World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(RESEARCH ARTICLE)

Check for updates

IDENTIFICATION OF A STRUCTURALLY AND EVOLUTIONARILY CONSERVED 3'→5' PROOFREADING EXONUCLEASE DOMAIN IN THE LARGE L PROTEIN OF HANTAVIRUSES- EMERGING HUMAN PATHOGENS

Peramachi Palanivelu *

Department of Molecular Microbiology, School of Biotechnology, Madurai Kamaraj University, Madurai – 625 021, India.
* Corresponding Author

World Journal of Advanced Research and Reviews, 2026, 32(01), 969–979

Publication history: Received on 06 September 2026; revised on 12 October 2026; accepted on 14 October 2026

Article DOI: https://doi.org/10.30574/wjarr.2026.32.1.2621

Copyright © 2026 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution License 4.0

ABSTRACT

The proofreading 3′→5′ exonuclease not only represents a crucial class of enzymes responsible for maintaining genome integrity across diverse groups of organisms, but also plays a critical role in ensuring replication fidelity. In the *non-segmented*, *positive-strand* bat-borne RNA viruses like Severe Acute Respiratory Syndrome Coronaviruses (SARS-CoVs and MERS-CoV), it is found along with the methyltransferase (MTase) gene of the non-structural protein 14 (NSP14). In contrast, in the *non-segmented*, but *negative-strand* bat-borne RNA viruses such as Marburg, Ebola and Nipah viruses, it has been identified together with a novel type of capping enzyme, viz. guanosine diphosphate polyribonucleotidyltransferase (PRNTase). Hantaviruses, one of the fatal and emerging human pathogens in recent times, belong to *segmented*, *negative-strand* RNA viruses and are primarily rodent-borne. Currently, no universally approved vaccine or antiviral therapy exists for all hantavirus infections. The presence of a proofreading (PR) exonuclease domain has not been reported in hantaviruses until now. Interestingly, a putative PR exonuclease domain has been identified for the first time in these viruses within the large L protein, preceding the RNA-dependent RNA polymerase (RdRp). The newly identified domain belongs to the DEDD superfamily, but exhibits a distinct structural deviation, notably the replacement of the first catalytic triad -**D**xE- in the Old World hantaviruses with an -**N**xE- motif in most of the New World hantaviruses, and the use of an invariant tyrosine instead of histidine at the proton acceptor site. This feature closely resembles the DEDD superfamily of PR exonucleases reported in bacterial DNA polymerases I and II and in the DNA-dependent DNA polymerases of DNA viruses. To the best of my knowledge, this study provides the first evidence for the presence of a typical DEDD superfamily of PR exonuclease domain in hantaviruses, and thus, expanding our understanding of the structure and catalytic mechanism of 3′→5′ PR exonucleases in *segmented*, *negative-strand* RNA viruses and further presenting a novel structural target for the development of universally effective antiviral therapies against fatal hantavirus infections.

Keywords: Hantaviruses, Old World hantaviruses, New World hantaviruses, 3'→5' proofreading exonucleases, DEDD superfamily of proofreading exonucleases.

1 INTRODUCTION

Viruses are ubiquitous in nature, infecting and proliferating in organisms across all kingdoms of life. They have evolved and adapted their lifecycles to the organisms they infect and, following successful infection, viruses rely predominantly on their host's cellular machinery for their survival and proliferation. Based on the nature of their genome, viruses are broadly classified into RNA and DNA viruses. Both types cause diseases in plants, animals and humans, ranging from endemic to pandemic proportions, resulting in major global healthcare crises and unprecedented economic losses, necessitating a deeper understanding of their pathogenic and replication mechanisms to develop effective therapeutic interventions [1]. However, the majority of the serious human diseases are caused by RNA viruses belonging to both negative- and positive-strand viruses [2]. For example, Measles, Mumps, Human influenza viruses (A, B and C), Hendra, Cedar, Ebola, Marburg, Nipah, Rabies, etc. are caused by the non-segmented, negative-strand RNA viruses whereas SARS-CoVs, MERS-CoV, Polio, Hepatitis C, Dengue, Zika, Yellow fever, Rubella, Encephalomyocarditis, HIVs, etc. belong to non-segmented, but positive-strand RNA viruses. Again, some of the RNA viruses possess *segmented*, single-stranded genomes like human influenza viruses (8 segments) and hantaviruses (3 segments), whereas some possess *segmented*, double-stranded RNA genomes like rotavirus and bluetongue virus.

Hantaviruses are emerging human pathogens that cause severe, often fatal, hemorrhagic fevers and pulmonary syndromes. They have been responsible for several human disease outbreaks in recent years. Furthermore, human-to-human transmission in recent years has established hantaviruses as one of the most dangerous viruses [3, 4]. In fact, hantaviruses are more serious and lethal than the SARS-CoV-2, as some of them have a greater mortality rate (~50%) than the latter (~1% global average). Mostly hantaviruses are not transmitted between humans. However, the Andes orthohantavirus in South America has demonstrated limited person-to-person transmission, particularly among those in close contact. For example, recently, an outbreak of the Andes orthohantavirus on a cruise ship infected 11 people and killed three earlier this year (2026). Hantavirus was first discovered following the Korean War. During the war, haemorrhagic fever with renal syndrome (HFRS) was a common ailment in soldiers stationed near the Hantan River. This illness is now considered as one of the most serious haemorrhagic diseases in the world, like Ebola haemorrhagic fever, due to its high mortality rate and epidemic potential. The first hantavirus was isolated in 1978 in South Korea and was named Hantaan virus. Based on the geographic distribution, hantaviruses are broadly classified into Old World hantaviruses, which typically cause HFRS and are prevalent in Africa, Asia, and Europe, (e.g., Hantaan virus, Seoul virus, Puumala virus, Dobrava-Belgrade virus, Amur virus, etc.) and New World hantaviruses (e.g., Andes virus, Sin Nombre virus, Laguna Negra virus, Black Creek Canal virus, Choclo virus, etc.) which are associated with hantavirus pulmonary syndrome (HPS) in the Americas [5, 6]. The hantaviruses are primarily rodent-borne and are the main source of human infections, although new hantaviruses have recently been reported or isolated from bats, moles and shrews [7, 8]. Spillover of these viruses to humans can result in one of the two serious illnesses. The case fatality rate of HFRS ranges from 1% to 15%, while for HPS, it is 30–60% [9]. Thus, both diseases can be fatal, but with differing mortality rates. Currently, there is neither an approved vaccine nor a specific and effective antiviral treatment available for human use. Importantly, hantaviruses can survive for 10 days at room temperature, 15 days in temperate environments and more than 18 days at 4 °C (39 °F), which facilitates their stable transmission [9, 10]. Therefore, understanding the structure of the hantavirus genome and the functions of the key viral proteins is critical for developing effective antiviral therapeutic agents to control the spread of the virus.

