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 Proofreading Domain within the Guanosine Diphosphate Polyribonucleotidyltransferase Domain in the Bat-borne Deadly Human Viruses: Marburg, Ebola and Nipah

Peramachi Palanivelu *

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

World Journal of Advanced Research and Reviews, 2026, 30(01), 2246-2261

Publication history: Received on 13 March 2026; revised on 21 April 2026; accepted on 23 April 2026

Article DOI: https://doi.org/10.30574/wjarr.2026.30.1.1047

Abstract

The proofreading 3'→5' exonuclease is an indispensable class of exonucleases that maintains the integrity of genomes in all organisms, from viruses to humans. It plays a crucial role in regulating the fidelity of genome replication and its diversity. Typically, they are placed on the same polypeptide along with the DNA polymerases as a separate domain or as an associated enzyme of the replicase multienzyme complexes. In the non-segmented, *positive-strand,* bat-borne RNA viruses like Severe Acute/Middle East Respiratory Syndrome Coronaviruses (SARS-CoVs and MERS-CoV), it is found placed along with the methyltransferase (MTase) gene of the non-structural protein, NSP14. However, in the non-segmented, *negative-strand,* bat-borne RNA viruses like Marburg, Ebola and Nipah viruses, it is not found along with the MTase gene, but identified within a novel type of capping enzyme, viz. guanosine diphosphate polyribonucleotidyltransferase (PRNTase). Despite this difference in genomic organization, the proofreading exonucleases in both groups belong to the DEDD superfamily and share a conserved active site and catalytic mechanism. Specifically, they utilise a completely conserved histidine residue that functions as a proton acceptor from a metal-activated water molecule, enabling the excision of incorrectly incorporated nucleotides during genome replication. To the best of my knowledge, this study provides the first evidence for the presence of a DEDD superfamily proofreading exonuclease domain embedded within the PRNTase domain in non-segmented, negative-sense bat-borne RNA viruses such as Marburg, Ebola, and Nipah viruses.

Keywords: Bat-borne human viruses; SARS-CoVs; Marburg virus; Ebola virus; Nipah virus; Guanosine diphosphate polyribonucleotidyltransferase; PRNTase; Methyltransferase; Proofreading exonucleases; DEDD superfamily of proofreading exonucleases.

1. Introduction

Viruses are ubiquitous in nature and are found to infect and proliferate in all kingdoms of life. They adapt their lifecycle to the organism they infect and after successful infection, they mainly make use of the host machineries for their survival and proliferation. Based on the genetic material they possess, they are broadly classified into RNA and DNA viruses. Both types of viruses cause diseases in plants, animals and humans of endemic and pandemic proportions, resulting in major global healthcare crises and unprecedented economic losses. However, the majority of human and animal diseases of endemic and pandemic nature are caused by the RNA viruses like SARS-CoVs, MERS-CoV, Marburg, Ebola, Nipah, Rabies, etc. [1]. The genome of RNA viruses is further classified into segmented and non-segmented types and further into positive- or negative-strand subtypes, depending on whether they possess a positive- or a negative-strand RNA genome. Both types cause severe viral diseases of both endemic and pandemic proportions, necessitating a deeper understanding of their pathogenic and replication mechanisms in order to develop effective therapeutic interventions [2]. For example, Measles, Mumps, Influenza viruses (A, B and C), Nipah, Hendra, Cedar, Ebola, Marburg, Rabies, etc.

* Corresponding author: Peramachi Palanivelu

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.

belong to the negative-strand RNA viruses and SARS-CoVs, MERS-CoV, Polio, Hepatitis C, Dengue, Zika, Yellow fever, Rubella, Encephalomyocarditis, HIVs, etc. belong to positive-strand RNA viral types. The Marburg, Ebola, Nipah viruses which are analysed in this communication are bat-borne viruses and belong to the same order *Mononegavirales*, as all of them possess a non-segmented, negative-strand genome and are found to infect both animals and humans. In addition to Marburg, Ebola, Nipah viruses, other members that also cause deadly human diseases in this order include rabies virus, human respiratory syncytial virus, measles and mumps viruses, etc. The order includes eleven viral families, viz. *Artoviridae*, *Bornaviridae*, *Filoviridae*, *Lispiviridae*, *Mymonaviridae*, *Nyamiviridae, Paramyxoviridae*, *Pneumoviridae*, *Rhabdoviridae*, *Sunviridae*, and *Xinmoviridae* [3]. Both Marburg and Ebola viruses belong to the same Family, viz. *Filoviridae,* and belong to the Genera, Marburgvirus and Ebolavirus, respectively, whereas the Nipah virus belongs to the Family, *Paramyxoviridae* and belongs to the Genus*, Henipavirus*. The Marburg and Ebola viruses cause severe haemorrhagic fever and death within a matter of days, whereas the Nipah virus causes acute respiratory infection and fatal encephalitis (swelling of the brain) with high fatality rates of up to 75%. Interestingly, all the three viruses use fruit bats as their natural reservoir and use primates and pigs as their intermediary hosts, respectively. Importantly, vaccines or antivirals are not available for these viruses, (vaccine is available only for the Zaire strain of Ebola virus and not for other strains of Ebola). The Marburg virus was first identified in **1967** in **Marburg**, Germany, and the Ebola virus was first identified in **1976** near the **Ebola** River in the Democratic Republic of the Congo (DRC) and the Nipah virus was first reported from Malaysia in 1998. All three viruses cause deadly human diseases, e.g., Marburg and Ebola viruses cause a severe haemorrhagic fever and multiorgan failure in humans, whereas the Nipah virus causes fatal encephalitis and brain damage. As the Nipah and Marburg viruses do not have any vaccine or antivirals to control them, the reported fatality rates are extremely high and reported up to 90% (as compared to the COVID-19 pandemic, which was only about 1% global average). It is interesting to note that bats are the natural reservoirs for these endemic/pandemic viruses. For example, reports confirm that SARS-CoV-1 and -2 and MERS-CoV originally originated from common bats like horseshoe bats (*Rhinolophus* sp.), whereas the other three viruses, viz. Marburg, Ebola and Nipah are from fruit bats.

1.1. Replication/Transcription and Proofreading Enzymes in Marburg, Ebola and Nipah Viruses

In these non-segmented, negative-strand RNA viruses, both transcription and replication are accomplished by the same enzyme, viz. by an RNA-dependent RNA polymerase (RdRp, EC 2.7.7.48). In fact, this enzyme forms a part of a large protein known as L-protein of *M*r of ~250 kDa. Other components of the L-protein are the guanosine diphosphate polyribonucleotidyltransferase (PRNTase, a capping enzyme), a Connector Domain (CD) and a Methyltransferase (MTase). The MTase exhibits a dual specificity i.e., it methylates the ribose at 2′-O position on the first nucleotide of the transcript and then methylates the capped 5'-guanylate at N-7 position. The CD and the C-terminal domain (CTD), both with nonenzymatic functions, flank the MTase. The phosphoprotein, P acts as a non-catalytic cofactor for the RdRp. It is interesting to note that a DEDD superfamily of PR exonuclease active site amino acids are found located within the PRNTase domain of the L-Protein (Fig. 1A). Interestingly, all these enzymes are considered for the development of potential antiviral drugs for effective control of these viruses, which is further elaborated in the following sections.

1.2. Viral RdRp as a Potential Drug Target

Consistent with their shared genome structure, the non-segmented, negative-strand RNA viruses have evolved similar ways to transcribe their genome into mRNAs and then to replicate to produce new genomes. A single viral RdRp (of the L-protein) performs both the transcription as well as replication of the viral genome. In the first phase of its lifecycle, i.e., after viral entry via receptor binding and membrane fusion, the viral mRNAs are transcribed, which are then capped, methylated by the other two enzymes of the L-protein. The poly-A tails are added to the mRNAs at their 3'-ends by a stuttering mechanism at a slippery stop-site present at the end of the viral genes. The capped and tailed mRNAs are then translated into viral proteins using the host machinery. In the transcription mode, the RdRp performs the sequential transcription of all the mRNAs using a 'termination-reinitiation' mechanism responding to 'gene-start' and 'gene-end' signals. Some polymerases could also disengage from the template at each gene junction, resulting in a decreasing abundance of transcripts from the 3' to the 5' end of the genome. Once enough nuclear capsid proteins (N) are made from its mRNA, the viral RNA is now tightly encapsidated by the N and now, the RdRp switches from transcription to the replication mode which is tightly linked to the intracellular N protein concentrations. It is interesting to note that in the replication mode, the polymerase replicates the whole viral genome without recognizing the transcription signals, and the replicated genome is not capped or polyadenylated. Thus, the viral RNA replication is a two-step process, i.e., in the first step, a full-length copy of the genome with positive polarity is synthesized, which is also known as the antigenomic RNA, the plus-strand, intermediate RNA. The antigenomic RNA is concurrently encapsidated by the viral nucleoprotein, N. The encapsidated antigenomic RNA now serves as a template for the synthesis of viral genomic RNA, i.e., the negative-strand genomic RNA [4]. Thus, the RdRp of these negative-strand RNA viruses plays a dual role, first as a transcriptase, and then as a replicase (except in retro viruses) in animal and human cells. Therefore, RdRps from these viruses have been the main target for antiviral drug development to control the proliferation of these viruses in animal and human cells.

1.3. PR Exonucleases and Their Classification

The viral RdRps (EC 2.7.7.48) belong to the Main class of 'Transferases' and are thus, involved in the transfer of ribonucleoside triphosphates (NTPs) to the 3' growing end of the nascent RNA chain [5]. Unlike the prokaryotic and eukaryotic RNA polymerases, the viral RdRps make a high rate of error (~10^{-4}) during copying of the viral genomes [6]. This is attributed mostly due to lack of an effective proofreading function in most of these RNA viruses. However, the lack of efficient proofreading activity also favours constant creation of new genetic variants to adapt to constantly changing environmental conditions. Interestingly, such mutations may result in 'loss-of-function' or 'gain-of-function' where both are exploited by these viruses for their successful adaptation to new environments [6, 7]. However, to maintain the integrity of their genomes and their identity, they are also equipped with a PR mechanism and is discussed further.

The PR 3'→5' exonucleases (EC 3.1.11.x) are an indispensable class of exonucleases, evolved to maintain the integrity of the genome in all living organisms. This is because, even one mistake in critical areas of an important protein or an enzyme is detrimental to the survival of organisms. These PR exonucleases usually exist either as an independent domain from the polymerase domain on the same polypeptide chain as a multifunctional enzyme (MFE) or as an independent, accessory subunit in the replicase multienzyme complex (MEC). For example, in the bacterial DNA polymerases I, three different enzymes are present on a single polypeptide as three distinct domains and exhibit three different activities, viz. i) polymerization, ii) proofreading, and iii) DNA repair. The second type of PR exonuclease exists as an independent subunit of a MEC, e.g., the ε-subunit (a Zn^{2+}-dependent 3′→5′ exonuclease) of the bacterial DNA replicases [8]]. However, in some viruses it is located along with nucleic acid modifying enzymes. For example, in SARS- and MERS-CoVs, the PR exonuclease is found in the non-structural protein (NSP14) along with the methyltransferase [9] (Fig. 2B). But, in the non-segmented, negative-strand viruses like Marburg, Ebola and Nipah, a PR exonuclease domain is identified, but not with the RdRp or with the MTase, but on a different enzyme, viz. witin the PRNTase (which transfers the 5'-monophosphate of the nascent RNA transcript onto a GDP acceptor) (Fig. 2A).

Based on the active site amino acids, the PR exonucleases are broadly classified into two superfamilies, viz. DEDD and Polymerase and Histidinol Phosphatase (PHP) [10,11]. The PR exonucleases of the bacterial and viral RdRps use four acidic amino acids in the active site, viz. DEDD, for metal-binding and catalysis, and hence, they are classified under the DEDD superfamily. Whereas the DNA polymerases X, the co-editing exonuclease of DNA polymerases III [11] and YcdX phosphoesterases [12] use PHP superfamily, which is essentially a Zn-based enzyme, with the active site amino acid structure, HxH-E-H-D. It uses His residues for metal-binding and catalysis [9]. It is interesting to note that the PA subunit of the human influenza viruses uses both types of PR exonucleases [13]. The DEDD superfamily consists of two subfamilies, viz. DEDD(Y) and DEDD(H), depending upon whether they employ an invariant Y or a H as the proton acceptor during catalysis [Table 2]. The RdRps in the positive-strand of SARS-CoVs, SARS-related CoVs and other human CoVs belong to the DEDD(H) subfamily and is located along with the MTase gene [9]. In the positive-strand viruses like Marburg, Ebola and Nipah, a similar PR exonuclease is identified within the PRNTase domain and it also belongs to the DEDD(H) type which has not been reported so far in these types of non-segmented, positive-strand RNA viruses. This 3'→5' PR exonuclease is also considered as a potential drug target to contain the proliferation of these deadly RNA viruses.

