstrated lytic growth in epithelial, endothelial, and fibroblastic cells (Kieff and Liebowitz, 1990). In lymphoblastic cells, viral replication is restricted to either T or B cells. The γ -Herpesvirinae subfamily is further divided into two genera, *Lymphocryptovirus* and *Rhadinovirus*. Among the HHVs, Epstein–Barr virus (EBV) is a member of the *Lymphocryptovirus* genus, while the newly discovered HHV8 is a member of the *Rhadinovirus* genus.

The icosahedral nucleocapsids of herpesviruses have diameter of about 1250 Å. The genomes are between 130 and 250 kbp. They consist of four distinct components: the core, capsid, tegument, and envelope (Roizman and Sears, 1993). The core consists of 162 capsomeres and a dsDNA genome arranged in a torus shape that is located inside an icosadeltahedral capsid that is

approximately 100 nm in size (Furlong et al., 1972). Tegument is an amorphous structure located between the capsid and the viral envelope that contains numerous proteins. Its structure is generally asymmetrical (Roffman et al., 1990). Presumably, the tegument is responsible for connecting the capsid to the envelope and acting as a reservoir for viral proteins that are required during the initial stages of viral infection (Batterson and Roizman, 1983; Pellett et al., 1985). The herpes virion's outermost structure is enveloped, which is derived from cell nuclear membranes and contains several viral glycoproteins (gps). The size of mature herpesviruses ranges from 120 to 300 nm, owing to differences in the size of the individual viral teguments. Cellular entry of infectious virus involves the attachment to a cell surface receptor followed by fusion of the outer viral envelope with a cell membrane. The genomic DNA is transported through the cell nuclear membrane into the nucleus, where transcription and replication occur.

4.5.1 *Herpesvirus Protease*

Herpesvirus protease is required for the life cycle of virus, as it carries out the maturational processing of the viral assembly protein. It was first identified in 1991 from studies on HSV1 and HCMV. The assembly protein is required to form a scaffold required for the formation of herpesvirus capsid. A temperature-sensitive mutant of HSV1 is not able to package the viral genome and produces empty capsids at the nonpermissive temperature (Gao et al., 1994; Preston et al., 1983), confirming the functional requirement of protease in completion of viral life cycle and spread of virus. The protease and assembly proteins are encoded by overlapping genes. The gene for the assembly protein uses the 3' portion of the gene for the protease. The maturational processing of the assembly protein occurs at M-site adjacent to C-terminus of assembly protein. Additionally, the protease catalyzes cleavage at R-site, which releases an N-terminal fragment of about 250 residues from the full-length protease gene product. The N-fragment retains all catalytic activity of the protease protein, and is generally referred to as the herpesvirus protease. In HCMV protease, two additional cleavages are catalyzed by an enzyme within the protease itself at the residues 143 and 209. The cleavage at residue 143 produces a two-chain form of the protease that is still catalytically active (Holwerda et al., 1994; O'Boyle et al., 1995). In solution, herpesvirus protease exists in monomer–dimer equilibrium, where the dimer is the only active form. The crystal structures of all herpesvirus proteases have a novel polypeptide backbone of serine proteases. It contains a central, mostly antiparallel, seven-stranded β -barrel, which is surrounded by eight helices. The active site of the protease contains a novel Ser132-His63-His157 catalytic triad. The first two residues, Ser132 and His63, are determined from biochemical and mutagenesis studies, while the third one, His157, is not fully understood, but its removal can lead to a loss of catalytic activity by >20,000-fold.

4.6 Human Immunodeficiency Virus of Type 1

The HIV is an enveloped retrovirus with a positive-sense RNA genome. It belongs to the genus *Lentiviridae* (Horwitz, 2001). HIV is of two types, HIV1 and HIV2, where HIV1 was found to be the causative agent of AIDS (Gallo and Montagnier, 1988). These viruses store their genetic information as ribonucleic acid (RNA), while most viruses do so as deoxyribonucleic acid (DNA). Before viral replication can take place, the RNA must be converted to DNA by reverse transcriptase (RT) enzyme. HIV1 has an outer envelope consisting of a lipid bilayer with spikes of glycoproteins (gp), gp41 and gp120. These gps are attached in such a way that gp120 protrudes from the surface of the virus. Inside this envelope is a nucleocapsid (p17), which surrounds a central core of protein, p24. The core contains two copies of ssRNA (the virus genome). Gene expression is believed to be regulated by proteins, p7 and p9, that are bound to the RNA. Multiple molecules of the enzyme reverse transcriptase are found in the core. This enzyme is responsible for converting the viral RNA into proviral DNA (Abbas et al., 2000). HIV particles are spherical with a diameter of about 1000 Å. RNA genome of nucleocapsid is about 13 kbp in length. Upon entry into the host cell, the viral RNA is transcribed to DNA that integrates with host genome to produce the provirus. HIV1 infection leads to impairment of immune cells, mainly CD4⁺ T cells, monocytes, and macrophages, that ultimately about after 10–15 years leads to AIDS, which is characterized by low CD4⁺ cell counts. It is estimated that about 36 million people worldwide are affected by this virus. HIV1 transmission takes place by sexual contact, contaminated needles, vertical transmission from mother to child, blood transfusions, blood products, and organ/tissue transplants (Negishi, 1993). Various neurological complications affecting the nervous system have been seen in HIV-infected individuals. These include rare opportunistic infections and neoplastic diseases, such as cerebral toxoplasmosis, cryptococcal meningitis, and primary central nervous system lymphoma, as well as syndromes caused by or directly related to HIV1 infection, such as dementia and various neuromuscular complications (Krebs et al., 2000). As the immune system of HIV1-infected people becomes weak, opportunistic infections can also occur by organisms, such as bacteria, viruses, fungi, and parasites (Schneider, 2000).