1.1 Classification of hantaviruses

Hantaviruses belong to the genus *Orthohantavirus* and the order *Bunyavirales* within the family of *Hantaviridae*. They are enveloped, negative-sense, single-stranded RNA viruses that cause two potentially lethal zoonotic syndromes, viz. HFRS and HPS as discussed elsewhere. It is estimated that, globally, around 1.5 million people get infected with hantaviruses each year with typical case fatality rates of up to 15% for HFRS and around 40% or higher for HPS [11]. Hantaviruses are further classified into Old World (HFRS-causing) and New World (HPS-causing) viruses [3]. The Old World hantaviruses are primarily found in Europe and Asia and typically cause HFRS with a mortality rate of up to 15%. Examples include, Hantaan virus, Seoul virus, Puumala virus, Dobrava-Belgrade virus, Amur virus, etc. The New World hantaviruses are more dangerous and prevalent in North and South America and are associated with HPS with mortality rates reaching up to 40%. Examples include, Sin Nombre virus, Andes virus, Laguna Negra virus, Black Creek Canal virus. Choclo virus, etc. [12]. Among the hantaviruses, only the New World Andes virus has been demonstrated to have the potential for human-to-human transmission [4].

1.2 3'→5' PR exonucleases and their classifications

PR exonucleases (EC 3.1.11.x) are ubiquitous enzymes that play a crucial role in repairing error(s) during replication of genomes and thus, maintain the integrity of genomes in organisms. They remove misincorporated wrong nucleotides in the 3′→ 5′ direction during DNA/RNA synthesis. Based on their conserved active site amino acids, these enzymes are broadly classified into two superfamilies, viz. the DEDD and the Polymerase and Histidinol Phosphatase (PHP) superfamilies [13, 14]. The DEDD superfamily is the most common type of PR enzyme and is reported from both DNA

and RNA polymerases from a large number of organisms from viruses to humans. The name derives from the four highly conserved acidic amino acid residues, viz. DEDD, in its active site for metal-binding and catalysis. The PHP superfamily, characterized by the active site amino acid structure, **H**x**H**-E-**H**-**D**, uses His residues for metal-binding and catalysis. Bacterial/Archaeal DNA polymerases X, bacterial DNA polymerases III (bacterial replicases) with a co-editing, built-in exonuclease [14, 15] and YcdX phosphoesterases [16] belong to PHP superfamily. Both superfamilies are essentially Zn-dependent enzymes and initiate the catalysis from a metal-bound water molecule [17]. It is interesting to note that certain viruses like the human influenza viruses use both types of PR exonucleases [18]. Furthermore, the DEDD superfamily is further classified into two subfamilies, viz. DEDD(Y) and DEDD(H), depending on whether they employ an invariant Y or a H as the proton acceptor during catalysis [Table 1].

Most of the RNA and DNA viruses possess a typical DEDD superfamily of 3′→5′ PR exoribonuclease in their genome and are well-characterized. For example, in SARS-CoVs and MERS-CoV, it is characterized in the ExoN of the non-structural protein 14 along with the methyltransferase domain [17]. But in the other bat-borne viruses such as Marburg, Ebola and Nipah, a DEDD superfamily of PR exonuclease domain was identified in the PRNTase (the capping enzyme) domain [19]. The DEDD superfamily of PR exonuclease has also been reported in DNA viruses like smallpox, monkeypox viruses, etc. [Table 1] suggesting the same PR mechanism is used by both types of viruses. In contrast to their ubiquitous presence in RNA and DNA viruses, in hantaviruses it was reported that there was no demonstrated, dedicated proofreading exonuclease [20, 21]. However, in the present communication, a putative DEDD superfamily of PR exonuclease domain, preceding the RdRp domain, has been identified featuring an invariant Y as the proton acceptor.

1.3 Importance of 3'→5' PR exonuclease as a potential drug target for controlling RNA viruses

Maintenance of genome stability is essential for all living organisms and it mainly relies on the error-free replication of genomes by the respective DNA and RNA replicases. The genome stability is equally important for viruses also. This is because too many uncorrected mutations during replications could lead to an "error catastrophe," where the viruses could literally mutate itself to death. To overcome this problem, the RNA and DNA polymerases/replicases exhibit strong discrimination between nucleoside triphosphates (NTPs) and deoxynucleoside triphosphates (dNTPs) and rarely insert a wrong nucleotide during replication of genomes and hence, their error rates are very, very low and is usually in the order of $\sim 10^{-4}$ to 10^{-6} and $\sim 10^{-6}$-10^{-9}, respectively [22]. To overcome even this minimal error rates during replications, these crucial polymerases of life are invariably associated with PR mechanisms that correct any insertion error(s) during replication. When an incorrect nucleotide gets incorporated during replication, the polymerases stall/pause, which in-turn activates the associated PR function, which promptly excises the misincorporated nucleotide. The DEDD superfamily of PR exonucleases operates in the 3’→5’ direction and excise any incorrectly incorporated nucleotide from the 3’-growing end, and thus, helping the polymerases to perform error-free replication of genomes. Following the excision of the wrong nucleotide, the correct nucleotide is inserted and the polymerases resume the replication. This critical PR step in living organisms and viruses ensures that the original DNA/RNA template is copied accurately and passed on to the next generation. For example, with the PR 3'→ 5' exonuclease activity, the error rate dramatically improves to an accuracy of about 10^{-7} to 10^{-8} (one mistake per 10 million to 100 million bases). And further, with cellular mismatch repair, the final error rate drops as low as 10^{-9} to 10^{-11} per base pair (i.e., about one mistake per 1 billion bases) [23]. Thus, the PR functions play a crucial role in maintaining the integrity of genomes in all organisms including the viruses. It is interesting to note that a SARS-CoV-2 PR exonuclease knockout mutant was unable to replicate, suggesting possibilities for development of effective antivirals directing the PR exonucleases to contain the spread of RNA viruses [17].