Maintenance of genome stability is very important for all living organisms and it relies mainly on the error-free replication of genomes by DNA and RNA polymerases. This is equally important for viruses also. 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 DNA and RNA polymerases exhibit strong discrimination for nucleoside triphosphates (NTPs) and deoxynucleoside triphosphates (dNTPs) and rarely insert a wrong nucleotide during the replication process and hence, their error rate is very, very minimum and is usually in the order of 10^{-6}-10^{-9} and 10^{-4} to 10^{-6}, respectively. To overcome even this minimal error rates during replications, these crucial polymerases are invariably associated with PR mechanisms to correct any insertion error(s) during genome replication. However, when a mismatch is encountered by the DNA or RNA polymerases during replication, the polymerases stall/pause, which in-turn activates the PR function that promptly excises the mismatch. The PR exonuclease belongs to 3'→5' type and they excise any wrongly added nucleotide from the 3'-growing end, and thus, helping the polymerases to perform error-free genome replication. Following the excision of the wrong base, the correct base is inserted and replication proceeds. This important PR step in living organisms ensures the original DNA/RNA template is copied without any mistake and passed on to the next generation. Thus, the PR functions play a crucial role in maintaining the integrity of genomes. It is interesting to note that the SARS-CoV-2 PR exonuclease knockout mutant was unable to replicate, suggesting possibilities for development of effective antivirals for PR exonucleases to contain these RNA viruses [9].

1.4. PRNTase: A Novel, Capping Enzyme of RNA Viruses as a Potential Drug Target

After, RdRp and PR exonuclease, the third line of potential antiviral development strategies are directed towards the capping enzyme, the PRNTase (EC 2.7.7.88). PRNTase catalyzes the transfer of 5′-triphosphate of nascent mRNAs to a GDP molecule, forming the essential cap structure on mRNAs with concomitant release of an inorganic pyrophosphate. This novel strategy is followed in a large number of the negative-strand RNA viruses, like, Marburg, Ebola, Nipah, Rabies, Vesicular Stomatitis Virus, etc. [14, 15], where the PRNTase is invariably a part of the L-protein. Interestingly, this mechanism of capping of an mRNA is distinctly different from the mRNA capping mechanism of the host cells and hence, is considered as a potential target for antiviral drug development without affecting the host cells. It should be noted that the 5′-cap structure is an essential component of all eukaryotic mRNAs, because it confers stability to mRNAs, facilitates the capped mRNAs for initiation of translation by recruiting the ribosome machinery, protects the mRNAs from degradation by exonucleases and also evades antiviral host immune response. It is interesting to note that two types of capping mechanisms operate in biological systems. For example, all known eukaryotic and some viral (dsRNA and DNA viruses) mRNAs are capped by a type of capping enzyme system that transfers a GMP moiety to the 5′-diphosphate-end of the acceptor mRNA (i. e., after the removal of the terminal γ-phosphate by a triphosphatase) to generate the cap structure, 5′-G-ppp-mRNA. However, the negative-strand RNA viruses of the order *Mononegavirales,* where most of the deadly human viruses belong to, possess an unconventional capping mechanism, in which a GDP is transferred (produced by the action of the GTPase associated with the L-protein) to the 5′-p of the mRNAs to from 5′-*G-pp*p-mRNA with the release of an inorganic pyrophosphate (*PPi*). This novel mRNA capping mechanism by PRNTase is present as an enzymatic domain in the multifunctional L-protein of the non-segmented, negative-strand RNA viruses like Nipah, Hendra, Cedar, Lyssavirus, Ebola, Marburg, Rabies, etc. In these viruses, the PRNTase use an invariant –HR- pair to add the 5′-cap structure to the nascent mRNAs by covalent catalysis [16]. In other words, the PRNTase domain carries out mRNA capping with a nucleophilic His residue of the invariant -HR- motif. A lone pair of electrons at the $N^{\varepsilon 2}$ position of the His nucleophilically attacks the α-phosphate in the 5′-triphophate group of the mRNA-start, resulting in the formation of a covalent enzyme-(hystidyl-$N^{\varepsilon 2}$)-pRNA intermediate (called L-pRNA) with the concomitant release of the inorganic pyrophosphate (PP_i). In the second step, the L-pRNA transfers the pRNA to the GDP molecule, resulting in the formation of the 5′ cap structure, 5′-Gppp-mRNA. The conserved R residue of the -HR- motif is also essential for the intermediate formation such as binding to the 5′-triphosphate group of 5′-ppp-mRNA and donating a proton to the pyrophosphate leaving group resulting in the formation of an inorganic pyrophosphate (PP_i) [16, 17]. Furthermore, Li et al. [18] reported that RNA capping with GTP is abolished with the H→A or R→A mutation of the -HR- motif. Interestingly, as the capping mechanism of these viruses is different from the host cell capping mechanism, it is also considered as a potential drug target to contain the spread of these viruses. It is interesting to note that a similar capping mechanism, i.e., the GDP addition to the 5'ppp-mRNA is also reported for the non-segmented, positive-strand, bat-borne viruses like the SARS-CoVs, suggesting that it is a common mechanism in these non-segmented RNA viruses [19]. Figures 2A and 2B show their organization in non-segmented, negative- and positive-strand RNA viruses, respectively.

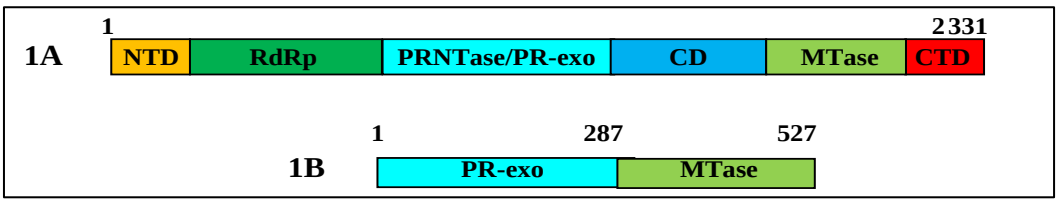


Figure 1A Organization of the PR exonuclease (PR-exo) domain in the PRNTase region of the L-protein in Marburg (2331), Ebola (2212) and Nipah (2244) viruses

Figure 1B Organization of PR exonuclease and MTase domains in the non-structural protein (NSP-14) of SARS- and MERS-CoVs. (Amino acids 1-287 represent PR exonuclease domain, and 288-527 represent N7-MTase domain in SARS-CoV-1 &-2 and 288-524 in MERS-CoV)

2. Materials and Methods

Full-length L-protein sequences from the Marburg, Ebola and Nipah viruses are obtained from PUBMED and SWISS-PROT databases. The advanced version of Clustal Omega was used for protein multiple sequence alignment (MSA) analysis. Along with the conserved motifs and sequence similarities identified by the bioinformatics analysis, the data already available from biochemical, SDM experiments and X-ray crystallographic analyses on these enzymes were used to arrive at the possible active amino acids in the PR exonucleases of the Marburg, Ebola and Nipah viruses.

3. Results and Discussion

3.1. Identification of PR Exonuclease Domain in the Marburg Virus

Figure 2 shows the PRNTase region in the L-protein of different strains of the Marburg virus (only the region required for the discussion is shown here). The Marburg virus strain Musoke (MARV/Mus) is a highly virulent variant of the Marburg virus identified from a 1980 human case in Kenya and it is used as the standard strain and highlighted in yellow. It is frequently used in research to study Marburg viral pathogenesis and evaluate medical countermeasures. Besides, Musoke is a cornerstone strain for understanding filovirus infections, which is characterized by a severe, often fatal haemorrhagic fever and multi-organ failure in humans and non-human primates. The tentative region that harbours the PRNTase region is shown with arrow marks. Interestingly, a complete set of conserved PR exonuclease active site amino acids are identified in this region and highlighted in blue. It is clear from the conserved motifs and amino acids that it belongs to the DEDD-exonuclease superfamily. It is interesting to note that the DEDD superfamily PR exonuclease is the most common PR enzyme among the DNA/RNA polymerases as shown in Table 1. The Marburg virus uses a completely conserved H within the region as the proton acceptor from a metal-bound water molecule to initiate the PR activity as shown in Fig. 6. The proposed mechanism of action for the DEDD superfamily of exonucleases is already explained by Palanivelu [20]. Additionally, a putative Zinc-binding motif (ZBM) is also observed in the PRNTase region of the Marburg virus (highlighted in orange).

CLUSTAL O (1.2.4) MSA of the PRNTase region of the L-proteins from different strains of the Marburg virus and from different regions

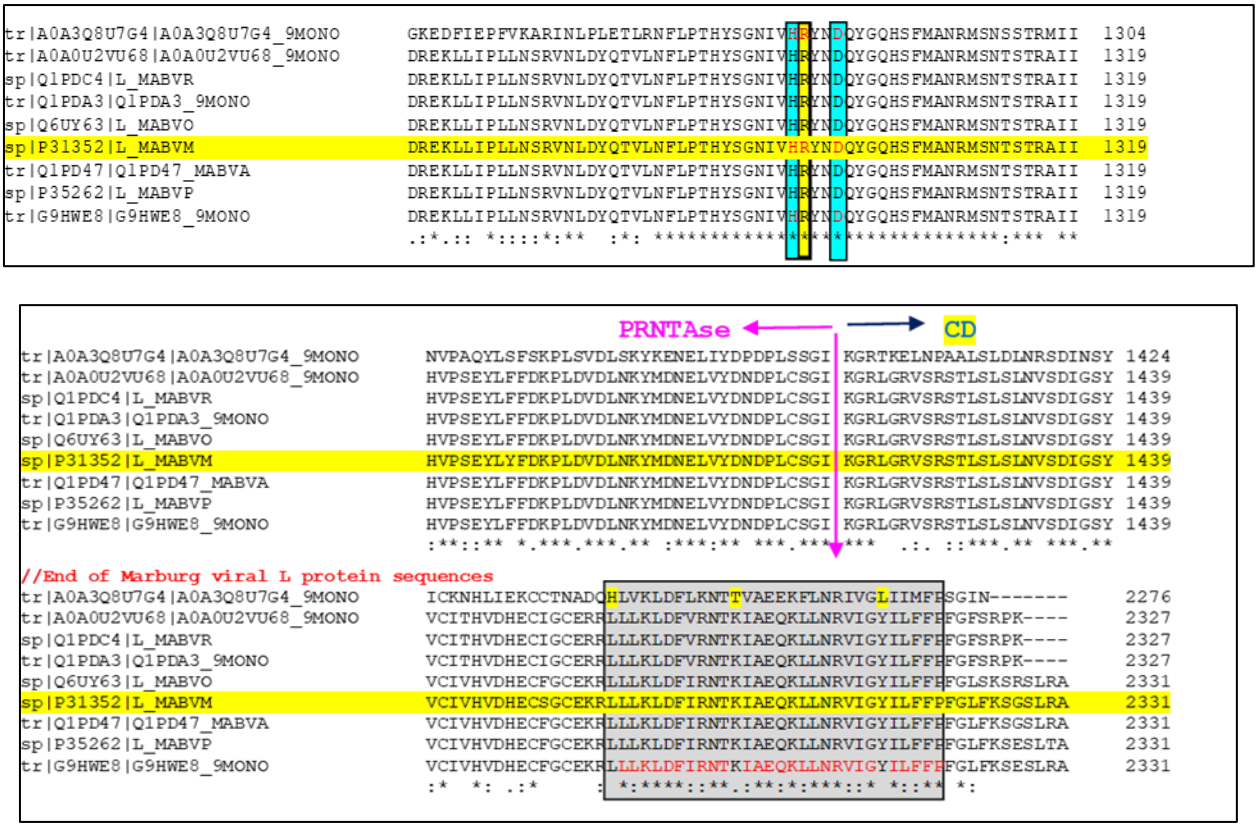
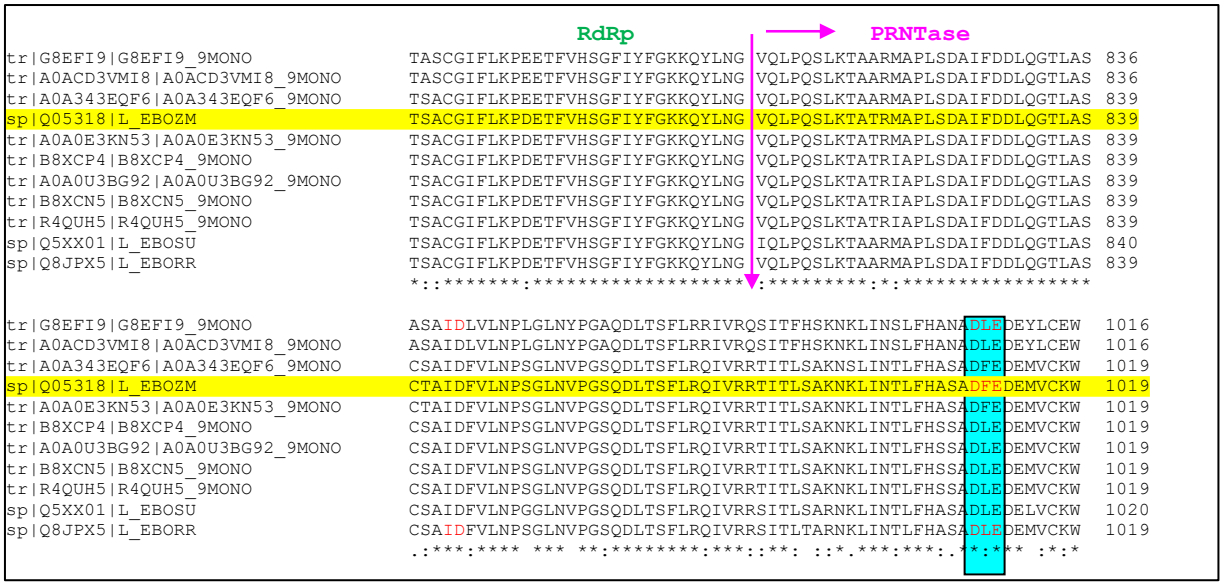