4.6.1 HIV1 Protease

HIV1 protease plays a major in viral maturation, which is important for the production of infectious virus particles. It is an aspartic protease and the catalytic site has the characteristic Asp-Thr-Gly (Asp25, Asp26, and Asp27) sequence common to all aspartic proteases. It is a symmetric homodimer consisting of 99 amino acids per monomer (Fig. 1.5). The protease contains a β -sheet with the two aspartic acids Asp25 from the two monomers, forming the central active site. The active site cavity, the flexible flaps, and the dimer interface are three important regions of the protease structure. Protease inhibitors are competitive inhibitors that bind at the active site of the protease with the flaps

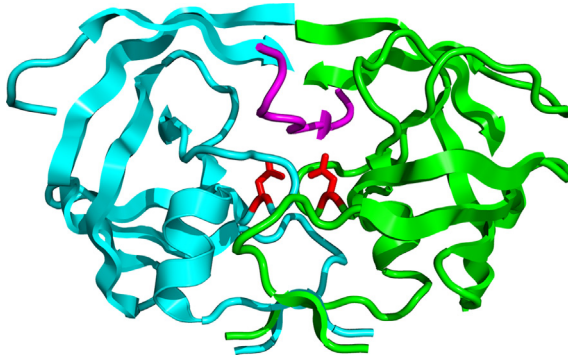


FIGURE 1.5 The human immunodeficiency virus type 1 (HIV1) protease dimer (*green and cyan*) with active site Asp25 (*red*) complexed with a polypeptide substrate (*magenta*). (Image available from Wikipedia: <https://en.wikipedia.org/w/index.php?curid=18176790>)

folded into a closed conformation over the active site. The flaps were seen in closed and open conformations in the crystal structures of inhibitor-bound, as well as free protease (Liu et al., 2006; Perryman et al., 2004; Rose et al., 1998; Scott and Schiffer, 2000; Spinelli et al., 1991). This process of opening and closing of the flaps enables the substrate to enter and leave the active site of protease. Inhibitor bound at the active site with the flaps in closed conformations keeps the enzyme in a locked-down state and prevents the processing of substrates. The protease active site cavity comprises the residues Arg8, Leu23, Asp25, Gly27, Ala28, Asp29, Asp30, Val32, Lys45, Ile47, Met46, Gly48, Gly49, Ile50, Phe53, Leu76, Thr80, Pro81, Val82, and Ile84. The residues forming the substrate-binding site are mainly hydrophobic in nature; the exceptions are the catalytic residues Asp25 and Asp29, which form hydrogen bonds with the peptide main chain groups, and Arg8, Asp30, and Lys45, which can interact with polar side chains or distal main chain groups in longer peptides (Tie et al., 2005).

4.7 Picornaviruses

Picornaviruses are small, nonenveloped viruses, having positive-sense ssRNA genome, and belong to the family of Picornaviridae (Racaniello, 2001). This family consists of 46 species having 26 genera, the best known are *Enterovirus* (poliovirus, rhinovirus, coxsackievirus, and echovirus), *Aphthovirus* (FMDV), *Cardiovirus* (encephalomyocarditis virus, EMCV), Thieler's virus, and *Hepatovirus* (HAV) (Knowles et al., 2012). The diameter of picornaviruses is about 300 Å. Replication begins with the attachment of the virus to a specific cellular receptor that leads to virus internalization and destabilization of the capsid and release of the genome from the endosome into the cytoplasm (Bergelson, 2010). After binding to the receptor, the virus uses host and viral proteins to complete its replication cycle (Fig. 1.6). In fact, to complete a round of infection,

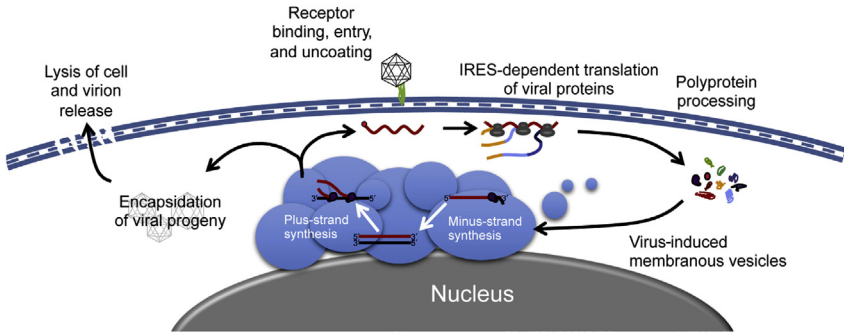


FIGURE 1.6 Overview of the picornavirus replication cycle. *IRES*, Internal ribosome entry site. (From Cathcart, A.L., Baggs, E.L., Semler, B.L., 2014. *Picornaviruses: Pathogenesis and Molecular Biology. Reference Module in Biomedical Sciences, third ed. Elsevier, UC, Irvine, (pp. 1–13.)*)

a number of complex activities are performed and coordinated. These include replication of the plus-strand RNA via minus-strand intermediates, translation, proteolytic processing, inhibition of host cell transcription/translation, virion assembly, and cell lysis.

4.7.1 Picornavirus Proteases

The viral genome, having about 8000 bases, encodes a single polyprotein, which is processed cotranslationally by proteases including the 3C, 2A, and leader proteases (Kay and Dunn, 1990; Lawson and Semler, 1990; Racaniello, 2001; Ryan and Flint, 1997). Most of the processing is carried out by 3C, while 2A and leader proteases are present in only some of the picornaviruses. The 2A protease catalyzes the release of structural polyprotein at the N-terminus, while the leader protease releases itself from the polyprotein located at C-terminus (Strelbel and Beck, 1986). The 2A and leader proteases also inhibit host cell protein synthesis by cleaving host cell eukaryotic initiation factors 4G (eIF4G) during translation from 5' capped mRNA of the host cells, whereas translation from viral mRNA is not affected (Racaniello, 2001). It has been found recently that cleaved eIF4G can still support translation from capped mRNA, but not as efficiently as from viral RNA (Ali et al., 2001). Additionally, other enzymes can also cleave eIF4G, including 3C protease of FMDV (Belsham et al., 2000), HIV protease (Ventoso et al., 2001), cellular caspases, and other proteases (Zamora et al., 2002).