1.4 Genome and L protein organizations in hantaviruses

Arenaviridae, *Bunyaviridae*, and *Orthomyxoviridae* are the principal families of *segmented*, negative-strand RNA viruses, and each of them include serious human pathogens. *Arenaviridae* have two genome segments, whereas *Bunyaviridae* and *Orthomyxoviridae* have three and six to eight genome segments, respectively. Arenaviruses and bunyaviruses replicate in the cytoplasm using a single subunit RdRp of the L protein, whereas the orthomyxoviruses replicate in the nucleus- an unusual trait of an RNA virus- using a heterotrimeric polymerase complex composed of PA, PB1 and PB2 subunits [18]. Hantaviruses are *segmented*, negative-strand, enveloped viruses which belong to the order *Bunyavirales,* Family, *Hantaviridae,* and Genus, *Orthohantavirus.* Unlike other human pathogenic viruses, they are primarily carried by rodents. The virions are generally spherical to pleomorphic, measuring approximately 80-160 nm in diameter, although the size may vary among species.

Thus, the segmented, tripartite RNA genome of hantaviruses consists of a small (S), a medium (M), and a large (L) segment and encodes only four structural proteins. The S segment encodes the nucleocapsid (N) protein, which encapsidates viral RNA genome segments to form ribonucleoprotein (RNP) complexes and plays key roles in viral replication and immune modulation. The M segment encodes a glycoprotein precursor that is post-translationally cleaved into two envelope glycoproteins, Gn and Gc, which are essential for viral attachment and entry into host cells. The multidomain L segment encodes a viral RdRp, responsible for both transcription and replication and a cap-snatching endonuclease, responsible for snatching mRNA caps from the host mRNAs for priming the viral mRNAs. The

genomic RNA segments are encapsidated by the nucleocapsid protein P, to form helical RNP complexes, which are enclosed within a host-derived lipid envelope studded with Gn and Gc glycoprotein spikes. [24]. The viral envelope is derived from host cell membranes and contains the above embedded glycoprotein spikes that facilitate interaction with host cell receptors. The tripartite genome segments form into circles via non-covalent interactions of the ends of the genome with an attached L protein to initiate transcription upon entry into host cells. The 3'- and 5'- terminal nucleotides are highly conserved and complementary in all three genome segments that form a conserved panhandle structure acting as the viral promoter in each of the genome segment. In addition to the coding region, each segment possesses short untranslated regions (UTRs) at the 5′ and 3′ termini. The 3' and 5' UTRs are highly conserved and regulate replication and transcription of the genomes [25].

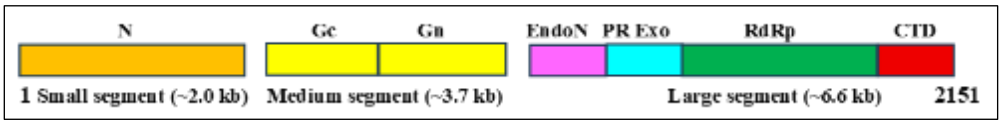


Figure 1: Tripartite genome organization and the tentative domain architecture of the L protein in hantaviruses

- The small segment encodes the nucleocapsid protein (N = ~ 50 kDa) which binds viral RNAs to form RNP complexes that contain a single RNA segment, many copies of N which bind and protect the RNA, and a single copy of L polymerase.
- The two surface glycoproteins, Gn and Gc, of the medium segment are the result of post-translational cleavage of the glycoprotein precursor (GPC = ~126 kDa) by the host protease and Gn and Gc represent the N- and C-terminal portions of GPC, respectively. They are embedded in the lipid bilayer and form characteristic spikes. They are involved in host cell attachment and entry.
- The large segment encodes the L protein, an approximately 250 kDa protein harbouring the RdRp. The L protein is a multifunctional protein that harbours an N-terminal endonuclease (EndoN), an RdRp, and a presumed cap binding domain (CBD) at the C-terminal (CTD) and organised as: Endonuclease 1-372; PR Exo 373-676; RdRp 677-1601; CTD 1602-2151.
- Importantly, each of the genome segments carries the L protein polymerase to initiate the transcription upon entry and uncoating. Thus, the L protein is associated with each genome segment as a part of the RNP complex. Nucleotides 1–10 of the 3′ and 5′ of the genome ends are complementary. The 5′ binds as a hook in a specific site located on the surface of the polymerase core region, whereas the 3′ binds either in the active site or in a distinct 3′ end of the secondary binding site. The complementary nucleotides around positions 11 to 20 form a distal duplex that protrudes from the polymerase core [26].

1.5 Lifecycle of hantaviruses and release of mature virions

Integrins, particularly the β3 integrins, are important cell surface receptors involved in hantaviral entry. The Old World hantaviruses use clathrin-dependent endocytosis, whereas the New World hantaviruses use clathrin-independent endocytosis. After entry into the host cells, the viral envelope fuses with the endosome or lysosome and releases the viral RNP complexes into the host cell cytoplasm. Complete hantaviral transcription and replication occur in the cytoplasm of the infected host cell. The associated L protein polymerase of each genome segment initiates transcription and replication [27]. As the hantavirus RdRp does not have an independent capping enzyme like most of the other RNA viruses, it snatches 5′-caps from host mRNAs and uses them for priming its own mRNA synthesis. Followed by transcription, the viral mRNAs are translated by host ribosomes and produce viral proteins. The glycoproteins undergo maturation in the Golgi apparatus [25]. Genome replication begins when sufficient nucleocapsid protein P, is available to encapsidate newly synthesized complementary positive-sense RNA intermediates (antigenomes), which serve as templates for the synthesis of negative-sense genomic RNA segments. Thus, the genome replication is a two-step process. In the first step, the viral RdRp transcribes the negative, single-stranded RNA segments into complementary, positive-sense strands (cRNAs), which are used as a template to produce new genomic viral RNA copies, in the second step. The Old World hantaviruses assemble in the Golgi apparatus and acquire their envelope from it, before being transported to the cell membrane to leave the cell via exocytosis, while the New World hantaviruses assemble near the cell membrane and obtain their envelope from it as they exit the cell by budding from its surface.

2 MATERIALS AND METHODS

Full-length L protein sequences from the Old World and New World hantaviruses were obtained from the PubMed and Swiss-Prot databases. The advanced version of Clustal Omega was used for MSA analysis [28, 29]. The conserved motifs and sequence similarities identified by the bioinformatics analysis were integrated with the data already available from biochemical, SDM, cryo-EM, and X-ray crystallographic analyses on the PR exonucleases from different organisms to identify the potential PR exonuclease active site amino acids in the L protein of the hantaviruses.