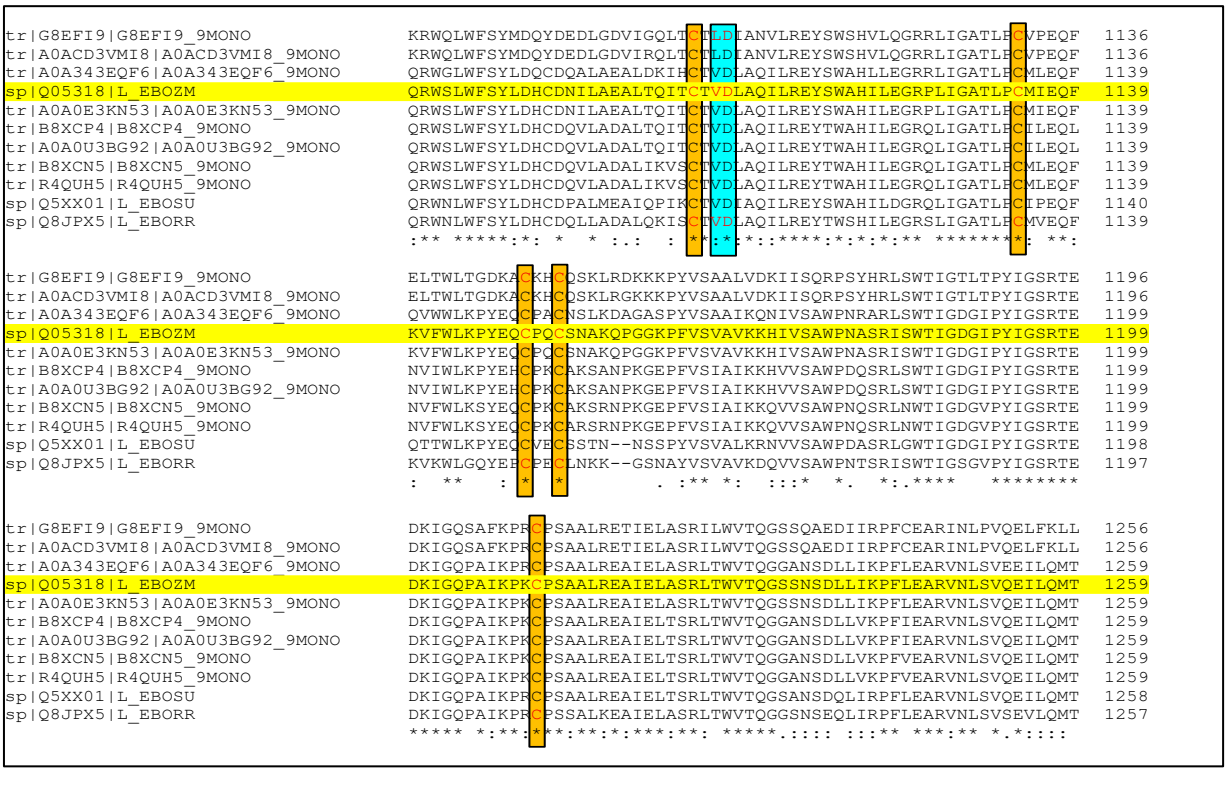
Figure 2 MSA of the PRNTase region of the L-protein from different strains of the Marburg virus and from different regions. CD, Connector Domain

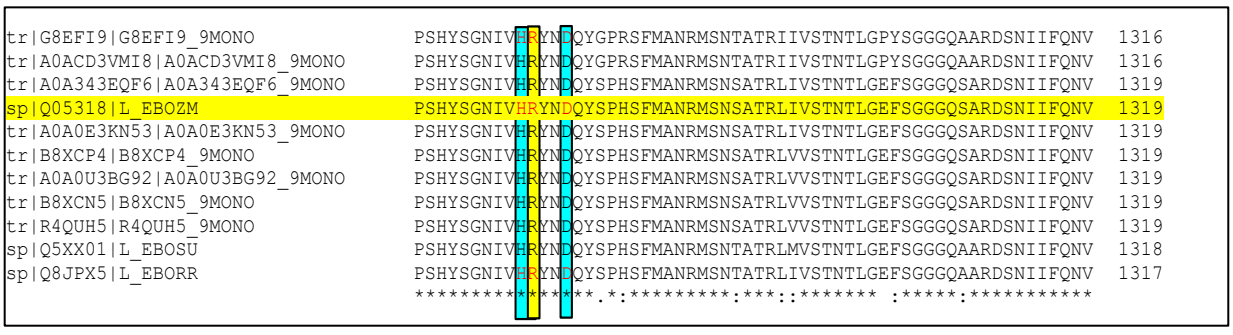
3.2. Identification of PR Exonuclease Domain in Ebola Virus

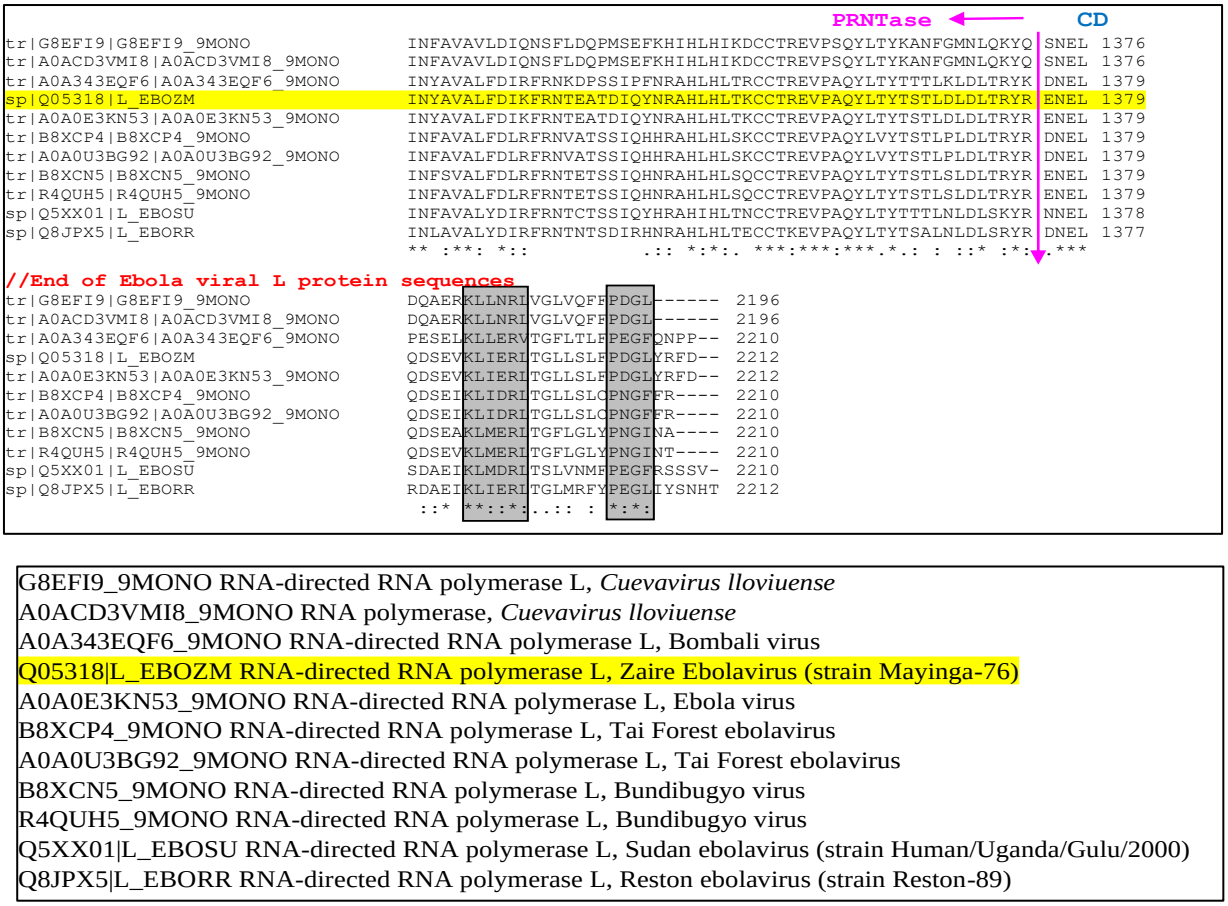
Figure 3 shows the PRNTase region in the L-protein of different strains of Ebola virus (only the region required for the discussion is shown here). Ebola virus and its genus was originally named after Zaire (now the DRC), the country where it was first described and is used as the standard strain and highlighted in yellow. The tentative region that harbours the PRNTase region is shown with arrow marks. Interestingly, a complete set of conserved PR exonuclease active site amino acids are identified in this region and highlighted in blue. It is clear from the conserved motifs and amino acids, the PR exonuclease belongs to the DEDD-exonuclease superfamily. It is interesting to note that the DEDD superfamily PR exonuclease is the most common enzyme among the DNA/RNA polymerases as shown in Table 1. The Ebola virus uses a completely conserved H within this region as the proton acceptor from a metal-bound water molecule to initiate the PR activity as shown in Fig. 6. The proposed mechanism for the DEDD superfamily of exonucleases is already explained by Palanivelu [20]. A putative ZBM is also observed in the Ebola virus PRNTase region and is highlighted in orange.

CLUSTAL O (1.2.4) MSA of the PRNTase region of the L-proteins of different strains of Ebola virus and from different regions.


RdRp
PRNTase
tr|G8EFI9|G8EFI9_9MONO TASCGIFLKPEETFVHSGFIYFGKKQYLNG VQLPQSLKTAARMAPLSDAIFDDLQGTLAS 836
tr|A0ACD3VMI8|A0ACD3VMI8_9MONO TASCGIFLKPEETFVHSGFIYFGKKQYLNG VQLPQSLKTAARMAPLSDAIFDDLQGTLAS 836
tr|A0A343EQF6|A0A343EQF6_9MONO TSACGIFLKPEETFVHSGFIYFGKKQYLNG VQLPQSLKTAARMAPLSDAIFDDLQGTLAS 839
sp|Q05318|L_EBOZM TSACGIFLKPDETFVHSGFIYFGKKQYLNG VQLPQSLKTATRMAPLSDAIFDDLQGTLAS 839
tr|A0A0E3KN53|A0A0E3KN53_9MONO TSACGIFLKPDETFVHSGFIYFGKKQYLNG VQLPQSLKTATRMAPLSDAIFDDLQGTLAS 839
tr|B8XCP4|B8XCP4_9MONO TSACGIFLKPDETFVHSGFIYFGKKQYLNG VQLPQSLKTATRIAPLSDAIFDDLQGTLAS 839
tr|A0A0U3BG92|A0A0U3BG92_9MONO TSACGIFLKPDETFVHSGFIYFGKKQYLNG VQLPQSLKTATRIAPLSDAIFDDLQGTLAS 839
tr|B8XCN5|B8XCN5_9MONO TSACGIFLKPDETFVHSGFIYFGKKQYLNG VQLPQSLKTATRIAPLSDAIFDDLQGTLAS 839
tr|R4QUH5|R4QUH5_9MONO TSACGIFLKPDETFVHSGFIYFGKKQYLNG VQLPQSLKTATRIAPLSDAIFDDLQGTLAS 839
sp|Q5XX01|L_EBOSU TSACGIFLKPDETFVHSGFIYFGKKQYLNG IQLPQSLKTAARMAPLSDAIFDDLQGTLAS 840
sp|Q8JPX5|L_EBORR TSACGIFLKPDETFVHSGFIYFGKKQYLNG VQLPQSLKTAARMAPLSDAIFDDLQGTLAS 839
tr|G8EFI9|G8EFI9_9MONO ASAIDLVLNPLGLNYPGAQDLTSFLRRIVRQSITFHSKNKLINSLFHANADLEDEYLCEW 1016
tr|A0ACD3VMI8|A0ACD3VMI8_9MONO ASAIDLVLNPLGLNYPGAQDLTSFLRRIVRQSITFHSKNKLINSLFHANADLEDEYLCEW 1016
tr|A0A343EQF6|A0A343EQF6_9MONO CSAIDFVLNPSGLNVPGSQDLTSFLRQIVRRTITLSAKNSLINTLFHASADFEDEMVCKW 1019
sp|Q05318|L_EBOZM CTAIDFVLNPSGLNVPGSQDLTSFLRQIVRRTITLSAKNKLINTLFHASADFEDEMVCKW 1019
tr|A0A0E3KN53|A0A0E3KN53_9MONO CTAIDFVLNPSGLNVPGSQDLTSFLRQIVRRTITLSAKNKLINTLFHASADFEDEMVCKW 1019
tr|B8XCP4|B8XCP4_9MONO CSAIDFVLNPSGLNVPGSQDLTSFLRQIVRRTITLSAKNKLINTLFHSSADLEDEMVCKW 1019
tr|A0A0U3BG92|A0A0U3BG92_9MONO CSAIDFVLNPSGLNVPGSQDLTSFLRQIVRRTITLSAKNKLINTLFHSSADLEDEMVCKW 1019
tr|B8XCN5|B8XCN5_9MONO CSAIDFVLNPSGLNVPGSQDLTSFLRQIVRRTITLSAKNKLINTLFHSSADLEDEMVCKW 1019
tr|R4QUH5|R4QUH5_9MONO CSAIDFVLNPSGLNVPGSQDLTSFLRQIVRRTITLSAKNKLINTLFHSSADLEDEMVCKW 1019
sp|Q5XX01|L_EBOSU CSAIDFVLNPGGLNVPGSQDLTSFLRQIVRRSITLSARNKLINTLFHASADLEDELVCKW 1020
sp|Q8JPX5|L_EBORR CSAIDFVLNPSGLNVPGSQDLTSFLRQIVRRSITLTARNKLINTLFHASADLEDEMVCKW 1019