The 3C and 2A proteases are cysteine proteases (Kräusslich and Wimmer, 1998), as they both contain a Gly-X-Cys-Gly motif, similar to that of the active site of chymotrysin-like serine proteases, which is Gly-Asp-Ser-Gly (Bazan and Fletterick, 1989; Choi et al., 1997; Zamora et al., 2002). Crystal structures of these proteases present in picornaviruses and other viruses, such as HAV and human rhinovirus 2 and 14, confirm that they are chymotrysin-like serine proteases (Allaire et al., 1994; Bergmann et al., 1997; Matthews

et al., 1994; Mosimann et al., 1997; Peterson et al., 1999; Seipelt et al., 1999). The 3C protease of human rhinovirus 14 (HRV14) has 182 residues, while 2A protease of HRV2 has 142 residues (Peterson et al., 1999)

The active site of 3C protease of HRV14 contains a catalytic triad of Cys147-His40-Glu71, while in HAV 3C protease Glu71 is replaced by Tyr143. The catalytic triad of 2A proteases of HRV2 contains Cys106, His18, and Asp35. Crystal structures of the FMDV leader protease has confirmed that it is also a papain-like cysteine protease having 150 residues, with a catalytic triad of Cys51, His148, and Asp175 (Gorbalenya et al., 1991; Guarne et al., 1998, 2000; Piccone et al., 1995; Roberts and Belsham, 1995; Seipelt et al., 1999). This leader protease contributes to the inhibition of host protein synthesis by cleaving eIF4G at a different site than that of 2A.

5 CONCLUSIONS

All of the prominent viruses, namely AdVs, alphaviruses, flaviviruses, HCV, herpesviruses, HIV1, and picornaviruses, contain proteases that play crucial roles in their replication and thus are important targets for the discovery of potent antiviral drugs. The process of protein catabolism by hydrolysis of peptide bond is catalyzed by the enzyme called protease. It may also be known as peptidase or proteinase. Different classes of protease can perform the same reaction by completely different catalytic mechanisms. The catalytic mechanism of these proteases involves either the activation of a water molecule, which performs a nucleophilic attack on the peptide bond to hydrolyze it or uses a nucleophilic residue in a catalytic triad to perform a nucleophilic attack that covalently links the protease to the substrate protein, releasing the first half of the product. Regarding the specificity of proteases, it is said that proteolysis is highly promiscuous, such that a wide range of protein substrates are hydrolyzed. Proteases, being proteins themselves, are cleaved by other protease molecules, sometimes of the same variety. Some proteases, such as TEV protease, have high specificity and cleave only a very restricted set of substrate sequences. Virus proteases are now extensively studied so that they can be exploited for drug development.

REFERENCES

- Abbas, A.K., Lichtman, A.H., Pober, J.S., 2000. Cellular and Molecular Immunology. W. B. Saunders, Philadelphia, PA.
- Ali, I.K., McKendrick, L., Morley, S.J., Jackson, R.J., 2001. Truncated initiation factor eIF4G lacking an eIF4E binding site can support capped mRNA translation. *EMBO J.* 20, 4233–4242.
- Allaire, M., Chernaia, M.M., Malcolm, B.A., James, M.N., 1994. Picornaviral 3C cysteine proteinases have a fold similar to chymotrypsin-like serine proteinases. *Nature* 369, 72–76.
- Alvarez, D.E., Lodeiro, M.F., Filomatori, C.V., Fucito, S., Mondotte, J.A., Gamarnik, A.V., 2006. Structural and functional analysis of dengue virus RNA. *Novartis Found. Symp.* 277, 120–132, (Discussion 132-125, 251-123).
- Baltimore, D., 1971. Expression of animal virus genomes. *Bacteriol. Rev.* 35 (3), 235–241.

- Batterson, W., Roizman, B., 1983. Characterization of the herpes simplex virion-associated factor responsible for the induction of alpha genes. *J. Virol.* 46, 371–377.
- Bazan, J.F., Fletterick, R.J., 1989. Detection of a trypsin-like serine protease domain in flaviviruses and pestiviruses. *Virology* 171, 637–639.
- Belsham, G.J., MacLnerney, G.M., Ross-Smith, N.J., 2000. Foot-and-mouth disease virus 3C protease induces cleavage of translation initiation factors eIF4A and eIF4G within infected cells. *J. Virol.* 74, 272–280.
- Bergelson, J.M., 2010. Receptors. In: Ehrenfeld, E., Domingo, E., Roos, R.P. (Eds.), *The Picornaviruses*. ASM Press, Washington, DC, pp. 73–86.
- Bergmann, E.M., Mosimann, S.C., Chernaia, M.M., Malcolm, B.A., James, M.N.G., 1997. The refined crystal structure of the 3C gene product from hepatitis A virus: specific proteinase activity and RNA recognition. *J. Virol.* 71, 2436–2448.
- Bollati, M., Alvarez, K., Assenberg, R., Baronti, C., Canard, B., Cook, S., Coutard, B., Decroly, E., de Lamballerie, X., Gould, E.A., Grard, G., Grimes, J.M., Hilgenfeld, R., Jansson, A.M., Malet, H., Mancini, E.J., Mastrangelo, E., Mattevi, A., Milani, M., Moureau, G., Neyts, J., Owens, R.J., Ren, J., Selisko, B., Speroni, S., Steuber, H., Stuart, D.I., Unge, T., Bolognesi, M., 2010. Structure and functionality in *Flavivirus* NS-proteins: perspectives for drug design. *Antiviral Res.* 87, 125–148.
- Brooks, A.J., Johansson, M., John, A.V., Xu, Y., Jans, D.A., Vasudevan, S.G., 2002. The interdomain region of dengue NS5 protein that binds to the viral helicase NS3 contains independently functional importin beta 1 and importin alpha/beta recognized nuclear localization signals. *J. Biol. Chem.* 277, 36399–36407.
- Cancedda, R., Shatkin, A.J., 1979. Ribosome-protected fragments from Sindbis 42-S and 26-S RNAs. *Eur. J. Biochem.* 94, 41–50.
- Chambers, T.J., McCourt, D.W., Rice, C.M., 1990a. Production of yellow fever virus proteins in infected cells: identification of discrete polyprotein species and analysis of cleavage kinetics using region-specific polyclonal antisera. *Virology* 177, 159–174.
- Chambers, T.J., Weir, R.C., Grakoui, A., McCourt, D.W., Bazan, J.F., Fletterick, R.J., Rice, C.M., 1990b. 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. *Proc. Natl. Acad. Sci. USA* 87, 8898–8902.
- Choi, H.K., Lu, G., Lee, S., Wengler, G., Rossmann, M.G., 1997. Structure of Semliki Forest virus core protein. *Proteins Struct. Funct. Genet.* 27, 345–359.
- Clum, S., Ebner, K.E., Padmanabhan, R., 1997. Cotranslational membrane insertion of the serine proteinase precursor NS2B–NS3(Pro) of dengue virus type 2 is required for efficient *in vitro* processing and is mediated through the hydrophobic regions of NS2B. *J. Biol. Chem.* 272, 30715–30723.
- de Cedrón, M.G., Ehsani, N., Mikkola, M.L., Kääriäinen, L., García, J.A., 1999. RNA helicase activity of Semliki Forest virus replicase protein NSP2. *FEBS Lett.* 448, 19–22.
- Dé, I., Sawicki, S.G., Sawicki, D.L., 1996. Sindbis virus RNA-negative mutants that fail to convert from minus-strand to plus-strand synthesis: role of the nsP2 protein. *J. Virol.* 70, 2706–2719.
- De Francesco, R., Steinkuhler, C., 2000. Structure and function of the hepatitis C virus NS3–NS4A serine proteinase. *Curr. Top. Microbiol. Immunol.* 242, 149–169.
- Decroly, E., Ferron, F., Lescar, J., Canard, B., 2012. Conventional and unconventional mechanisms for capping viral mRNA. *Nat. Rev. Microbiol.* 10, 51–65.
- Ding, J., McGrath, W.J., Sweet, R.M., Mangel, W.F., 1996. Crystal structure of the human adenovirus proteinase with its 11 amino acid cofactor. *EMBO J.* 15, 1778–1783.