3 RESULTS AND DISCUSSION

3.1 Identification of PR exonuclease domain in hantaviruses

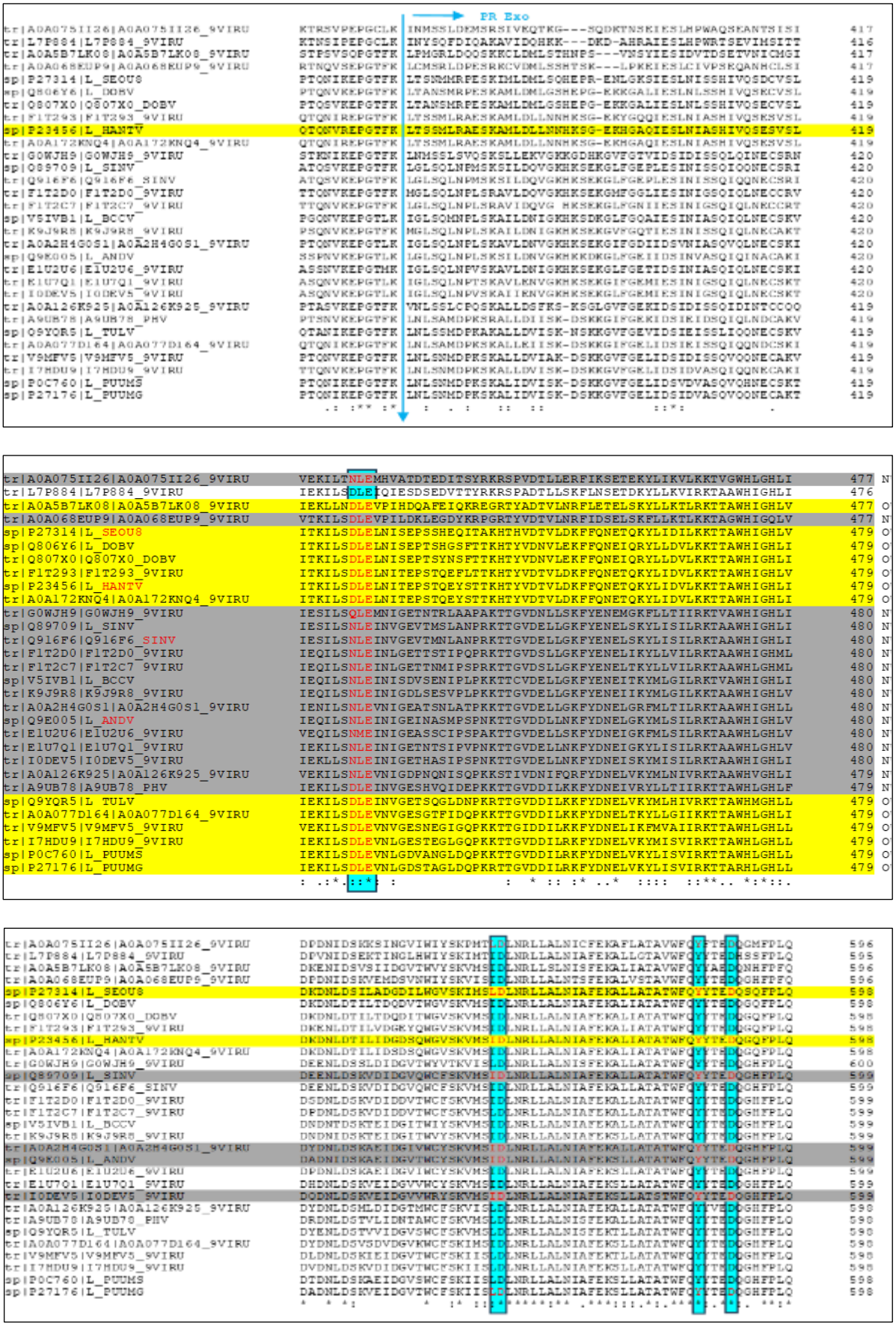
Figure 2 shows the tentative PR exonuclease domain harbouring the highly conserved active site amino acids in the L protein of Old World and New World hantaviruses, (Only the region relevant to the discussion is shown here, and the PR exonuclease and RdRp domains are indicated by arrows). It has been previously reported that hantavirus RdRp lacks PR activity [20, 21]. However, a complete set of highly conserved putative PR exonuclease active site amino acids (highlighted in light blue) are identified in these viruses in the L protein, preceding the RdRp domain. Furthermore, it is clear from the conserved motifs and amino acids that the PR exonuclease belongs to the DEDD superfamily of PR exonuclease which are the most common PR enzyme among DNA and RNA polymerases/ replicases (Table 1). The DEDD superfamily uses either an invariant Y or a H as the proton acceptor to initiate cleavage of the phosphodiester bond of the incorrectly incorporated nucleotide during the polymerization process. In hantaviruses, a completely conserved Y within the PR exonuclease domain could serve as the proton acceptor and hence, belongs to the DEDD(Y) subfamily. During the exonucleolytic cleavage, the invariant Y could accept the proton from a metal-bound water molecule to initiate the PR activity as illustrated in Fig. 3. The proposed mechanism of action for the DEDD superfamily of PR exonucleases has been previously explained by Palanivelu [30].

It is interesting to note that the first conserved triad -**DxE**- of the DEDD superfamily of exonucleases differs markedly between the Old World and New World hantaviruses. For instance, the Old World hantaviruses use the typical -**DxE**- triad (highlighted in yellow) whereas the New World hantaviruses use an **NxE** (and rarely a **Q** in the place of **N**) (highlighted in grey) (Fig. 2). The preferential use of an N instead of a D in the New World hantaviruses is interesting and also intriguing as they cause HPS, a severe and potentially a fatal disease affecting the lungs. Notably, the Andes virus, which belongs to the New World hantaviruses, is the only hantavirus capable of sustained human-to-human transmission, use -**N**xE- triad. The second important difference that is observed among the hantaviruses is that the PR exonuclease uses an invariant **Y** as the proton acceptor, which differs from the other bat-borne human pathogenic viruses like SARS-CoVs, MERS-CoV, human influenza viruses, RSV, Marburg, Ebola and Nipah viruses where they all use an invariant **H** as the proton acceptor (Table 1). Therefore, this DEDD superfamily of PR exonuclease with an invariant Y as the proton acceptor is in close agreement with the DEDD superfamily of PR exonuclease reported in DNA-dependent DNA polymerases (DdDps) rather than in the RdRps (Table 1), suggesting different evolutionary pathways. (However, only the Hepatitis A virus (a non-segmented, positive-strand RNA virus) RdRp uses an invariant Y) (Table 1). Furthermore, it also differs from the other *segmented*, negative-strand RNA viruses such as the human influenza viruses, where they use an invariant H as the proton acceptor and additionally possess a PHP type superfamily of PR exonuclease also [18].