tr|G8EFI9|G8EFI9_9MONO KRWQLWFSYMDQYDEDLGDVIGQLTCTLDIANVLREYSWSHVLQGRRLIGATLFCVPEQF 1136
tr|A0ACD3VMI8|A0ACD3VMI8_9MONO KRWQLWFSYMDQYDEDLGDVIRQLTCTLDIANVLREYSWSHVLQGRRLIGATLFCVPEQF 1136
tr|A0A343EQF6|A0A343EQF6_9MONO QRWGLWFSYLDQCDQALAEALDKIHCTVDLAQILREYSWAHLLEGRRLIGATLFCMLEQF 1139
sp|Q05318|L_EBOZM QRWSLWFSYLDHCDNILAEALTQITCTVDLAQILREYSWAHILEGRPLIGATLPCMIEQF 1139
tr|A0A0E3KN53|A0A0E3KN53_9MONO QRWSLWFSYLDHCDNILAEALTQITCTVDLAQILREYSWAHILEGRPLIGATLFCMIEQF 1139
tr|B8XCP4|B8XCP4_9MONO QRWSLWFSYLDHCDQVLADALTQITCTVDLAQILREYTWAHILEGRQLIGATLFCILEQL 1139
tr|A0A0U3BG92|A0A0U3BG92_9MONO QRWSLWFSYLDHCDQVLADALTQITCTVDLAQILREYTWAHILEGRQLIGATLFCILEQL 1139
tr|B8XCN5|B8XCN5_9MONO QRWSLWFSYLDHCDQVLADALIKVSCTVDLAQILREYTWAHILEGRQLIGATLFCMLEQF 1139
tr|R4QUH5|R4QUH5_9MONO QRWSLWFSYLDHCDQVLADALIKVSCTVDLAQILREYTWAHILEGRQLIGATLFCMLEQF 1139
sp|Q5XX01|L_EBOSU QRWNLWFSYLDHCDPALMEAIQPIKCTVDIAQILREYSWAHILDGRQLIGATLFCIPEQF 1140
sp|Q8JPX5|L_EBORR QRWNLWFSYLDHCDQLLADALQKISCTVDLAQILREYTWSHILEGRSLIGATLFCMVEQF 1139
tr|G8EFI9|G8EFI9_9MONO ELTWLTGDKACKHCQSKLRDKKKPYVSAALVDKIISQRPSYHRLSWTIGTLTPYIGSRTE 1196
tr|A0ACD3VMI8|A0ACD3VMI8_9MONO ELTWLTGDKACKHCQSKLRGKKKPYVSAALVDKIISQRPSYHRLSWTIGTLTPYIGSRTE 1196
tr|A0A343EQF6|A0A343EQF6_9MONO QVWWLKPYEQCPACNSLKDAGASPYVSAAIKQNIVSAWPNRARLSWTIGDGIPYIGSRTE 1199
sp|Q05318|L_EBOZM KVFWLKPYEQCPQCSNAKQPGGKPFVSVAVKKHIVSAWPNASRISWTIGDGIPYIGSRTE 1199
tr|A0A0E3KN53|A0A0E3KN53_9MONO KVFWLKPYEQCPQCSNAKQPGGKPFVSVAVKKHIVSAWPNASRISWTIGDGIPYIGSRTE 1199
tr|B8XCP4|B8XCP4_9MONO NVIWLKPYEHCPKCAKSANPKGEPFVSIAIKKHVVSAWPDQSRLSWTIGDGIPYIGSRTE 1199
tr|A0A0U3BG92|A0A0U3BG92_9MONO NVIWLKPYEHCPKCAKSANPKGEPFVSIAIKKHVVSAWPDQSRLSWTIGDGIPYIGSRTE 1199
tr|B8XCN5|B8XCN5_9MONO NVFWLKSYEQCPKCAKSRNPKGEPFVSIAIKKQVVSAWPNQSRLNWTIGDGVPYIGSRTE 1199
tr|R4QUH5|R4QUH5_9MONO NVFWLKSYEQCPKCARSRNPKGEPFVSIAIKKQVVSAWPNQSRLNWTIGDGVPYIGSRTE 1199
sp|Q5XX01|L_EBOSU QTTWLKPYEQCVECSSTN--NSSPYVSVALKRNVVSAWPDASRLGWTIGDGIPYIGSRTE 1198
sp|Q8JPX5|L_EBORR KVKWLGQYEHCPECLNKK--GSNAYVSVAVKDQVVSAWPNTSRISWTIGSGVPYIGSRTE 1197
tr|G8EFI9|G8EFI9_9MONO DKIGQSAFKPRCPSAALRETIELASRILWVTQGSSQAEDIIRPFCEARINLPVQELFKLL 1256
tr|A0ACD3VMI8|A0ACD3VMI8_9MONO DKIGQSAFKPRCPSAALRETIELASRILWVTQGSSQAEDIIRPFCEARINLPVQELFKLL 1256
tr|A0A343EQF6|A0A343EQF6_9MONO DKIGQPAIKPRCPSAALREAIELASRLTWVTQGGANSDLLIKPFLEARVNLSVEEILQMT 1259
sp|Q05318|L_EBOZM DKIGQPAIKPKCPSAALREAIELASRLTWVTQGSSNSDLLIKPFLEARVNLSVQEILQMT 1259
tr|A0A0E3KN53|A0A0E3KN53_9MONO DKIGQPAIKPKCPSAALREAIELASRLTWVTQGSSNSDLLIKPFLEARVNLSVQEILQMT 1259
tr|B8XCP4|B8XCP4_9MONO DKIGQPAIKPKCPSAALREAIELTSRLTWVTQGGANSDLLVKPFIEARVNLSVQEILQMT 1259
tr|A0A0U3BG92|A0A0U3BG92_9MONO DKIGQPAIKPKCPSAALREAIELTSRLTWVTQGGANSDLLVKPFIEARVNLSVQEILQMT 1259
tr|B8XCN5|B8XCN5_9MONO DKIGQPAIKPKCPSAALREAIELTSRLTWVTQGGANSDLLVKPFVEARVNLSVQEILQMT 1259
tr|R4QUH5|R4QUH5_9MONO DKIGQPAIKPKCPSAALREAIELTSRLTWVTQGGANSDLLVKPFVEARVNLSVQEILQMT 1259
sp|Q5XX01|L_EBOSU DKIGQPAIKPRCPSAALREAIELTSRLTWVTQGSANSDQLIRPFLEARVNLSVQEILQMT 1258
sp|Q8JPX5|L_EBORR DKIGQPAIKPRCPSSALKEAIELASRLTWVTQGGSNSEQLIRPFLEARVNLSVSEVLQMT 1257


tr|G8EFI9|G8EFI9_9MONO PSHYSGNIVHRYNDQYGPRSFMANRMSNTATRIIVSTNTLGPYSGGGQAARDSNIIFQNV 1316
tr|A0ACD3VMI8|A0ACD3VMI8_9MONO PSHYSGNIVHRYNDQYGPRSFMANRMSNTATRIIVSTNTLGPYSGGGQAARDSNIIFQNV 1316
tr|A0A343EQF6|A0A343EQF6_9MONO PSHYSGNIVHRYNDQYSPHSFMANRMSNSATRLIVSTNTLGEFSGGGQSARDSNIIFQNV 1319
sp|Q05318|L_EBOZM PSHYSGNIVHRYNDQYSPHSFMANRMSNSATRLIVSTNTLGEFSGGGQSARDSNIIFQNV 1319
tr|A0A0E3KN53|A0A0E3KN53_9MONO PSHYSGNIVHRYNDQYSPHSFMANRMSNSATRLIVSTNTLGEFSGGGQSARDSNIIFQNV 1319
tr|B8XCP4|B8XCP4_9MONO PSHYSGNIVHRYNDQYSPHSFMANRMSNSATRLVVSTNTLGEFSGGGQSARDSNIIFQNV 1319
tr|A0A0U3BG92|A0A0U3BG92_9MONO PSHYSGNIVHRYNDQYSPHSFMANRMSNSATRLVVSTNTLGEFSGGGQSARDSNIIFQNV 1319
tr|B8XCN5|B8XCN5_9MONO PSHYSGNIVHRYNDQYSPHSFMANRMSNSATRLVVSTNTLGEFSGGGQSARDSNIIFQNV 1319
tr|R4QUH5|R4QUH5_9MONO PSHYSGNIVHRYNDQYSPHSFMANRMSNSATRLVVSTNTLGEFSGGGQSARDSNIIFQNV 1319
sp|Q5XX01|L_EBOSU PSHYSGNIVHRYNDQYSPHSFMANRMSNTATRLMVSTNTLGEFSGGGQAARDSNIIFQNV 1318
sp|Q8JPX5|L_EBORR PSHYSGNIVHRYNDQYSPHSFMANRMSNTATRLIVSTNTLGEFSGGGQAARDSNIIFQNV 1317



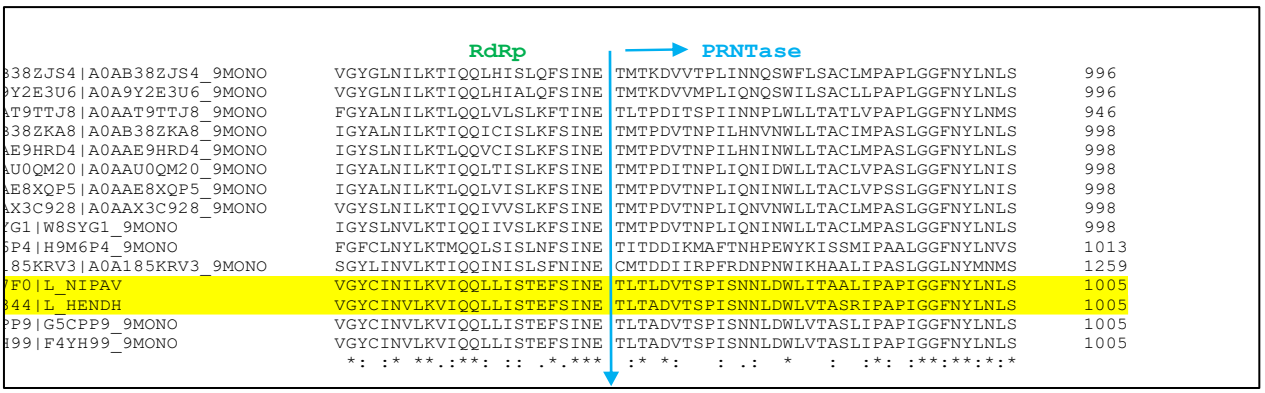
G8EFI9_9MONO RNA-directed RNA polymerase L, *Cuevavirus lloviuense*
A0ACD3VMI8_9MONO RNA polymerase, *Cuevavirus lloviuense*
A0A343EQF6_9MONO RNA-directed RNA polymerase L, Bombali virus
Q05318|L_EBOZM RNA-directed RNA polymerase L, Zaire Ebolavirus (strain Mayinga-76)
A0A0E3KN53_9MONO RNA-directed RNA polymerase L, Ebola virus
B8XCP4_9MONO RNA-directed RNA polymerase L, Tai Forest ebolavirus
A0A0U3BG92_9MONO RNA-directed RNA polymerase L, Tai Forest ebolavirus
B8XCN5_9MONO RNA-directed RNA polymerase L, Bundibugyo virus
R4QUH5_9MONO RNA-directed RNA polymerase L, Bundibugyo virus
Q5XX01|L_EBOSU RNA-directed RNA polymerase L, Sudan ebolavirus (strain Human/Uganda/Gulu/2000)
Q8JPX5|L_EBORR RNA-directed RNA polymerase L, Reston ebolavirus (strain Reston-89)