- Dong, H., Chang, D.C., Hua, M.H., Lim, S.P., Chionh, Y.H., Hia, F., Lee, Y.H., Kukkaro, P., Lok, S.M., Dedon, P.C., Shi, P.Y., 2012. 20-O methylation of internal adenosine by *Flavivirus* NS5 methyltransferase. *PLoS Pathog.* 8, e1002642.
- Dong, H., Fink, K., Zust, R., Lim, S.P., Qin, C.F., Shi, P.Y., 2014. *Flavivirus* RNA methylation. *J. Gen. Virol.* 95, 763–778.
- Egloff, M.P., Benarroch, D., Selisko, B., Romette, J.L., Canard, B., 2002. An RNA cap (nucleoside-20-O)-methyltransferase in the *Flavivirus* RNA polymerase NS5: crystal structure and functional characterization. *EMBO J.* 21, 2757–2768.
- Fairman-Williams, M.E., Guenther, U.P., Jankowsky, E., 2010. SF1 and SF2 helicases: family matters. *Curr. Opin. Struct. Biol.* 20, 313–324.
- 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. *J. Virol.* 65, 2467–2475.
- Fang, D.A., Huang, X.M., Zhang, Z.Q., Xu, D.P., Zhou, Y.F., Zhang, M.Y., Liu, K., Duan, J.R., Shi, W.G., 2013. Molecular cloning and expression analysis of chymotrypsin-like serine protease from the redclaw crayfish (*Cherax quadricarinatus*): a possible role in the junior and adult innate immune systems. *Fish Shellfish Immunol.* 34, 1546–1552.
- Filomatori, C.V., Lodeiro, M.F., Alvarez, D.E., Samsa, M.M., Pietrasanta, L., Gamarnik, A.V., 2006. A 50 RNA element promotes dengue virus RNA synthesis on a circular genome. *Genes Dev.* 20, 2238–2249.
- Froshauer, S.A., Kartenbeck, J., Helenius, A., 1988. Alphavirus RNA replicase is located on the cytoplasmic surface of endosomes and lysosomes. *J. Cell. Biol.* 107, 2075–2086.
- Fuller, S.D., 1987. The T = 4 envelope of Sindbis virus is organized by interactions with a complementary T = 3 capsid. *Cell* 48, 923–934.
- Furlong, D., Swift, H., Roizman, B., 1972. Arrangement of herpesvirus deoxyribonucleic acid in the core. *J. Virol.* 10, 1071–1074.
- Gallo, R.C., Montagnier, L., 1988. AIDS in 1988. *Sci. Am.* 259, 41–48.
- Gao, M., Matusick-Kumar, L., Hurlburt, W., DiTusa, S.F., Newcomb, W.W., Brown, J.C., McCann, III, P.J., Deckman, I., Colonno, R.J., 1994. The protease of herpes simplex virus type 1 is essential for functional capsid formation and viral growth. *J. Virol.* 68, 3702–3712.
- Glanville, N.T., Ranki, M., Morser, J., 1976. Initiation of translocation directed by 42S and 26S RNAs from Semliki Forest virus in vitro. *Proc. Natl. Acad. Sci. USA* 73, 3059–3063.
- Gorbalenya, A.E., Koonin, E.V., 1993. Helicases: amino acid sequence comparisons and structure–function relationships. *Curr. Opin. Struct. Biol.* 3, 419–429.
- Gorbalenya, A.E., Donchenko, A.P., Koonin, E.V., Blinov, V.M., 1989. N-terminal domains of putative helicases of flavi- and pestiviruses may be serine proteases. *Nucleic Acids Res.* 17, 3889–3897.
- Gorbalenya, A.E., Koonin, E.V., Lai, M.M., 1991. Putative papain-related thiol proteases of positive-strand RNA viruses. Identification of rubi- and aphthovirus proteases and delineation of a novel conserved domain associated with proteases of rubi-, alpha- and coronaviruses. *FEBS Lett.* 288, 201–205.
- Guarne, A., Hampoelz, B., Glaser, W., Carpena, X., Tormo, J., Fita, I., Skern, T., 2000. Structural and biochemical features distinguish the foot-and-mouth disease virus leader proteinase from other papain-like enzymes. *J. Mol. Biol.* 302, 1227–1240.
- Guarne, A., Tormo, J., Kirchweber, R., Pfistermueller, D., Fita, I., Skern, T., 1998. Structure of the foot-and-mouth disease virus leader protease: a papain-like fold adapted for self-processing and eIF4G recognition. *EMBO J.* 17, 7469–7479.