The two characteristic motifs, viz. –SLYPS- and –YGDTDS- are invariably found in the eukaryotic replicative polymerases (pols) α and δ, from yeasts, higher fungi, plants, and animals. These motifs are placed before and after the proposed polymerase active site, and are implicated in dNTP- and metal-binding, respectively. Interestingly, they are also reported from a large number of other DdDps [30]. For example, in addition to their presence in the eukaryotic replicative DNA pols α and δ, they are also found in the *E. coli* DNA pol II and in many viral DNA pols like poxviruses (Smallpox, Monkeypox, etc.), Vaccinia, Epstein-Barr virus, Human Cytomegalo virus, Human Herpes Simplex viruses 1 and 2, Adenovirus 2, *E. coli* phage RB69, *Bacillus subtilis* phage Φ29, suggesting their possible origin from viruses and prokaryotes. (In contrast, these motifs are absent in the eukaryotic replicative leading-strand synthesis enzyme, pol ε, suggesting it is closer to the prokaryotic DNA pols I, and DNA pols III (replicases), where they also do not possess these two characteristic motifs. These findings further suggest two distinct origins of the eukaryotic replicative pols [31]. It is interesting to note a functionally similar conserved motif -GVYPS- is observed (highlighted in yellow in the RdRp domain of the hantaviruses) (Figure 2) in the Old and New World hantaviruses, further suggesting their origin from the B-family of DdDps, a group of highly accurate enzymes with built-in 3'→5'PR activity primarily responsible for nuclear DNA replication across all domains of life (eukaryotes, bacteria, and archaea) as well as many DNA viruses including poxviral DdDps (-SLYPN-) [32].

Interestingly, the C-terminal ends of the L protein are highly conserved among hantaviruses, featuring an invariant aromatic amino acid diad, **-FY-** (highlighted), at the extreme C- termini. This remarkable conservation suggests that the C-terminal region may have a functional role, possibly involving interactions with the viral genome termini (Fig. 2). Interestingly, a similar organization is observed in human influenza viruses, where each RNA genome segment is associated with multiple nucleocapsid protein molecules and a polymerase heterotrimer composed of PB2, PB1, and PA. The viral polymerase binds the conserved terminal sequences of each viral RNA genome segment.

CLUSTAL O (1.2.4) MSA of Old World (highlighted in yellow) and New World (highlighted in grey) hantaviruses from various regions of the world.



NB: O and N, Old World and New World hantaviruses, respectively.

A0A075II26_9VIRU RNA-directed RNA polymerase L, Thottapalayam virus (Vellore, Tamil Nadu)
L7P884_9VIRU RNA-directed RNA polymerase L, Kilimanjaro virus (Tanzania, Africa)
A0A5B7LK08_9VIRU RNA-directed RNA polymerase L, Lena virus (Russia)
A0A068EUP9_9VIRU RNA-directed RNA polymerase L, Camp Ripley virus (USA)
P27314|L_SEOU8 RNA-directed RNA polymerase L, Seoul virus (strain 80-39) (S. Korea)
Q806Y6|L_DOBV RNA-directed RNA polymerase L, Dobrava-Belgrade orthohantavirus (Germany)
Q807X0|Q807X0_DOBV RNA-directed RNA polymerase L, Kurkino virus (Germany)
F1T293_9VIRU RNA-directed RNA polymerase L, Amur virus (East Asia)
P23456|L_HANTV RNA-directed RNA polymerase L, Hantaan virus (strain 76-118) (East Asia)
A0A172KNQ4_9VIRU RNA-directed RNA polymerase L, Orthohantavirus hantanense (S. Korea & China)
G0WJH9_9VIRU RNA-directed RNA polymerase L, Rockport virus (Texas, USA)
Q89709|L_SINV RNA-directed RNA polymerase L, Sin Nombre orthohantavirus (North America)
Q916F6_SINV RNA-directed RNA polymerase L Convict Creek 107 virus (North America)
F1T2D0_9VIRU RNA-directed RNA polymerase L, Huitzilac virus (Mexico)
F1T2C7_9VIRU RNA-directed RNA polymerase L, Carrizal virus (Mexico)
V5IVB1|L_BCCV RNA-directed RNA polymerase L, Black Creek Canal orthohantavirus (USA)
K9J9R8_9VIRU RNA-directed RNA polymerase L, Catacamas virus (Honduras)
A0A2H4G0S1_9VIRU RNA-directed RNA polymerase L, Choclo virus (S. America)
Q9E005|L_ANDV RNA-directed RNA polymerase L Andes orthohantavirus (S. America)
E1U2U6_9VIRU RNA-directed RNA polymerase L Maporal virus (S. America)
E1U7Q1_9VIRU RNA-directed RNA polymerase L, Rio Mamore hantaviruse (South America)
I0DEV5_9VIRU RNA-directed RNA polymerase L, Maripa virus (French Guiana)
A0A126K925_9VIRU RNA-directed RNA polymerase L, Fugong virus (China)
A9UB78_PHV RNA-directed RNA polymerase L, Prospect Hill virus (USA)
Q9YQR5|L_TULV RNA-directed RNA polymerase L, Tula orthohantavirus (Europe and Asia)
A0A077D164_9VIRU RNA-directed RNA polymerase L, Yuanjiang virus (China)
V9MFV5|V9MFV5_9VIRU RNA-directed RNA polymerase L, Muju virus (Korea)
I7HDU9_9VIRU RNA-directed RNA polymerase L, Hokkaido virus (Japan)
P0C760|L_PUUMS RNA-directed RNA polymerase L, Puumala virus (Europe & Russia)
P27176|L_PUUMG RNA-directed RNA polymerase L, Puumala virus (Europe & Russia)

Figure 2: MSA of Old World (highlighted in yellow) and New World hantaviruses (highlighted in grey) from various regions around the world.

3.2 PR exonuclease active site structures from the viral, prokaryotic, and eukaryotic DNA and RNA polymerases

Table 1 presents the PR exonuclease active site structures from viral, prokaryotic, and eukaryotic DNA and RNA polymerases. It is evident from the Table that not only a chemical basis is involved in the design of PR exonuclease active sites, but are also highly conserved and adapted in all organisms, from viruses to humans throughout the extensive evolutionary stages. Members of the DEDD superfamily utilize either an Y or a H as the proton acceptor and accepts the proton from a metal-bound water molecule to initiate the phosphodiester bond cleavage reaction.

Table 1: The DEDD superfamily of PR exonuclease active site structures (confirmed and putative) from viral, prokaryotic, and eukaryotic DNA/RNA polymerases.