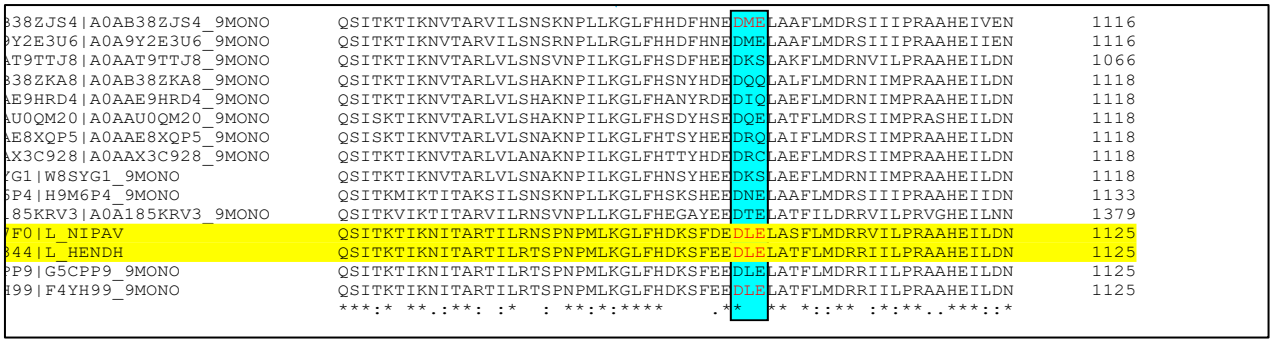
Figure 3 MSA of the PRNTase region of the L-protein of different strains of Ebola virus and from different regions

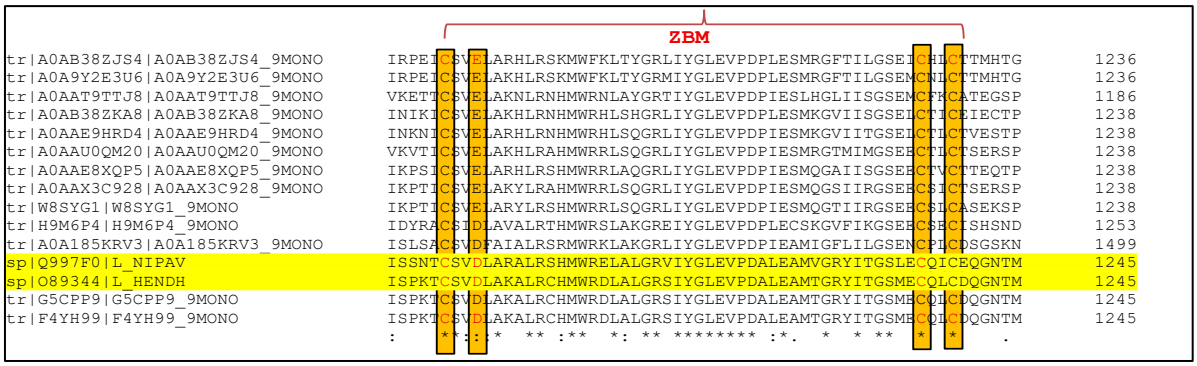
3.3. Identification of PR Exonuclease Domain in Nipah virus

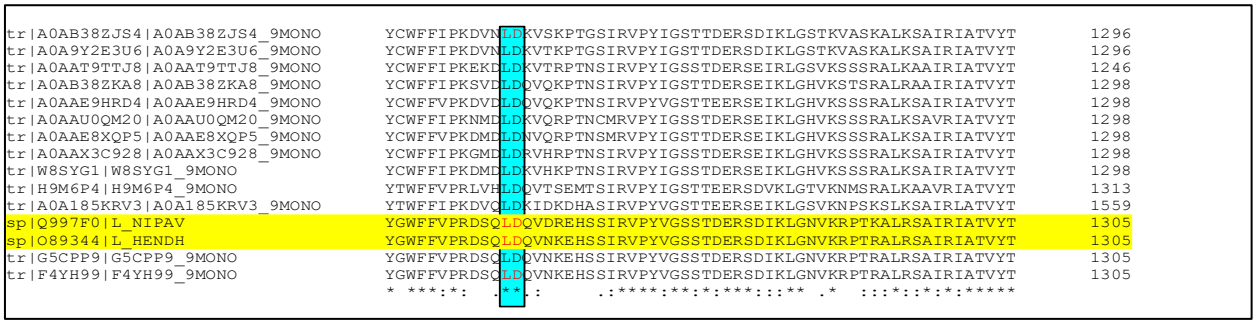
Figure 4 shows the PRNTase region in the L-protein of different strains of Nipah virus (only the region required for the discussion is shown here). The Nipah virus which was first reported from Malaysia is used as the standard strain and is highlighted in yellow. The tentative region that harbours the PRNTase region is shown with arrow marks. Interestingly, a complete set of conserved PR exonuclease active site amino acids are identified in this region and highlighted in blue. It is clear from the conserved motifs and amino acids, the PR exonuclease belongs to the DEDD-exonuclease superfamily. It is interesting to note that the DEDD superfamily PR exonuclease is the most common enzyme among the DNA/RNA polymerases as shown in Table 1. The Nipah virus uses a completely conserved H as the proton acceptor from a metal-bound water molecule to initiate the PR activity as shown in Fig. 6. The proposed mechanism for the DEDD superfamily of exonucleases is already explained by Palanivelu [20]. The Nipah viral PRNTase region exhibits a putative ZBM (highlighted in orange) as observed in the Marburg and Ebola viruses.

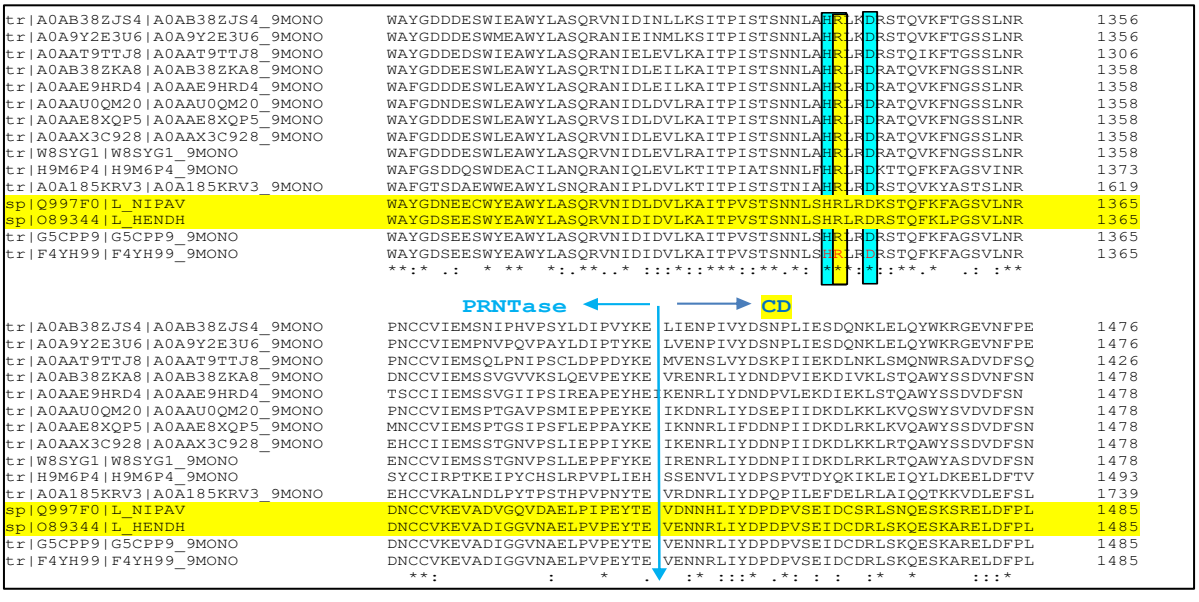
CLUSTAL O (1.2.4) MSA of the PRNTase region of the L proteins from different Nipah viral strains.






ZBM




PRNTase
CD

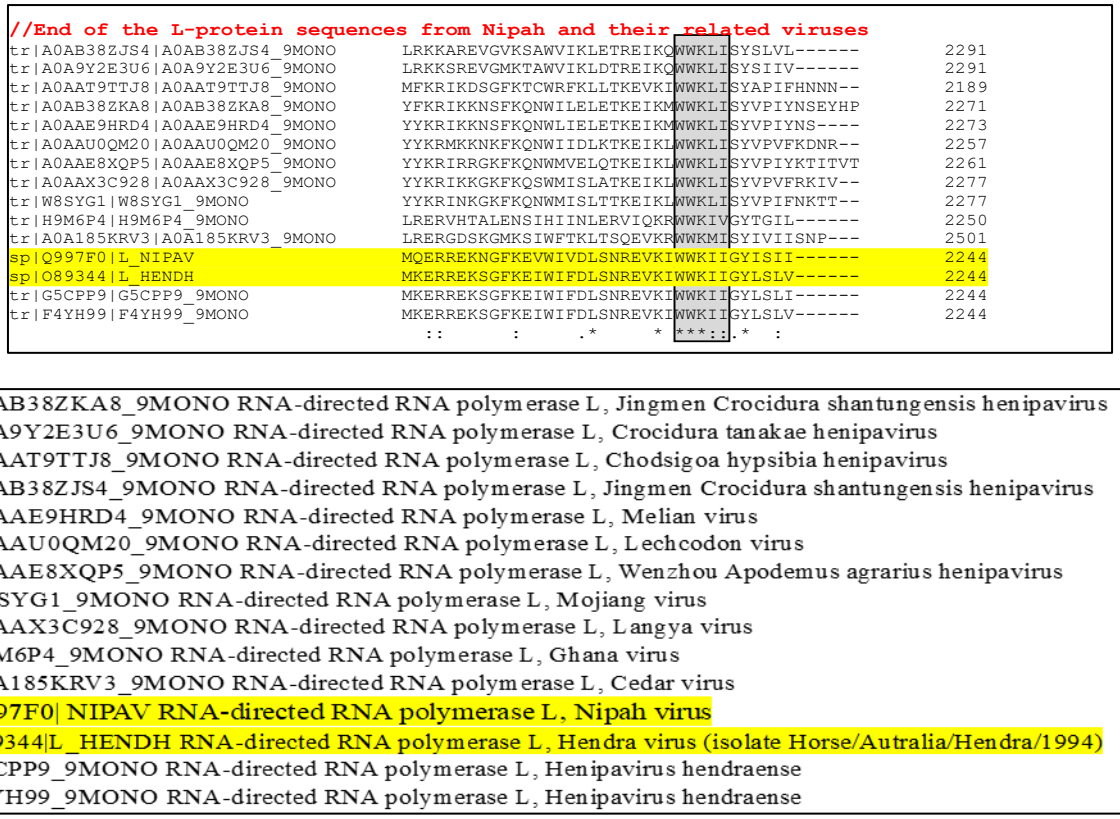
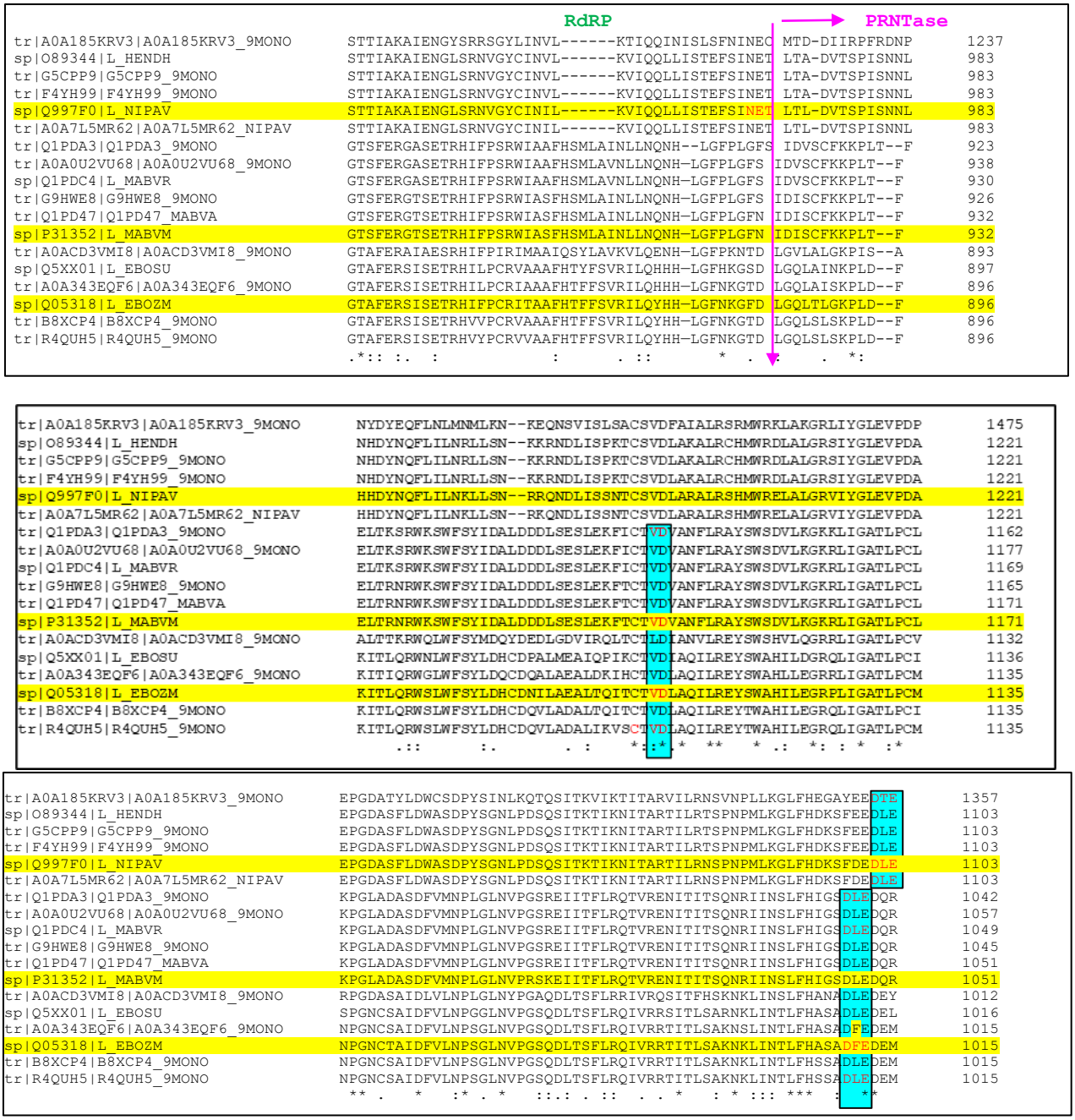