- Gubler, D.J., Kuno, G., Markoff, L., 2007. Flaviviruses. In: Knipe, D.M., Howley, P.M. (Eds.), *Fields Virology*. fifth ed. Lippincott, Williams, and Wilkins, Philadelphia, PA, pp. 1153–1252.
- Hardy, W.R., Strauss, J., 1989. Processing the nonstructural polyproteins of Sindbis virus: non-structural proteinase is in the C-terminal half of nsP2 and functions both in *cis* and in *trans*. *J. Virol.* 63, 4653–4664.
- Holwerda, B.C., Wittwer, A.J., Duffin, K.L., Smith, C., Toth, M.V., Cam, L.S., Wiegand, R.C., Bryant, M.L., 1994. Activity of two-chain recombinant human cytomegalovirus protease. *J. Biol. Chem.* 269, 25911–25915.
- Horwitz, M., 2001. *Fields, B.M., Knipe, D.M., Howley, P.M. (Eds.), Fields Virology*, vol. 2, Lippincott William and Wilkins, Philadelphia, PA, (International Committee on taxonomy of viruses, 2002, Lentivirus National Institute of Health).
- Jan, L.R., Yang, C.S., Trent, D.W., Falgout, B., Lai, C.J., 1995. Processing of Japanese encephalitis virus non-structural proteins: NS2B–NS3 complex and heterologous proteases. *J. Gen. Virol.* 76 (Pt. 3), 573–580.
- Karpe, Y.A., Aher, P.P., Lole, K.S., 2011. NTPase and 5′-RNA triphosphatase activities of Chikungunya virus nsP2 protein. *PLoS One* 6, e22336.
- Kay, J., Dunn, B.M., 1990. Viral proteinases: weakness in strength. *Biochim. Biophys. Acta* 1048, 1–18.
- Khromykh, A.A., Kondratieva, N., Sgro, J.Y., Palmenberg, A., Westaway, E.G., 2003. Significance in replication of the terminal nucleotides of the *Flavivirus* genome. *J. Virol.* 77, 10623–10629.
- Kieff, E., Liebowitz, D., 1990. Epstein-Barr virus and its replication. In: *Fields, B.M., Knipe, D.M., Chanock, R.M., Hirsch, M.S., Melnick, J.L., Monath, T.P. et al., (Eds.), Fields Virology*. second ed. Raven Press, New York, NY, pp. 1889–1920.
- Kim, Y.M., Gayen, S., Kang, C., Joy, J., Huang, Q., Chen, A.S., Wee, J.L., Ang, M.J., Lim, H.A., Hung, A.W., Li, R., Noble, C.G., Lee, T., Yip, A., Wang, Q.Y., Chia, C.S., Hill, J., Shi, P.Y., Keller, T.H., 2013. NMR analysis of a novel enzymatically active unlinked dengue NS2B–NS3 protease complex. *J. Biol. Chem.* 288, 12891–12900.
- Kimberlin, D.W., 1998. Human herpesviruses 6 and 7: identification of newly recognized viral pathogens and their association with human disease. *Pediatr. Infect. Dis. J.* 17, 59–68.
- Knowles, N.J., Hovi, T., Hyypiä, T., King, A.M.Q., Lindberg, A.M., Pallansch, M.A., Palmenberg, A.C., Simmonds, P., Skern, T., et al., 2012. Picornaviridae. In: *King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J. (Eds.), Virus Taxonomy: Classification and Nomenclature of Viruses: Ninth Report of the International Committee on Taxonomy of Viruses*. Elsevier, San Diego, CA, pp. 855–880.
- Krebs, F.C., Ross, H., McAllister, J., Wigdahl, B., 2000. HIV-1-associated central nervous system dysfunction. *Adv. Pharmacol.* 49, 315–385.
- Kräusslich, H.G., Wimmer, E., 1998. Viral proteinases. *Annu. Rev. Biochem.* 57, 701–754.
- Kujala, P., Ikäheimonen, A., Ehsani, N., Vihinen, H., Kääriäinen, L., Auvinen, P., 2001. Biogenesis of the Semliki Forest virus RNA replication complex. *J. Virol.* 75, 3873–3884.
- LaStarza, M.W., Lemm, J.A., Rice, C.M., 1994. Genetic analysis of the nsP3 region of Sindbis virus: evidence for roles in minus-strand and subgenomic RNA synthesis. *J. Virol.* 68, 5781–5791.
- Lawson, M.A., Semler, B.L., 1990. Picornavirus protein processing—enzymes, substrates, and genetic regulation. *Curr. Top. Microbiol. Immunol.* 161, 49–87.
- Lescar, J., Luo, D., Xu, T., Sampath, A., Lim, S.P., Canard, B., Vasudevan, S.G., 2008. Towards the design of antiviral inhibitors against flaviviruses: the case for the multifunctional NS3 protein from Dengue virus as a target. *Antiviral Res.* 80, 94–101.
- Leung, J.Y., Ng, M.M., Chu, J.J., 2011. Replication of alphaviruses: a review on the entry process of alphaviruses into cells. *Adv. Virol.* 2011, (Article ID 249640).