Organism	PR exonuclease active site
1. Prokaryotic DEDD family PR Exonuclease Active Site Structure	
E. coli DNA Pol I (MFE)	-D^{355}TE^{357}-----YD^{424}-----Y^{497}----D^{501}-
E. coli DNA Pol II (MFE)	-D^{156}IE^{158}------FD^{229}-----Y^{331}----D^{335}-
E. coli DNA Pol III (Replicase, ε-subunit) (MSE)	-D^{12}TE^{14}-------FD^{102}-----H^{162}----D^{167}-
Tth DNA pol III (Replicase, ε-subunit) (MSE)	-D^{77}LE---------FD^{161}-----H^{214}----D^{219}-
2. Mitochondrial DNA Polymerase γ from Yeasts, Plants and Animals	
Yeasts (*S. cerevisiae*)	-D^{171}VE^{173}------YD^{230}-------Y^{343}---D^{347}-
Plants (*A. thaliana* DNA pol IA, Mito/Chplast)	-D^{294}TE^{296}------FD^{308}-------Y^{470}----D^{474}-
Plants (*A. thaliana* DNA pol IB, Mito/Chplast)	-D^{272}TE^{274}------FD^{346}-------Y^{448}----D^{452}-
Animals (*Homo sapiens*)	-D^{198}VE^{200}------FD^{274}-------Y^{395}----D^{399}-
3. Eukaryotic Nuclear Replicative DNA Polymerases δ and ε from Yeasts, Plants and Animals	
Yeasts (*S. cerevisiae* δ DNA pol)	-DIE^{323}-------FD^{407}-------Y^{516}----D^{520}-
Yeasts (*S. cerevisiae* ε DNA pol	-DIE^{292}-------FD^{383}-------Y^{473}----D^{477}-
Plants (*A. thaliana* δ DNA pol)	-DIE^{314}-------FD^{398}-------Y^{507}----D^{511}-
Plants (*A. thaliana* ε DNApol	-DIE^{249}-------FD^{340}-------Y^{431}----D^{435}-
Animals (*H. sapiens* δ DNA pol)	-D^{316}IE^{318}-- -FD^{402}-------Y^{511}----D^{515}-
Animals (*H. sapiens* ε DNA pol	-D^{275}IE^{277}----F^{367}D^{368}----Y^{458}S^{459}---D^{462}-
4. Nuclear-encoded Organellar Eukaryotic RNA Polymerases from Yeasts, Plants and Animals	
Plants (*A. thaliana*, NE-RNA pol, Chplast)	-D^{548}VE----------LD^{572}F--------NH^{586}L---D^{590}LC-
Plants (*A. thaliana*, NE-RNA pol, Mito)	-DIE^{533}----------VD^{555}F--------NH^{569}L---D^{573}LC-
Animals (*H. sapiens*, NE-RNA pol, Mito)	-D^{540}AE----------MD^{801}F-------NH^{815}L---D^{819}VA
5. DEDD-superfamily of PR Exonuclease Active Site Structures in Viral Polymerases (DdDp/RdRp)	
T4 Phage (DdDp)	-D^{112}IE^{114}------FD^{219}------Y^{320}---D^{324}-
(dsDNA) Smallpox Virus (DdRp)	-D^{166}IE---------FD^{268}------Y^{457}---D^{461}-
(dsDNA) Vaccinia Virus (DdRp)	-D^{166}IE---------FD^{268}------Y^{458}---D^{462}-
(dsDNA) Mpox Virus (DdRp)	-D^{166}IE---------FD^{268}------Y^{458}---D^{462}-
(dsDNA) Human Herpes Simplex Virus (DdDp)	-D^{368}IE---------FD^{471}------Y^{577}---D^{581}-
(ns, + strand RNA) SARS-CoV-1 (#NSP14-Exon)*	-D^{90}VE^{92}-------FE^{191}--------H^{268}----D^{273}-
(ns, + strand RNA) MERS-CoV (#NSP14-Exon)*	-D^{90}VE^{92}-------FE^{191}--------H^{268}----D^{273}-
(ns, + strand RNA) SARS-CoV-2 (#NSP14-Exon)*	-D^{90}VE^{92}-------FE^{191}--------H^{268}----D^{273}-
(ns, + strand RNA) Hepatitis A Virus (RdRp)	-D^{233}LE--------FD^{363}--------Y^{483}----D^{487}-
(s, - strand RNA) Human Influenza Virus A: H1N1 (PA subunit)	-D^{347}IE---------LD^{425}-------H^{510}---D^{514}-
(s, - strand RNA) Human Influenza Virus B (PA subunit)	-D^{194}IE---------LD^{420}-------H^{506}---D^{510}-
(s, - strand RNA) Human Influenza Virus C (PA subunit)	-D^{84}LE---------ID^{288}--------H^{494}---D^{498}-
(ns, - strand RNA) Human Respiratory Syncytial Virus (RdRp)	-D^{664}LE--------LD^{794}-------H^{829}---D^{833}-
(ns, - strand RNA) Marburg Virus (PRNTase)	-D^{1013}LE-------LD^{1096}------H^{1294}R---D^{1298}-
(ns, - strand RNA) Ebola Virus (PRNTase)	-D^{1010}FE-------VD^{1108}------H^{1269}R---D^{1273}-
(ns, - strand RNA) Nipah Virus (PRNTase)	-D^{1101}LE-------LD^{1258}------H^{1347}R---D^{1351}-
(s, - strand RNA) Old World hantaviruses^ (Hantaan)	-D^{426}LE--------ID^{563}--------Y^{587}Y---D^{591}QG-
(s, - strand RNA) Old World hantaviruses^ (Seoul)	-D^{426}LE--------LD^{563}-------Y^{587}Y---D^{591}QS-
(s, - strand RNA) New World hantaviruses^ (Andes)	-N^{427}LE--------ID^{564}--------Y^{588}Y---D^{592}QG-
(s, - strand RNA) New World hantaviruses^ (Sin Nombre)	-N^{427}LE--------ID^{564}--------Y^{588}Y---D^{592}QG-

Adapted from Palanivelu [19].