Figure 4 MSA of the PRNTase region of the L- protein of Nipah and its related viruses

3.4. 'Mix and Match' MSA analysis of PR Exonuclease Domain from All the Three Bat-borne Viruses: Marburg, Ebola and Nipah viruses

Figure 5 shows the 'Mix and Match' MSA of all the three bat-borne human viruses, viz. Marburg, Ebola and Nipah viruses (only the region required for the discussion is shown here). The representative strains of the three viruses are highlighted in yellow. The tentative PRNTase region in all three viruses is marked in arrows. The conserved motifs and amino acids of the PR exonuclease are highlighted in light blue. The typical, invariant DxE motif reported in all DEDD superfamily exonuclease superfamily is completely conserved in all three viruses, but not perfectly aligned. This is also true for the second conserved amino acid D, which slightly deviates for the Nipah virus. However, the next two active site amin acids, viz. the H and last D are completely conserved and aligned in all three viruses. It is interesting to note that the conserved -HR- motif of the PRNTase active site is also completely aligned in all three, suggesting a dual role for the H residue in both exonuclease and capping activities. The proposed overall structure of the DEDD(H) superfamily of exonuclease in all three viruses is shown in Fig. 6.

CLUSTAL O (1.2.4) 'Mix and Match' MSA of the PRNTase region of the L-protein of Nipah, Marburg and Ebola and their related viruses


RdRP
PRNTase
tr|A0A185KRV3|A0A185KRV3_9MONO STTIAKAIENGYSRRSGYLINVL------KTIQQINISLSFNINEC MTD-DIIRPFRDNP 1237
sp|O89344|L_HENDH STTIAKAIENGLSRNVGYCINVL------KVIQQLLISTEFSINET LTA-DVTSPISNNL 983
tr|G5CPP9|G5CPP9_9MONO STTIAKAIENGLSRNVGYCINVL------KVIQQLLISTEFSINET LTA-DVTSPISNNL 983
tr|F4YH99|F4YH99_9MONO STTIAKAIENGLSRNVGYCINVL------KVIQQLLISTEFSINET LTA-DVTSPISNNL 983
sp|Q997F0|L_NIPAV STTIAKAIENGLSRNVGYCINIL------KVIQQLLISTEFSINET LTL-DVTSPISNNL 983
tr|A0A7L5MR62|A0A7L5MR62_NIPAV STTIAKAIENGLSRNVGYCINIL------KVIQQLLISTEFSINET LTL-DVTSPISNNL 983
tr|Q1PDA3|Q1PDA3_9MONO GTSFERGASETRHIFPSRWIAAFHSMLAINLLNQNH--LGFPLGFS IDVSCFKKPLT--F 923
tr|A0A0U2VU68|A0A0U2VU68_9MONO GTSFERGASETRHIFPSRWIAAFHSMLAVNLLNQNH-LGFPLGFS IDVSCFKKPLT--F 938
sp|Q1PDC4|L_MABVR GTSFERGASETRHIFPSRWIAAFHSMLAVNLLNQNH-LGFPLGFS IDVSCFKKPLT--F 930
tr|G9HWE8|G9HWE8_9MONO GTSFERGTSETRHIFPSRWIASFHSMLAINLLNQNH-LGFPLGFS IDISCFKKPLT--F 926
tr|Q1PD47|Q1PD47_MABVA GTSFERGTSETRHIFPSRWIASFHSMLAINLLNQNH-LGFPLGFN IDISCFKKPLT--F 932
sp|P31352|L_MABVM GTSFERGTSETRHIFPSRWIASFHSMLAINLLNQNH-LGFPLGFN IDISCFKKPLT--F 932
tr|A0ACD3VMI8|A0ACD3VMI8_9MONO GTAFERAIAESRHIFPIRIMAAIQSYLAVKVLQENH-LGFPKNTD LGVLALGKPIS--A 893
sp|Q5XX01|L_EBOSU GTAFERSISETRHILPCRVAAAFHTYFSVRILQHHH-LGFHKGSD LGQLAINKPLD--F 897
tr|A0A343EQF6|A0A343EQF6_9MONO GTAFERSISETRHILPCRIAAAFHTFFSVRILQHHH-LGFNKGTD LGQLAISKPLD--F 896
sp|Q05318|L_EBOZM GTAFERSISETRHIFPCRITAAFHTFFSVRILQYHH-LGFNKGFD LGQLTLGKPLD--F 896
tr|B8XCP4|B8XCP4_9MONO GTAFERSISETRHVVPCRVAAAFHTFFSVRILQYHH-LGFNKGTD LGQLSLSKPLD--F 896
tr|R4QUH5|R4QUH5_9MONO GTAFERSISETRHVYPCRVVAAFHTFFSVRILQYHH-LGFNKGTD LGQLSLSKPLD--F 896

tr|A0A185KRV3|A0A185KRV3_9MONO NYDYEQFLNLMNMLKN--KEQNSVISLSACSVDFAIALRSRMWRKLAKGRLIYGLEVPDP 1475
sp|O89344|L_HENDH NHDYNQFLILNRLLSN--KKRNDLISPKTCSVDLAKALRCHMWRDLALGRSIYGLEVPDA 1221
tr|G5CPP9|G5CPP9_9MONO NHDYNQFLILNRLLSN--KKRNDLISPKTCSVDLAKALRCHMWRDLALGRSIYGLEVPDA 1221
tr|F4YH99|F4YH99_9MONO NHDYNQFLILNRLLSN--KKRNDLISPKTCSVDLAKALRCHMWRDLALGRSIYGLEVPDA 1221
sp|Q997F0|L_NIPAV HHDYNQFLILNKLLSN--RRQNDLISSNTCSVDLARALRSHMWRELALGRVIYGLEVPDA 1221
tr|A0A7L5MR62|A0A7L5MR62_NIPAV HHDYNQFLILNKLLSN--RKQNDLISSNTCSVDLARALRSHMWRELALGRVIYGLEVPDA 1221
tr|Q1PDA3|Q1PDA3_9MONO ELTKSRWKSWFSYIDALDDDLSESLEKFICTVDVANFLRAYSWSDVLKGKRLIGATLPCL 1162
tr|A0A0U2VU68|A0A0U2VU68_9MONO ELTKSRWKSWFSYIDALDDDLSESLEKFICTVDVANFLRAYSWSDVLKGKRLIGATLPCL 1177
sp|Q1PDC4|L_MABVR ELTKSRWKSWFSYIDALDDDLSESLEKFICTVDVANFLRAYSWSDVLKGKRLIGATLPCL 1169
tr|G9HWE8|G9HWE8_9MONO ELTRNRWKSWFSYIDALDDDLSESLEKFTCTVDVANFLRAYSWSDVLKGKRLIGATLPCL 1165
tr|Q1PD47|Q1PD47_MABVA ELTRNRWKSWFSYIDALDDDLSESLEKFTCTVDVANFLRAYSWSDVLKGKRLIGATLPCL 1171
sp|P31352|L_MABVM ELTRNRWKSWFSYIDALDDDLSESLEKFTCTVDVANFLRAYSWSDVLKGKRLIGATLPCL 1171
tr|A0ACD3VMI8|A0ACD3VMI8_9MONO ALTTKRWQLWFSYMDQYDEDLGDVIRQLTCTLDIANVLREYSWSHVLQGRRLIGATLPCV 1132
sp|Q5XX01|L_EBOSU KITLQRWNLWFSYLDHCDPALMEAIQPIKCTVDIAQILREYSWAHILDGRQLIGATLPCI 1136
tr|A0A343EQF6|A0A343EQF6_9MONO KITIQRWGLWFSYLDQCDQALAEALDKIHCTVDLAQILREYSWAHLLEGRRLIGATLPCM 1135
sp|Q05318|L_EBOZM KITLQRWSLWFSYLDHCDNILAEALTQITCTVDLAQILREYSWAHILEGRPLIGATLPCM 1135
tr|B8XCP4|B8XCP4_9MONO KITLQRWSLWFSYLDHCDQVLADALTQITCTVDLAQILREYTWAHILEGRQLIGATLPCI 1135
tr|R4QUH5|R4QUH5_9MONO KITLQRWSLWFSYLDHCDQVLADALIKVSCTVDLAQILREYTWAHILEGRQLIGATLPCM 1135