- Li, J., Lim, S.P., Beer, D., Patel, V., Wen, D., Tumanut, C., Tully, D.C., Williams, J.A., Jiricek, J., Priestle, J.P., Harris, J.L., Vasudevan, S.G., 2005a. Functional profiling of recombinant NS3 proteases from all four serotypes of Dengue virus using tetrapeptide and octapeptide substrate libraries. *J. Biol. Chem.* 280, 28766–28774.
- Li, X.D., Sun, L., Seth, R.B., Pineda, G., Chen, Z.J., 2005b. Hepatitis C virus protease NS3/4A cleaves mitochondrial antiviral signaling protein off the mitochondria to evade innate immunity. *Proc. Natl. Acad. Sci. USA* 102, 17717–17722.
- Lim, S.P., Wang, Q.Y., Noble, C.G., Chen, Y.L., Dong, H., Zou, B., Yokokawa, F., Nilar, S., Smith, P., Beer, D., Lescar, J., Shi, P.Y., 2013. Ten years of dengue drug discovery: progress and prospects. *Antiviral Res.* 100, 500–519.
- Lindenbach, B.D., Thiel, H.J., Rice, C.M., 2007. *Flaviviridae: The Viruses and Their Replication*. Lippincott-Raven Publishers, Philadelphia, PA.
- Liu, F., Kovalevsky, A.Y., Louis, J.M., Boross, P.I., Wang, Y.F., Harrison, R.W., Weber, I.T., 2006. Mechanism of drug resistance revealed by the crystal structure of the unliganded HIV-1 protease with F53L mutation. *J. Mol. Biol.* 358, 1191–1199.
- Luo, D., Wei, N., Doan, D.N., Paradkar, P.N., Chong, Y., Davidson, A.D., Kotaka, M., Lescar, J., Vasudevan, S.G., 2010. Flexibility between the protease and helicase domains of the dengue virus NS3 protein conferred by the linker region and its functional implications. *J. Biol. Chem.* 285, 18817–18827.
- Malet, H., Massé, N., Selisko, B., Romette, J.-L., Alvarez, K., Guillemot, J.-C., Tolou, H., Yap, T.L., Vasudevan, S.G., Lescar, J., Canard, B., 2008. The *Flavivirus* polymerase as a target for drug discovery. *Antiviral Res.* 80, 23–35.
- Mancini, E.J., Clarke, M., Gowen, B., Rutten, T., Fuller, S.D., 2000. Cryo-electron microscopy reveals the functional organization of an enveloped virus, Semliki forest virus. *Mol. Cell.* 5, 255–266.
- Marcin, S., Marcin, S., 2013. Viral Proteases as targets for drug design. *Curr. Pharm. Design* 19, 1126–1153.
- Markoff, L., 2003. 5′- and 3′-noncoding regions in *Flavivirus* RNA. *Adv. Virus Res.* 59, 177–228.
- Martin, M.A., Knipe, D.M., Fields, B.N., Howley, P.M., Griffin, D., Lamb, R., 2007. *Fields' Virology*. Wolters Kluwer Health/Lippincott Williams and Wilkins, Philadelphia, PA, (p. 2395).
- Matthews, D.A., Smith, W.W., Ferre, R.A., Condon, B., Budahazi, G., McElroy, H.E., Gribskov, C.L., Worland, S., 1994. Structure of human rhinovirus 3C protease reveals a trypsin-like polypeptide fold, RNA-binding site, and means for cleaving precursor polyprotein. *Cell* 77, 761–771.
- Matusan, A.E., Pryor, M.J., Davidson, A.D., Wright, P.J., 2001. Mutagenesis of the dengue virus type 2 NS3 protein within and outside helicase motifs: effects on enzyme activity and virus replication. *J. Virol.* 75, 9633–9643.
- Melançon, P., Garoff, H., 1987. Processing of the Semliki forest virus structural polyprotein: role of the capsid F protease. *J. Virol.* 61, 1301–1309.
- Merits, A., Vasiljeva, L., Ahola, T., Kaariainen, L., Auvinen, P., 2001. Proteolytic processing of Semliki forest virus-specific nonstructural polyprotein by nsP2 protease. *J. Gen. Virol.* 82, 765–773.
- Miller, R.H., Purcell, R.H., 1990. Hepatitis C virus shares amino acid sequence similarity with pestiviruses and flaviviruses as well as members of two plant virus supergroups. *Proc. Natl. Acad. Sci. USA* 87, 2057.
- Morgan, C., Howe, C., Rose, H.M., 1961. Structure and development of viruses as observed in the electron microscope. V. Western equine encephalomyelitis virus. *J. Exp. Med.* 113, 219–234.