- MFE, Multifunctional enzyme; MSU, Multisubunit enzyme; ns, non-segmented; s, segmented
- Eukaryotic replicative α pol (DNA primase) lack the PR exonuclease active site.
- Tth*, Thermus thermophilus*; NE, Nuclear encoded; Chplast, Chloroplast; Mito, Mitochondria
- The confirmed PR exonuclease active site amino acids (by SDM and Crystallography) are highlighted in dark blue/light blue, respectively.
- *All possess the same PR exonuclease domain in the NSP14 and their active site amino acids are completely conserved in Bat-RaTG13, Pangolin, Civet, SARS-CoV-1, MERS-CoV, SARS-CoV2 and HCoV-NL63.
- #NSP14 is a multifunctional protein with two distinct activities, an N-terminal PR 3'→5' exoribonuclease (ExoN) and a C-terminal N7-methyltransferase (N7-MTase), both are crucial for life cycle of coronaviruses, suggesting NSP14 Exon as one of the prominent targets for the development of antiviral drugs.
- ^ Present work

3.3 Proposed PR exonuclease active site structures in the Old World and New World hantaviruses

As discussed elsewhere, the DEDD superfamily of 3'→5' PR exonucleolytic activity is essential for the faithful replication of genomes and the maintenance of their integrity. It exists as an intrinsic domain in bacterial DNA polymerases I and II and in eukaryotic replicases (α and δ) or as an auxiliary subunit (ε subunit) in bacterial DNA polymerases III (replicases), a MSU enzyme [33]. In hantaviruses, the 3'→5' PR exonuclease is identified as an intrinsic domain preceding the RdRp of the L protein and thus, exists in the same polypeptide as a part of the MFE system as reported in bacterial DNA polymerases I, II and eukaryotic replicases (α and δ) rather than found in other non-segmented and positive- strand RNA viruses where it acts as an auxiliary enzyme as discussed elsewhere (Table 1). Figures 3A and 3B illustrate the proposed active site amino acids in the 3'→5' PR exonuclease domain of Old World and New World hantaviruses. In general, with the PR 3'→5' exonuclease, the error rate improves to an accuracy to about 10^{-7} to 10^{-8} (one mistake per 10 million to 100 million bases) [23]. Even with such a stringent error-correction mechanism associated with the DNA/RNA polymerases, some replication errors escape the repair mechanism and thus, become permanent mutation(s). These mutation(s) could be advantageous to the viruses (and thus maintained) or deleterious to them (and eliminated). In fact, the advantageous mutations serve as the basis for the evolution of new variants. The proposed two-metal ion mechanism in the active site of the viral 3'→5' PR exonuclease is based on Beese and Steitz's observation, who described an analogous active site architecture in the PR exonuclease domain of *E. coli* DNA polymerase I [34]. A general mechanism of action for the PR exonuclease has already been reported by the author [30]. It is not clear whether the preferential use of **N** instead of **D** in the first catalytic triad in almost all the New World hantaviruses confers any advantage for them in the PR mechanism and infection specificity.

Hantaan virus (OW) -D^{426}LE^{428}-------ID^{563}-------Y^{587}Y---D^{591}QG-

Andes orthohantavirus (NW) -N^{427}LE^{429}-------ID^{564}-------Y^{588}Y---D^{592}QG-

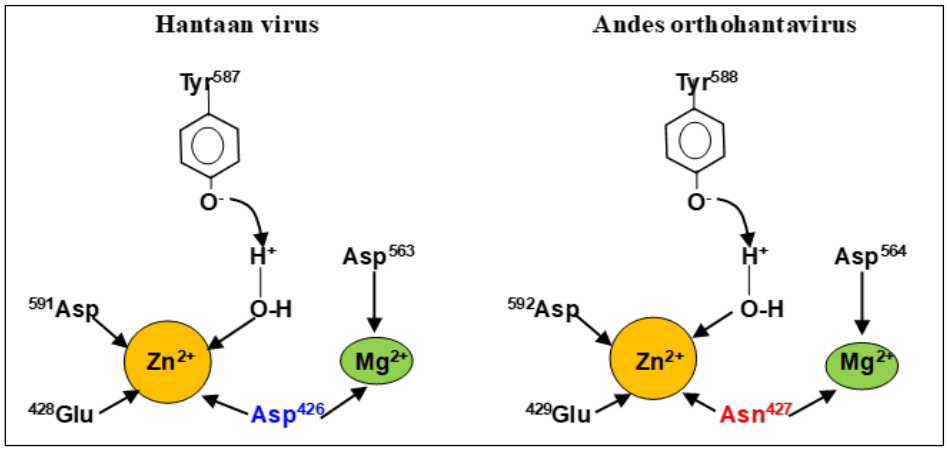


Figure 3: Proposed amino acids at the PR exonuclease active sites in the Old World (e.g., Hantaan virus) and New World hantaviruses (e.g., Andes orthohantavirus).

4 CONCLUSIONS

This study details the discovery of a regular PR mechanism embedded within the replication machinery of hantaviruses to maintain the replication fidelity and integrity of their genomes. The putative PR exonuclease from these viruses belongs to the DEDD superfamily of exonucleases. Notably, it uses an invariant Y as the proton acceptor in the catalytic

site, unlike most of the RNA viral RdRps, which use an invariant H. In this respect, the hantavirus PR exonuclease is structurally closer to the PR exonucleases of high-fidelity DdDps than to most other viral PR exonucleases of RdRps. Although the involvement of each of the DEDD catalytic amino acids has been confirmed experimentally in RNA/DNA polymerases from many different organisms, it still needs to be validated in hantaviruses. The complete conservation of the DEDD superfamily of PR exonuclease catalytic domain underscores the evolutionary significance of this functionally conserved mechanism for maintaining replication fidelity from viruses to humans. Furthermore, this discovery provides a strong rationale for developing antivirals targeting the PR exonuclease domain as a potential strategy to control the spread of these emerging human pathogens.

STATEMENTS ON COMPLIANCE WITH ETHICAL STANDARDS

Acknowledgments

The author wishes to thank Dr. N. Srinivasan, former Professor, Department of Endocrinology, Post Graduate Institute of Basic Medical Sciences, University of Madras, Chennai, for useful suggestions on the manuscript.

Disclosure of conflict of interest

The author declares no conflicts of interest.

REFERENCES

[1] Mohanty P, Panda P, Acharya RK, Pande B, Bhaskar L, Verma HK. Emerging perspectives on RNA virus-mediated infections: from pathogenesis to therapeutic interventions. World J Virol. 2023; 12: 242-255.

[2] Poltronieria P, Sunb B, Mallardoc M. RNA Viruses: RNA Roles in Pathogenesis, Coreplication and Viral Load. Curr Genomics. 2015; 16: 327-335.

[3] Kumar ASA, John J. Hantavirus: A Comprehensive Contemporary Review of Virology, Global Epidemiology, Clinical Variants, Diagnosis, Treatment, and Intensive Care Management. Indian J Crit Care Med. 2026; 30:430–442.