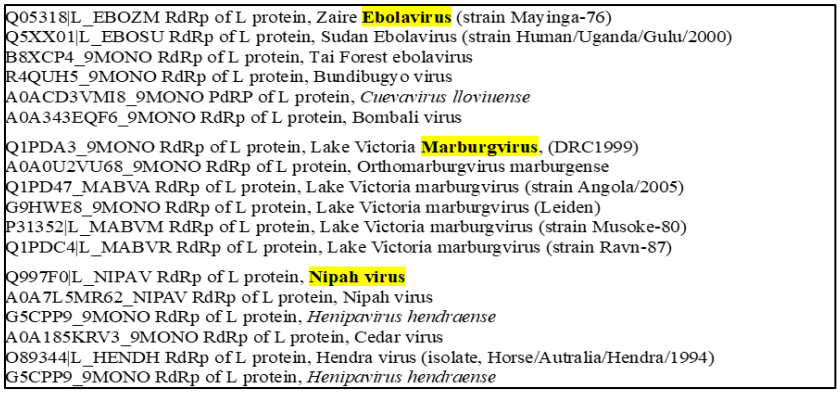
tr|A0A185KRV3|A0A185KRV3_9MONO EPGDATYLDWCSDPYSINLKQTQSITKVIKTITARVILRNSVNPLLKGLFHEGAYEEDTE 1357
sp|O89344|L_HENDH EPGDASFLDWASDPYSGNLPDSQSITKTIKNITARTILRTSPNPMLKGLFHDKSFEEDLE 1103
tr|G5CPP9|G5CPP9_9MONO EPGDASFLDWASDPYSGNLPDSQSITKTIKNITARTILRTSPNPMLKGLFHDKSFEEDLE 1103
tr|F4YH99|F4YH99_9MONO EPGDASFLDWASDPYSGNLPDSQSITKTIKNITARTILRTSPNPMLKGLFHDKSFEEDLE 1103
sp|Q997F0|L_NIPAV EPGDASFLDWASDPYSGNLPDSQSITKTIKNITARTILRNSPNPMLKGLFHDKSFDEDLE 1103
tr|A0A7L5MR62|A0A7L5MR62_NIPAV EPGDASFLDWASDPYSGNLPDSQSITKTIKNITARTILRNSPNPMLKGLFHDKSFDEDLE 1103
tr|Q1PDA3|Q1PDA3_9MONO KPGLADASDFVMNPLGLNVPGSREIITFLRQTVRENITITSQNRIINSLFHIGSDLEDQR 1042
tr|A0A0U2VU68|A0A0U2VU68_9MONO KPGLADASDFVMNPLGLNVPGSREIITFLRQTVRENITITSQNRIINSLFHIGSDLEDQR 1057
sp|Q1PDC4|L_MABVR KPGLADASDFVMNPLGLNVPGSREIITFLRQTVRENITITSQNRIINSLFHIGSDLEDQR 1049
tr|G9HWE8|G9HWE8_9MONO KPGLADASDFVMNPLGLNVPGSREIITFLRQTVRENITITSQNRIINSLFHIGSDLEDQR 1045
tr|Q1PD47|Q1PD47_MABVA KPGLADASDFVMNPLGLNVPGSREIITFLRQTVRENITITSQNRIINSLFHIGSDLEDQR 1051
sp|P31352|L_MABVM KPGLADASDFVMNPLGLNVPRSKEIITFLRQTVRENITITSQNRIINSLFHIGSDLEDQR 1051
tr|A0ACD3VMI8|A0ACD3VMI8_9MONO RPGDASAIDLVLNPLGLNYPGAQDLTSFLRRIVRQSITFHSKNKLINSLFHANADLEDEY 1012
sp|Q5XX01|L_EBOSU SPGNCSAIDFVLNPGGLNVPGSQDLTSFLRQIVRRSITLSARNKLINTLFHASADLEDEL 1016
tr|A0A343EQF6|A0A343EQF6_9MONO NPGNCSAIDFVLNPSGLNVPGSQDLTSFLRQIVRRTITLSAKNSLINTLFHASADFEDEM 1015
sp|Q05318|L_EBOZM NPGNCTAIDFVLNPSGLNVPGSQDLTSFLRQIVRRTITLSAKNKLINTLFHASADFEDEM 1015
tr|B8XCP4|B8XCP4_9MONO NPGNCSAIDFVLNPSGLNVPGSQDLTSFLRQIVRRTITLSAKNKLINTLFHSSADLEDEM 1015
tr|R4QUH5|R4QUH5_9MONO NPGNCSAIDFVLNPSGLNVPGSQDLTSFLRQIVRRTITLSAKNKLINTLFHSSADLEDEM 1015

```
tr|A0A185KRV3|A0A185KRV3_9MONO      IEAMIGFLILGSENCPL-----------------CD----SGSKNYTWFFIPKDVQLDK      1513
sp|O89344|L_HENDH                   LEAMTGRYITGSMECQL-----------------CD----QGNTMYGWFFVPRDSQLDQ      1259
tr|G5CPP9|G5CPP9_9MONO              LEAMTGRYITGSMECQL-----------------CD----QGNTMYGWFFVPRDSQLDQ      1259
tr|F4YH99|F4YH99_9MONO              LEAMTGRYITGSMECQL-----------------CD----QGNTMYGWFFVPRDSQLDQ      1259
sp|Q997F0|L_NIPAV                   LEAMVGRYITGSLECQI-----------------CE----QGNTMYGWFFVPRDSQLDQ      1259
tr|A0A7L5MR62|A0A7L5MR62_NIPAV      LEAMVGRYITGSLECQI-----------------CE----QGNTMYGWFFVPRDSQLDQ      1259
tr|Q1PDA3|Q1PDA3_9MONO              LEQFNVKWVNLSEDLKEQFKLSSDLGSPTDLLQYDCNGLHSKGADNAELNYVSCALDRKI      1222
tr|A0A0U2VU68|A0A0U2VU68_9MONO      LEQFNVKWVNLSEDLKEQFKLSSDLGSPTDLLQYDCNGLHSKGADNAELNYVSCALDRKI      1237
sp|Q1PDC4|L_MABVR                   LEQFNVKWVNLSEDLKEQFKLSSDLGSPTDLLQYDCNGLHSKGADNAELNYVSCALDRKI      1229
tr|G9HWE8|G9HWE8_9MONO              LEQFKVKWINLSEDLREQFNLSSESESTINLLPYDCKELRLGRSNDTELNYVSCALDRKV      1225
tr|Q1PD47|Q1PD47_MABVA              LEQFKVKWINLSEDLREQFNLSSDAESTINFLPYDCKELRLGGSNDTELNYVSCALDRKV      1231
sp|P31352|L_MABVM                   LEQFEVKWINLSEDLREQFNLSSDSKSTINLLPYDCKELRLEGSNDTELNYVSCALDRKV      1231
tr|A0ACD3VMI8|A0ACD3VMI8_9MONO      PEQFELTWLTGDKACKH-----------------CQ----SKLRGKKKPYVSAALVDKI      1170
sp|Q5XX01|L_EBOSU                   PEQFQTTWLKPYEQCVE-----------------CS----STN--NSSPYVSVALKRNV      1172
tr|A0A343EQF6|A0A343EQF6_9MONO      LEQFQVWWLKPYEQCPA-----------------CN----SLKDAGASPYVSAAIKQNI      1173
sp|Q05318|L_EBOZM                   IEQFKVFWLKPYEQCPQ-----------------CS----NAKQPGGKPFVSVAVKKHI      1173
tr|B8XCP4|B8XCP4_9MONO              LEQLNVIWLKPYEHCPK-----------------CA----KSANPKGEPFVSIAIKKHV      1173
tr|R4QUH5|R4QUH5_9MONO              LEQFNVFWLKSYEQCPK-----------------CA----RSRNPKGEPFVSIAIKKQV      1173
                                     * :    :                          *             ::      .
```

```
tr|A0A185KRV3|A0A185KRV3_9MONO      DAEWWEAWYLSNQRANIPLDVLKTITPISTSTNIAHRLRDRSTQVKYASTSLNRVSRHVT      1625
sp|O89344|L_HENDH                   EESWYEAWYLASQRVNIDIDVLKAITPVSTSNNLSHRLRDRSTQFKLPGSVLNRVSRYVN      1371
tr|G5CPP9|G5CPP9_9MONO              EESWYEAWYLASQRVNIDIDVLKAITPVSTSNNLSHRLRDRSTQFKFAGSVLNRVSRYVN      1371
tr|F4YH99|F4YH99_9MONO              EESWYEAWYLASQRVNIDIDVLKAITPVSTSNNLSHRLRDRSTQFKFAGSVLNRVSRYVN      1371
sp|Q997F0|L_NIPAV                   EECWYEAWYLASQRVNIDLDVLKAITPVSTSNNLSHRLRDKSTQFKFAGSVLNRVSRYVN      1371
tr|A0A7L5MR62|A0A7L5MR62_NIPAV      EECWYEAWYLASQRVNIDLDVLKAITPVSTSNNLSHRLRDKSTQFKFAGSVLNRVSRYVN      1371
tr|Q1PDA3|Q1PDA3_9MONO              ADREKLLIPLLNSRVNLDYQTVLNFLPTHYSGNIVHRYNDQYGQHSFMANRMSNTSTRAI      1342
tr|A0A0U2VU68|A0A0U2VU68_9MONO      ADREKLLIPLLNSRVNLDYQTVLNFLPTHYSGNIVHRYNDQYGQHSFMANRMSNTSTRAI      1357
sp|Q1PDC4|L_MABVR                   ADREKLLIPLLNSRVNLDYQTVLNFLPTHYSGNIVHRYNDQYGQHSFMANRMSNTSTRAI      1349
tr|G9HWE8|G9HWE8_9MONO              ADREKLLIPLLNSRVNLDYQTVLNFLPTHYSGNIVHRYNDQYGQHSFMANRMSNTSTRAI      1345
tr|Q1PD47|Q1PD47_MABVA              ADREKLLIPLLNSRVNLDYQTVLNFLPTHYSGNIVHRYNDQYGQHSFMANRMSNTSTRAI      1351
sp|P31352|L_MABVM                   ADREKLLIPLLNSRVNLDYQTVLNFLPTHYSGNIVHRYNDQYGQHSFMANRMSNTSTRAI      1351
tr|A0ACD3VMI8|A0ACD3VMI8_9MONO      SQAEDIIRPFCEARINLPVQELFKLLPSHYSGNIVHRYNDQYGPRSFMANRMSNTATRII      1290
sp|Q5XX01|L_EBOSU                   ANSDQLIRPFLEARVNLSVQEILQMTPSHYSGNIVHRYNDQYSPHSFMANRMSNTATRLM      1292
tr|A0A343EQF6|A0A343EQF6_9MONO      ANSDLLIKPFLEARVNLSVEEILQMTPSHYSGNIVHRYNDQYSPHSFMANRMSNSATRLI      1293
sp|Q05318|L_EBOZM                   SNSDLLIKPFLEARVNLSVQEILQMTPSHYSGNIVHRYNDQYSPHSFMANRMSNSATRLI      1293
tr|B8XCP4|B8XCP4_9MONO              ANSDLLVKPFIEARVNLSVQEILQMTPSHYSGNIVHRYNDQYSPHSFMANRMSNSATRLV      1293
tr|R4QUH5|R4QUH5_9MONO              ANSDLLVKPFVEARVNLSVQEILQMTPSHYSGNIVHRYNDQYSPHSFMANRMSNSATRLV      1293
                                              : . * *:  : :  : *   * *: ** .*:    .  .. :.. :
```

```
                                    PRNTase ←          → CD
tr|A0A185KRV3|A0A185KRV3_9MONO      VKALNDLPYTPSTHPVPNYTE|VRDNRLIYDPQPILEFDELRLAIQQT----------KKV      1733
sp|O89344|L_HENDH                   VKEVADIGGVNAELPVPEYTE|VENNRLIYDPDPVSEIDCDRLSKQES----------KAR      1479
tr|G5CPP9|G5CPP9_9MONO              VKEVADIGGVNAELPVPEYTE|VENNRLIYDPDPVSEIDCDRLSKQES----------KAR      1479
tr|F4YH99|F4YH99_9MONO              VKEVADIGGVNAELPVPEYTE|VENNRLIYDPDPVSEIDCDRLSKQES----------KAR      1479
sp|Q997F0|L_NIPAV                   VKEVADVGQVDAELPIPEYTE|VDNNHLIYDPDPVSEIDCSRLSNQES----------KSR      1479
tr|A0A7L5MR62|A0A7L5MR62_NIPAV      VKEVADVGQVDAELPIPEYTE|VDNNHLIYDPDPVSEIDCSRLSNQES----------KSR      1479
tr|Q1PDA3|Q1PDA3_9MONO              TRHVPSEYLFFDKPLDVDLNK|YMDNELVYDNDPLCSGIKGRLGRVSRSTLSLSLNVSDIG      1461
tr|A0A0U2VU68|A0A0U2VU68_9MONO      TRHVPSEYLFFDKPLDVDLNK|YMDNELVYDNDPLCSGIKGRLGRVSRSTLSLSLNVSDIG      1476
sp|Q1PDC4|L_MABVR                   TRHVPSEYLFFDKPLDVDLNK|YMDNELVYDNDPLCSGIKGRLGRVSRSTLSLSLNVSDIG      1468
tr|G9HWE8|G9HWE8_9MONO              TRHVPSEYLFFDKPLDVDLNK|YMDNELVYDNDPLCSGIKGRLGRVSRSTLSLSLNVSDIG      1464
tr|Q1PD47|Q1PD47_MABVA              TRHVPSEYLFFDKPLDVDLNK|YMDNELVYDNDPLCSGIKGRLGRVSRSTLSLSLNVSDIG      1470
sp|P31352|L_MABVM                   TRHVPSEYLYFDKPLDVDLNK|YMDNELVYDNDPLCSGIKGRLGRVSRSTLSLSLNVSDIG      1470
tr|A0ACD3VMI8|A0ACD3VMI8_9MONO      TREVPSQYLTYKANFGMNLQK|YQSNELIYDKDPLRCGLTCRVSVENFFLKHHFKNSVEGE      1409
sp|Q5XX01|L_EBOSU                   TREVPAQYLTYTTTLNLDLSK|YRNNELIYDSDPLRGGLNCNLSIDSPLMKGPRLNIIEDD      1411
tr|A0A343EQF6|A0A343EQF6_9MONO      TREVPAQYLTYTTTLKLDLTR|YKDNELIYDNDPLRGGLNCNLTFDHPLFKGERLNIIEED      1412
sp|Q05318|L_EBOZM                   TREVPAQYLTYTSTLDLDLTR|YRENELIYDSNPLKGGLNCNISFDNPFFQGKRLNIIEDD      1412
tr|B8XCP4|B8XCP4_9MONO              TREVPAQYLVYTSTLPLDLTR|YRDNELIYDDNPLRGGLNCNLSFDNPLFKGQRLNIIEED      1412
tr|R4QUH5|R4QUH5_9MONO              TREVPAQYLTYTSTLSLDLTR|YRENELIYDNNPLKGGLNCNLSFDNPLFKGQRLNIIEED      1412
                                    .: :             :  .  .*.*:** :*:      .:               .
```

```
//End of the L protein sequences from all the three viruses
tr|A0A185KRV3|A0A185KRV3_9MONO      P---------      2501
sp|O89344|L_HENDH                   ----------      2244
tr|G5CPP9|G5CPP9_9MONO              ----------      2244
tr|F4YH99|F4YH99_9MONO              ----------      2244
sp|Q997F0|L_NIPAV                   ----------      2244
tr|A0A7L5MR62|A0A7L5MR62_NIPAV      ----------      2244
tr|Q1PDA3|Q1PDA3_9MONO              GFSRPK----      2351
tr|A0A0U2VU68|A0A0U2VU68_9MONO      GFSRPK----      2366
sp|Q1PDC4|L_MABVR                   GFSRPK----      2358
tr|G9HWE8|G9HWE8_9MONO              GLFKSESLRA      2358
tr|Q1PD47|Q1PD47_MABVA              GLFKSGSLRA      2364
sp|P31352|L_MABVM                   GLFKSGSLRA      2364
tr|A0ACD3VMI8|A0ACD3VMI8_9MONO      GL--------      2196
sp|Q5XX01|L_EBOSU                   GFRSSSV---      2210
tr|A0A343EQF6|A0A343EQF6_9MONO      GFQNPP----      2210
sp|Q05318|L_EBOZM                   GLYRFD----      2212
tr|B8XCP4|B8XCP4_9MONO              GFFR------      2210
tr|R4QUH5|R4QUH5_9MONO              GINT------      2210
```