- Morikawa, K., Lange, C.M., Gouttenoire, J., Meylan, E., Brass, V., Penin, F., Moradpour, D., 2011. Non-structural protein 3-4A: the Swiss army knife of hepatitis C virus. *J. Viral Hepat.* 18, 305–315.
- Mosimann, S.C., Chernaia, M.M., Sia, S., Plotch, S., James, M.N.G., 1997. Refined X-ray crystallographic structure of the poliovirus 3C gene product. *J. Mol. Biol.* 273, 1032–1047.
- Murray, C.L., Jones, C.T., Rice, C.M., 2008. Architects of assembly: roles of Flaviviridae non-structural proteins in virion morphogenesis. *Nat. Rev. Microbiol.* 6, 699–708.
- Negishi, M., 1993. Preventive methods against HIV transmission. *Nippon Rinsho* 51, 509–511.
- Noble, C.G., Shi, P.Y., 2012. Structural biology of dengue virus enzymes: towards rational design of therapeutics. *Antiviral Res.* 96, 115–126.
- O'Boyle, II, D.R., Wager-Smith, K., Stevens, III, J.T., Wein Heimer, S.P., 1995. The effect of internal autocleavage on kinetic properties of the human cytomegalovirus protease catalytic domain. *J. Biol. Chem.* 270, 4753–4758.
- Oda, K., 2012. New families of carboxyl peptidases: serine-carboxyl peptidases and glutamic peptidases. *J. Biochem.* 151, 13–25.
- Pastorino, B.A., Peyrefitte, C.N., Almeras, L., Grandadam, M., Rolland, D., Tolou, H.J., Bessaud, M., 2008. Expression and biochemical characterization of nsP2 cysteine protease of Chikungunya virus. *Virus Res.* 131, 293–298.
- Paul, D., Bartenschlager, R., 2013. Architecture and biogenesis of plus-strand RNA virus replication factories. *World J. Virol.* 2, 32–48.
- Pellett, P.E., McKnight, J.L., Jenkins, F.J., Roizman, B., 1985. Nucleotide sequence and predicted amino acid sequence of a protein encoded in a small herpes simplex virus DNA fragment capable of trans-inducing alpha genes. *Proc. Natl. Acad. Sci. USA* 82, 5870–5874.
- Perryman, A.L., Lin, J.H., McCammon, J.A., 2004. HIV-1 protease molecular dynamics of a wild-type and of the V82F/I84V mutant: possible contributions to drug resistance and a potential new target site for drugs. *Protein Sci.* 13, 1108–1123.
- Peterson, J.F.W., Cherne, M.M., Liebig, H.D., Skern, T., Kuechler, E., James, M.N.G., 1999. The structure of the 2A proteinase from a common cold virus: a proteinase responsible for the shut-off of host-cell protein synthesis. *EMBO J.* 18, 5643.
- Phong, W.Y., Moreland, N.J., Lim, S.P., Wen, D., Paradkar, P.N., Vasudevan, S.G., 2011. Dengue protease activity: the structural integrity and interaction of NS2B with NS3 protease and its potential as a drug target. *Biosci. Rep.* 31, 399–409.
- Piccone, M.E., Zellner, M., Kumosinki, T.F., Mason, P.W., Grubman, M.J.J., 1995. Identification of the active-site residues of the L proteinase of foot-and-mouth disease virus. *Virology* 69, 4950–4956.
- Polacek, C., Foley, J.E., Harris, E., 2009. Conformational changes in the solution structure of the dengue virus 5' end in the presence and absence of the 3' untranslated region. *J. Virol.* 83, 1161–1166.
- Preston, V.G., Coates, J.A.V., Rixon, F.J., 1983. Identification and characterization of a herpes simplex virus gene product required for encapsidation of virus DNA. *J. Virol.* 45, 1056–1064.
- Pyle, A.M., 2008. Translocation and unwinding mechanisms of RNA and DNA helicases. *Annu. Rev. Biophys.* 37, 317–336.
- Racaniello, V.R., 2001. Picornaviridae: the viruses and their replication. Knipe, D.M., Howley, P.M. (Eds.), *Fields Virology*, vol. 1, Lippincott Williams and Wilkins, Philadelphia, PA.
- Raney, K.D., Sharma, S.D., Moustafa, I.M., Cameron, C.E., 2010. Hepatitis C virus non-structural protein 3 (HCV NS3): a multifunctional antiviral target. *J. Biol. Chem.* 285, 22725–22731.
- Rao, S.T., Rossmann, M.G., 1973. Comparison of super-secondary structures in proteins. *J. Mol. Biol.* 76, 241–256.
- Ray, D., Shah, A., Tilgner, M., Guo, Y., Zhao, Y., Dong, H., Deas, T.S., Zhou, Y., Li, H., Shi, P.Y., 2006. West Nile virus 5'-cap structure is formed by sequential guanine N-7 and ribose 2'-O methylations by nonstructural protein 5. *J. Virol.* 80, 8362–8370.

- Roberts, P.J., Belsham, G.J., 1995. Identification of critical amino acids within the foot and mouth disease virus leader protein, a cysteine protease. *Virology* 213, 140–146.
- Roffman, E., Albert, J.P., Goff, J.P., Frenkel, N., 1990. Putative site for the acquisition of human herpesvirus 6 virion tegument. *J. Virol.* 64, 6308–6313.
- Roizman, B., Sears, A.E., 1993. Herpes simplex viruses and their replication. In: Roizman, B., Whitley, R.J., Lopez, C. (Eds.), *The Human Herpesviruses*. first ed. Raven Press, New York, NY, pp. 11–68.
- Roizman, B., Carmichael, L.E., Deinhardt, F., de-The, G., Nahmias, A.J., Plowright, W., Rap, F., Sheldrick, P., Takahashi, M., Wolf, K., 1981. Herpesviridae. Definition, provisional nomenclature, and taxonomy. The Herpesvirus Study Group, the International Committee on Taxonomy of Viruses. *Intervirolgy* 16, 201–217.
- Rose, R.B., Craik, C.S., Stroud, R.M., 1998. Domain flexibility in retroviral proteases: structural implications for drug resistant mutations. *Biochemistry* 37, 2607–2621.
- Rosenberg, S., 2001. Recent advances in the molecular biology of hepatitis C virus. *J. Mol. Biol.* 313, 451–464.
- Rowe, W.P., Huebner, R.J., Gilmore, L.K., Parrott, R.H., Ward, T.G., 1953. Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. *Proc. Soc. Exp. Biol. Med.* 84 (3), 570–573.
- Russell, W.C., 2005. Adenoviruses. In: Mahy, B.W.J., ter Meulen, V. (Eds.), *Topley and Wilson's Microbiology and Microbial Infections*. ninth ed. Hodder Arnold, London, pp. 439–447.
- Ryan, M., Flint, M., 1997. Virus-encoded proteinases of the picornavirus super-group. *J. Gen. Virol.* 78, 699–723.
- Ryman, K.D., Klimstra, W.B., 2008. Host responses to alphavirus infection. *Immunol. Rev.* 225, 27–45.
- Salonen, A.H., Vasiljeva, L., Merits, A., Kääriäinen, L., Magden, J., Jokitalo, E., 2003. Properly folded nonstructural polyprotein directs the Semliki forest virus replication complex to the endosomal compartment. *J. Virol.* 77, 1691–1702.
- Sampath, A., Padmanabhan, R., 2009. Molecular targets for *Flavivirus* drug discovery. *Antiviral Res.* 81, 6–15.
- Santolini, E., Migliaccio, G., La Monica, N., 1994. Biosynthesis and biochemical properties of the hepatitis C virus core protein. *J. Virol.* 68, 3631–3641.
- Schneider, M.M., 2000. Change in natural history of opportunistic infections in HIV-infected patients. *Neth. J. Med.* 57, 1–3.
- Scott, W.R., Schiffer, C.A., 2000. Curling of flap tips in HIV-1 protease as a mechanism for substrate entry and tolerance of drug resistance. *Structure* 8, 1259–1265.
- Segerman, A., Arnberg, N., Erikson, A., Lindman, K., Wadell, G., 2003. There are two different species B adenovirus receptors: sBAR, common to species B1 and B2 adenoviruses, and sB2AR, exclusively used by species B2 adenoviruses. *J. Virol.* 77, 1157–1162.
- Seipelt, J., Guarne, A., Bergmann, E., James, M., Sommergruber, W., Fita, I., Skern, T., 1999. The structures of picornaviral proteinases. *Virus Res.* 62, 159–168.
- Simmons, D.T., Strauss, J., 1972. Replication of Sindbis virus. I. Relative size and genetic content of 26 S and 49 S RNA. *J. Mol. Biol.* 71, 599–613.
- Spinelli, S., Liu, Q.Z., Alzari, P.M., Hirel, P.H., Poljak, R.J., 1991. The three-dimensional structure of the aspartyl protease from the HIV-1 isolate BRU. *Biochimie* 73, 1391–1396.
- Story, R.M., Steitz, T.A., 1992. Structure of the recA protein–ADP complex. *Nature* 355, 374–376.
- Strauss, J.H., Strauss, E.G., 1994. The alphaviruses: gene expression, replication, and evolution. *Microbiol. Rev.* 58, 491–562.
- Strauss, E.G., De Groot, R.J., Levinson, R., Strauss, J.H., 1992. Identification of the active site residues in the nsP2 proteinase of Sindbis virus. *Virology* 191, 932–940.