[4] Toledo J, Haby MM, Reveiz L, Sosa Leon L, Angerami R, Aldighieri S. Evidence for Human-to-Human Transmission of Hantavirus: A Systematic Review. J Infect Dis. 2022; 226: 1362–1371. doi:10.1093/infdis/jiab461.

[5] Khan A, Khan AS. Hantaviruses: A tale of two hemispheres. Panminerva Medica. 2003; 45: 43–51.

[6] Schmaljohn, C. Molecular biology of hantaviruses. 1996; p 63-90. In R. M. Elliot (ed.), The Bunyaviridae. Plenum Press, New York, NY.

[7] Weiss S, Witkowski PT, Auste B, Nowak K, Weber N, Fahr J, et al. Hantavirus in bat, Sierra Leone. Emerg Infect Dis. 2012; 18:159–161.

[8] Song JW, Baek LJ, Schmaljohn CS, Yanagihara R. Thottapalayam virus, a prototype shrew-borne hantavirus. Emerg Infect Dis. 2007; 13:980–985.

[9] Chen R, Gong H, Wang X, Sun M, Ji Y, Tan S, Chen J, Shao J, Liao M. Zoonotic Hantaviridae with Global Public Health Significance. Viruses. 2023; 15: 1705. doi:10.3390/v15081705.

[10] Riccò M, Peruzzi S, Ranzieri S, Magnavita N. Occupational Hantavirus Infections in Agricultural and Forestry Workers: A Systematic Review and Metanalysis. Viruses. 2021; 13: 2150. doi:10.3390/v13112150.

[11] Jonsson CB, Figueiredo LT, Vapalahti O. A global perspective on hantavirus ecology, epidemiology, and disease. Clin Microbiol Rev. 2010; 23:412–441.

[12] Avšič-Županc T, Saksida A, Korva M. Hantavirus infections. Clin Microbiol Infect. 2019; 21S, e6–e16.

[13] Zuo Y, Deutscher MP. Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. Nucleic Acids Res. 2001; 29:1017–1026.

[14] Nagpal S, Nair DT. The PHP domain of PolX from *Staphylococcus aureus* aids high fidelity DNA synthesis through the removal of misincorporated deoxyribo-, ribo- and oxidized nucleotides. Nat Sci Rep. 2021; 11:4178.

[15] Stano NM, Chen J, McHenry CS. A co-proofreading Zn^{2+}-dependent exonuclease within a bacterial replicase. Nat Struct Mol Biol. 2006; 13:458–459.

[16] Teplyakov A, Obmolova G, Khil PP, Howard AJ, Camerini-Otero RD, Gilliland GL. Crystal Structure of the *Escherichia coli* YcdX Protein Reveals a Trinuclear Zinc Active Site. Proteins: Struct Funct Genet. 2003; 51:315–318.

[17] Palanivelu P. An overview of the proofreading functions in bacteria and SARS-Coronaviruses. Int J Biochem Res Rev. 2021; 30: 33-62.

[18] Palanivelu P. Identification of DEDD and PHP Superfamilies of Proofreading Exonucleases in the Acidic Protein Subunit PA of RNA Polymerase of Human Influenza Viruses, World J Adv Res Rev. 2022; 16: 804–824.

[19] Palanivelu P. Identification of a Proofreading Domain within the Guanosine Diphosphate Polyribonucleotidyltransferase Domain in the Bat-borne Deadly Human Viruses: Marburg, Ebola and Nipah. World J Adv Res Rev. 2026; 30: 2246-2261.

[20] Cheng E, Mir MA. Signatures of Host mRNA 5' Terminus for Efficient Hantavirus Cap Snatching. J Virol. 2012; 86:10173–10185.

[21] Jonsson CB, Schmaljohn CS. Replication of hantaviruses. Curr Top Microbiol Immunol. 2000; 256:15–32. doi: 10.1007/978-3-642-56753-7_2.

[22] Elena SF, Sanjuan R. Adaptive value of high mutation rates of RNA viruses: Separating causes from consequences. J. Virol. 2005; 79:11555–11558.

[23] Pray L. DNA Replication and Causes of Mutation. Nature Education. 2008; 1:214.

[24] Overbosch FW, Eggink D, Waalewijn B. Hantavirus: Past, present and future. Life 2023; 13:647. doi: 10.2217/fvl.13.113.

[25] Jacob AT, Ziegler BM, Farha SM, Vivian LR, Zilinski CA, Armstrong AR, Burdette AJ, Beachboard DC, Stobart CC. Sin Nombre Virus and the Emergence of Other Hantaviruses: A Review of the Biology, Ecology, and Disease of a Zoonotic Pathogen. Biology (Basel). 2023; 12: 1143. doi:10.3390/biology12111413.

[26] Trouilleton QD, Barata-García S, Arragain B, Reguera J, Malet H. Structures of active Hantaan virus polymerase uncover the mechanisms of Hantaviridae genome replication. Nat Commun. 2023; 14:2954. doi.org/10.1038/s41467-023-38555-w.

[27] Meier K, Thorkelsson SR, Quemin ER, Rosenthal M. Hantavirus Replication Cycle - An Updated Structural Virology Perspective. Viruses. 2021; 13: 1561. doi:10.3390/v13081561.

[28] Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soeding J, Thompson JD, Higgins DG. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011; 7: 539. doi: 10.1038/msb.2011.75.

[29] Sievers F, Higgins DG. Clustal Omega for making accurate alignments of many protein sequences. Protein Sci. 2018; 27:135-145. doi.org/10.1002/pro.3290.

[30] Palanivelu P. DNA polymerases– An insight into their active sites and mechanism of action, In: Recent Advances in Biological Research (2019), Chapter 2, Vol 1, pp 1-39, SCIENCEDOMAIN International Book Publishers, UK, ISBN: 9788193422441, doi: 10.9734/bpi/rabr/v1.

[31] Palanivelu P. Replicative δ polymerases from plants and animals possess very similar polymerase, proofreading and regulatory domains. World J Adv Res Rev. 2024; 23:2239–2260.

[32] Palanivelu P. An Insight into the Replication Mechanism of Smallpox and Monkeypox Viruses. World J Adv Res Rev. 2024; 23: 2239-2260.

[33] Palanivelu P. Replication of Genomes in Eukaryotes: Analyses of the Structure and Functions of the Replicative ε Polymerases from Yeasts, Plants and Animals. World J Adv Res Rev. 2025; 26: 545-578.

[34] Beese LS, Steitz TA. Structural basis for the 3'-5' exonuclease activity of *Escherichia coli* DNA polymerase I: a two-metal-ion mechanism. EMBO J. 1991; 10:25–33.