*The PRNTase domain is tentatively marked based on the X-ray crystallographic data of the Nipah viral L-protein

Figure 5 'Mix and Match' MSA of the PRNTase regions of the L-proteins of the Marburg, Ebola and Nipah and their related viruses

3.5. PR Exonuclease Active Site Structures from the Viral, Prokaryotic, and Eukaryotic DNA/RNA Polymerases

Table 1 shows the PR exonuclease active site structures from viral, prokaryotic, and eukaryotic DNA/RNA polymerases. It is clear from the Table that the PR exonuclease active site is not only completely conserved, but also adapted in all the organisms from viruses to humans. The DEDD superfamily either uses an Y or a H as the proton acceptor from a metal-bound water molecule to initiate the reaction [20].

Table 1 DEDD superfamily of PR exonuclease active site structures from viral, prokaryotic, and eukaryotic DNA/RNA polymerases.

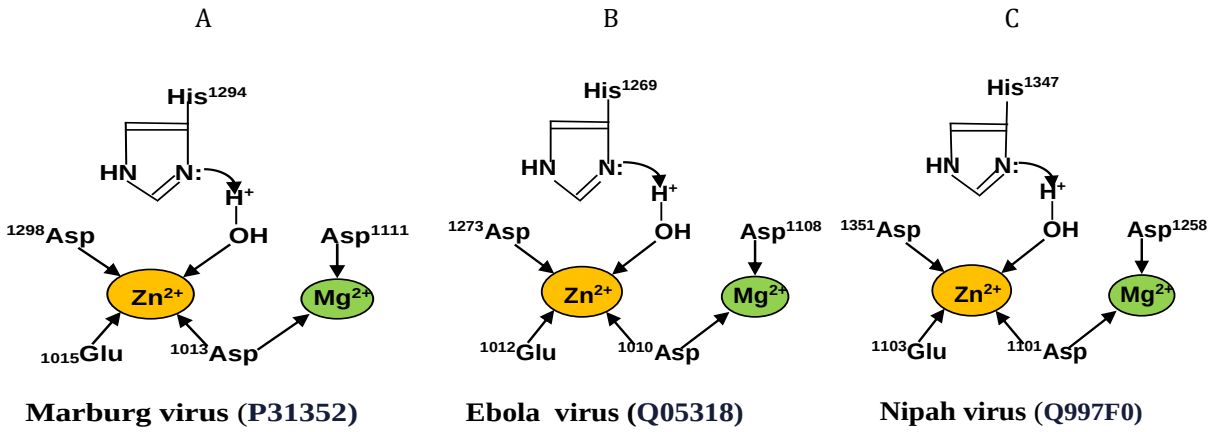
Organism	PR exonuclease active site am
1. Prokaryotic DEDD family PR Exonuclease Active Site Structure	
E. coli DNA Pol I (MFE)	-D^{355}T^{357}-----Y^{424}-----Y^{497}-----D^{501}-
E. coli DNA Pol II (MFE)	-D^{156}I^{158}------FD^{229}-----Y^{331}----D^{335}-
E. coli DNA Pol III (Replicase, ε-subunit)	-D^{12}T^{14}-------FD^{103}-----H^{162}----D^{167}-
Tth DNA pol III (Replicase, ε-subunit)	-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, Mitochondria/Chplastic)	-D^{234}TE^{236}------FD^{369}-------Y^{470}----D^{474}-
Plants (*A. thaliana* DNA pol IB, Mitochondria/Chplastic)	-D^{275}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, Chpalst)	-D^{542}VE----------LD^{577}F--------NH^{666}L---D^{599}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 (DNA pol)	-D^{112}I^{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}-
(+ strand RNA) SARS-CoV-1 (#NSP14-Exon)*	-D^{90}VE^{92}------FD^{191}--------H^{268}----D^{273}-
(+ strand RNA) MERS-CoV (#NSP14-Exon)*	-D^{90}VE^{92}------FD^{191}--------H^{268}----D^{273}-
(+ strand RNA) SARS-CoV-2 (#NSP14-Exon)*	-D^{90}VE^{92}------FD^{191}--------H^{268}----D^{273}-
(+ strand RNA) Hepatitis A Virus (RdRp)	-D^{233}LE-------FD^{363}--------Y^{483}---D^{487}-
(- strand RNA) Human Influenza Virus A: H1N1 (PA subunit)	-D^{347}IE---------LD^{425}-------H^{510}---D^{514}-
(- strand RNA) Human Influenza Virus B (PA subunit)	-D^{194}IE---------LD^{420}-------H^{506}---D^{510}-
(- strand RNA) Human Influenza Virus C (PA subunit)	-D^{84}LE---------ID^{288}--------H^{494}---D^{498}-
(- strand RNA) Human Respiratory Syncytial Virus (RdRp)	-D^{664}LE--------LD^{794}-------H^{829}---D^{833}-
(- strand RNA) Marburg Virus (PRNTase)	-D^{1013}LE-------LD^{1096}------H^{1294}R---D^{1298}-
(- strand RNA) Ebola Virus (PRNTase)	-D^{1010}FE-------VD^{1108}------H^{1269}R---D^{1273}-
(- strand RNA) Nipah Virus (PRNTase)	-D^{1101}LE-------LD^{1258}-------H^{1347}R---D^{1351}-

Adapted from Palanivelu [21].

Tth, Thermus thermophilus; NE, Nuclear encoded; Chplast, Chloroplast; Mito, Mitochondria.
Eukaryotic replicative α polymerases (DNA primases) lack the PR exo active site.
The confirmed PR exo active site amino acids by SDM experiments and X-ray Crystallography data are highlighted in dark blue/light blue, respectively.
*All possess the same PR exonuclease domain and their active site amino acids are completely conserved in Bat-RaTG13, Pangolin, Civet, SARS-CoV-1, MERS-CoV, SARS-CoV2 and Human CoV-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 the life cycle of Coronaviruses.

3.5.1. Proposed PR exonuclease active site structures in Marburg, Ebola and Nipah viruses

As discussed elsewhere, the DEDD superfamily of 3'→5' PR exonucleolytic activity is essential for faithful replication of the viral RNA genomes in the non-segmented, negative-strand viruses. It does not exist as an intrinsic domain or as an auxiliary subunit as reported for most of the replicative DNA polymerases of prokaryotes and eukaryotes [22]. In contrast, in these viruses, the 3'→5' PR exonuclease is placed along with the modifying enzymes like methyltransferases in the positive-strand viruses, like SARS-CoVs and MERS-CoV, or along with the PRNTase in the negative-strand viruses like of Marburg, Ebola, Nipah, Hendra, Cedar, etc. Figures 6A, B and C show the proposed active site amino acids in the 3'→5' PR exonuclease active sites of Marburg, Ebola and Nipah viruses which interestingly follow the same structural similarity to the already reported in a large number of 3'→5' PR exonucleases in DNA/RNA polymerases, both as MFEs or MECs [22]. However, some replication errors escape these stringent repair mechanisms like the PR exonucleases (that act during the replication process) and mismatch repair system (that operates after the replication process) and thus, become permanent mutations. Such mutation(s) could be advantageous (maintained) or deleterious (eliminated) to the viruses. In fact, the advantageous mutations serve as the basis for the evolution of new variants. A general mechanism of action for the PR exonuclease is already reported by the author [20]. The proposed two-metal ion in the active site in the viral 3'→5' PR exonuclease is based on Beese and Steitz's observation, who described an analogous active site architecture in the exonuclease domain of *E. coli* DNA polymerase I [23].



Figures 6 A, B & C. Proposed amino acids at the PR exonuclease active sites in the PRNTase regions of Marburg (A), Ebola (B) and Nipah (C) viruses

4. Conclusions

These findings demonstrate for the first time that these non-segmented, negative-strand RNA viruses also possess a typical PR mechanism to maintain the integrity of their genomes and their identities. This discovery provides a strong rationale for developing antivirals targeting the PR exonuclease domain as a potential strategy to control the spread of these deadly human viruses. Furthermore, the complete conservation of the DEDD superfamily of PR exonucleases from viruses to humans underscores the evolutionary significance of this catalytic domain. Moreover, these viruses utilise an unconventional capping enzyme, the PRNTase, which is structurally and mechanistically distinct from the host mRNA capping machinery. This difference presents an attractive and selective target for antiviral drug development against these viruses. The widespread distribution of reservoir hosts, coupled with increasing deforestation, multiple transmission routes, high case fatality rates, and the absence of effective therapeutics and vaccines, place these viruses, such as Marburg, Ebola, and Nipah, among those with significant pandemic potential. Additionally, recurring outbreaks and their high mortality rates continue to pose a major global public health challenge.

Compliance with ethical standards

Acknowledgments

The author wishes to thank Dr. P. Sathiyamoorthy, Chief Scientist, GEM Research Foundation, Chennai, for useful suggestions on the manuscript.

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

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

[2] 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.

[3] Amarasinghe GK, Aréchiga Ceballos NG, Banyard AC, Basler CF, Bavari S, Bennett AJ, et al. "Taxonomy of the order Mononegavirales: Arch Virol. 2018;163: 2283–2294.

[4] Kleiner VA, Fearns R. How does the polymerase of non-segmented negative strand RNA viruses commit to transcription or genome replication? J. Virol. 2014; 98: e0033224. https://doi.org/10.1128/jvi. 00332-24.

[5] Palanivelu, P. RNA-dependent RNA Polymerases of Severe Acute Respiratory Syndrome-Related Coronaviruses-An Insight into their Active Sites and Mechanism of Action. Int J Biochem Res Rev. 2020; 29: 29-52.

[6] Sanjuán R, Nebot MR, Chirico N, Mansky LM, Belshaw R. Viral mutation rates. J Virol. 84; 2010: 9733–9748.

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

[8] 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.

[9] Palanivelu P. An overview of the proofreading functions in bacteria and SARS-Coronaviruses. Int J Biochem Res Rev. 30(3); 33-62 (2021).

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

[11] 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.

[12] 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.

[13] 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.

[14] Ogino M, Ito N, Sugiyama M, Ogino T. The rabies virus L protein catalyzes mRNA capping with GDP polyribonucleotidyltransferase activity. Viruses. 2016; 8:144.

[15] Shuman S. Structure, mechanism, and evolution of the mRNA capping apparatus. Prog Nucleic Acid Res. Mol Biol. 2001; 66:1–40.

[16] Neubauer J, Ogino M, Green TJ, and Ogino T. 2016. Signature motifs of GDP polyribonucleotidyltransferase, a non-segmented negative strand RNA viral mRNA capping enzyme, domain in the L protein are required for covalent enzyme–pRNA intermediate formation. Nucleic Acids Res. *2016;* 44:330–341.

[17] Ogino T, Yadav SP, Banerjee AK. Histidine-mediated RNA transfer to GDP for unique mRNA capping by vesicular stomatitis virus RNA polymerase. Proc Natl Acad Sci (USA). 2010; 107: 3463 3468.

[18] Li J, Rahmeh A, Morelli M, Whelan SP. A conserved motif in region v of the large polymerase proteins of non-segmented negative-sense RNA viruses that is essential for mRNA capping. J. Virol. 2008; 82: 775–784.

[19] Park GJ, Osinski A, Hernandez G, Eitson JL, Majumdar A, Tonelli M, Henzler-Wildman K, et al. The mechanism of RNA capping by SARS-CoV-2. Nature. 2022; 609:793-800.

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

[21] Palanivelu P. An insight into the replication mechanism of smallpox and monkeypox viruses. World J Adv Res Rev. 2024; 23: 2239-2260.

[22] 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.

[23] 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.