- Strauss, E.G., Rice, C.M., Strauss, J., 1984. Complete nucleotide sequence of the genomic RNA of Sindbis virus. *Virology* 133, 92–110.
- Strelbel, K., Beck, F., 1986. A second protease of foot-and-mouth disease virus. *J. Virol.* 58, 893–899.
- Sun, J., Cao, N., Zhang, X.M., Yang, Y.S., Zhang, Y.B., Wang, X.M., Zhu, H.L., 2011. Oxadiazole derivatives containing 1,4-benzodioxan as potential immunosuppressive agents against RAW264.7 cells. *Bioorg. Med. Chem.* 19, 4895–4902.
- Tay, M.Y., Saw, W.G., Zhao, Y., Chan, K.W., Singh, D., Chong, Y., Forwood, J.K., Ooi, E.E., Gruber, G., Lescar, J., Luo, D., Vasudevan, S.G., 2015. The C-terminal 50 amino acid residues of Dengue NS3 protein are important for NS3-7NS5 interaction and viral replication. *J. Biol. Chem.* 290, 2379–2394.
- Tie, Y., Boross, P.I., Wang, Y.F., Gaddis, L., Liu, F., Chen, X., Tozser, J., Harrison, R.W., Weber, I.T., 2005. Molecular basis for substrate recognition and drug resistance from 1.1 to 1.6 angstroms resolution crystal structures of HIV-1 protease mutants with substrate analogs. *FEBS J.* 272, 5265–5277.
- Ventoso, I., Blanco, R., Perales, C., Carrasco, L., 2001. HIV-1 protease cleaves eukaryotic initiation factor 4G and inhibits cap-dependent translation. *Proc. Natl. Acad. Sci. USA* 98, 12966–12971.
- Villordo, S.M., Gamarnik, A.V., 2008. Genome cyclization as strategy for *Flavivirus* RNA replication. *Virus Res.* 139, 230–239.
- Wang, C.C., Huang, Z.S., Chiang, P.L., Chen, C.T., Wu, H.N., 2009. Analysis of the nucleoside triphosphatase, RNA triphosphatase, and unwinding activities of the helicase domain of dengue virus NS3 protein. *FEBS Lett.* 583 (4), 691–696.
- Wu, E., Nemerow, G.R., 2004. Virus yoga the role of flexibility in virus host cell recognition. *Trends Microbiol.* 12, 162–168.
- Wu, C.F., Wang, S.H., Sun, C.M., Hu, S.T., Syu, W.J., 2003. Activation of dengue protease auto-cleavage at the NS2B–NS3 junction by recombinant NS3 and GSTNS2B fusion proteins. *J. Virol. Methods* 114, 45–54.
- Xu, T., Sampath, A., Chao, A., Wen, D., Nanao, M., Chene, P., Vasudevan, S.G., Lescar, J., 2005. Structure of the Dengue virus helicase/nucleoside triphosphatase catalytic domain at a resolution of 2.4 Å. *J. Virol.* 79, 10278–10288.
- Yap, T.L., Xu, T., Chen, Y.L., Malet, H., Eglhoff, M.-P., Canard, B., Vasudevan, S.G., Lescar, J., 2007. The crystal structure of the Dengue virus RNA-dependent RNA polymerase at 1.85 Å resolution. *J. Virol.* 81, 4753–4765.
- Yu, I.M., Zhang, W., Holdaway, H.A., Li, L., Kostyuchenko, V.A., Chipman, P.R., Kuhn, R.J., Rossmann, M.G., Chen, J., 2008. Structure of the immature dengue virus at low pH primes proteolytic maturation. *Science* 319, 1834–1837.
- Yusof, R., Clum, S., Wetzel, M., Murthy, H.M., Padmanabhan, R., 2000. Purified NS2B/NS3 serine protease of dengue virus type 2 exhibits cofactor NS2B dependence for cleavage of substrates with dibasic amino acids in vitro. *J. Biol. Chem.* 275, 9963–9969.
- Zamora, M., Marissen, W.E., Lloyd, R.E., 2002. Multiple eF4GI specific protease activities present in uninfected and polio-virus infected cells. *J. Virol.* 76, 165–177.
- Zhang, L., Mohan, P.M., Padmanabhan, R., 1992. Processing and localization of Dengue virus type 2 polyprotein precursor NS3–NS4A–NS4B–NS5. *J. Virol.* 66, 7549–7554.

FURTHER READING

- Ralston, R., Thudium, K., Berger, K., Kuo, C., Gervase, B., Hall, J., Selby, M., Kuo, G., Houghton, M., Choo, Q.L., 1993. Characterization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia viruses. *J. Virol.* 67, 6753–6761.