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 HNH endonuclease domain in the basic protein 2 subunit of the polymerase of human influenza viruses

Peramachi Palanivelu *

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

World Journal of Advanced Research and Reviews, 2022, 15(03), 425–442

Publication history: Received on 16 August 2022; revised on 21 September 2022; accepted on 23 September 2022

Article DOI: https://doi.org/10.30574/wjarr.2022.15.3.0944

Abstract

RNA polymerases of human influenza viruses A, B and C do not have a capping enzyme, as other RNA viruses, to cap their mRNAs for translation in the host cells. So, they employ a unique 'cap-snatching' mechanism, where the mRNA cap structures are snatched from the host cell mRNAs and used as a primer to initiate its own mRNA synthesis. One of the RNA polymerase subunits, the polymerase basic protein subunit 2 (PB2), is shown to involve in the 'cap-snatching' mechanism. The active sites for the 'cap-snatching' and endonuclease by the PB2 were analyzed by multiple sequence alignment (MSA) analysis and corroborated with the results available from biochemical, site-directed mutagenesis (SDM) and X-ray crystallographic techniques. It is found that the PB2 subunit in all three human influenza viruses habours both the cap-binding motif (CBM) and a HNH/N type endonuclease domain. The CBM is aromatic amino acid rich and the HNH/N is a –DH- based endonuclease in influenza viruses A and C and a –DQ- based one in influenza virus B. The invariant H is proposed to act as a general base to initiate catalysis and the invariant first N is implicated in nucleotide binding. In addition, the nuclear localization signals were also identified in all three human influenza viruses. By sequence similarity, similar HNH/N domain are found in the RNA cleaving CRISPR-Cas13a/13b and CRISPR-Cas12a endoribonucleases. The identification of HNH/N domain in all three human influenza viruses suggests that the PB2 subunit itself could cleave the cap structures from the host cell mRNAs, which are subsequently used as primers to initiate viral mRNA synthesis. These results will facilitate the optimization of endonuclease inhibitors as potential new anti-influenza drugs, and could also help in developing new drugs for flu treatments in the future.

Keywords: Human influenza viruses; RNA Polymerase basic protein subunit 2; HNH endonuclease; Cap-snatching; CRISPR-Cas13a endoribonuclease; CRISPR-Cas12a endoribonuclease.

1. Introduction

Influenza is an infectious disease caused by influenza viruses. Influenza viral infection is a highly contagious, airborne that generally causes acute respiratory illness resulting in variable degrees of systemic symptoms from mild fatigue to respiratory failure and death. Therefore, influenza viral infections are a major public health concern worldwide. The A and B viruses cause seasonal influenza every year, which affect approximately 5–10% of the adult and 20–30% of the pediatric population. A report by World Health Organization estimates that seasonal flu may result in 2,90,000-6,50,000 deaths every year due to respiratory diseases alone. (This estimate does not take into account of deaths from other diseases such as cardiovascular disease, which can be influenza-related). The first flu viral pandemic took place in 1918 and was referred to as "Spanish" flu. It was the deadliest, claiming an estimated ~50 million lives worldwide in less than a year [1]. The second flu epidemic was reported in 2019-2020. The Center for Disease Control and Prevention (CDC), USA, estimates, the 2019–2020 had caused 39 million to 56 million flu illnesses and 24,000 to 62,000 deaths and found it was mainly caused by the influenza A virus, H1N1. Only few other pandemics have caused such a devastating effect,

* Corresponding author: Peramachi Palanivelu
Department of Molecular Microbiology, School of Biotechnology, Madurai Kamaraj University, Madurai – 625 021, India.

Copyright © 2022 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution Liscense 4.0.

like the ones which were caused by the SARS-Coronaviruses (SARS-CoVs) including the current epidemic caused by SARS-CoV-2.

Influenza viruses belong to the family, Orthomyxoviridae. In this family, there are four genera which are named A, B, C and D. Influenza A, B and C viruses infect humans whereas D virus infects animals. Among them, infections by influenza A virus is most common and severe and causes the flu pandemics. Influenza A virus is an avian virus, which is capable of jumping species and infects both humans and other animals. Wild aquatic birds and other animal species like, pigs, ferrets, horses, seals, whales, minks, giant anteaters, cats and dogs are also found to be the reservoir for the influenza A virus. While influenza A virus shows animal to human transmissions, influenza B and C viruses show very limited host range and appear predominantly in humans [2]. Interestingly, influenza viruses, B and C, do not have animal reservoirs like influenza A virus [3]. Out of the three most widely known human influenza viruses, the influenza A and B viral infections peak in winter months, causing substantial morbidity and mortality in humans and a considerable financial burden worldwide, whereas the influenza C virus [4] cause sporadic outbreaks in humans causing only mild upper respiratory infections [5, 6]. Influenza D virus primarily affects cattle and is not known to infect or cause any major illness in humans [3].

1.1. Human Transmission

Influenza viral infections are mediated by viral surface glycoproteins. Types A and B have two surface glycoproteins, viz. haemagglutinin (H) and neuraminidase (N). (H is a Type I integral membrane glycoprotein that binds to cell-surface receptors and facilitates fusion between the viral envelope and endosomal membrane, whereas N is a Type II integral membrane glycoprotein, which facilitates viral release from cells by removing sialic acid from sialyloligosaccharides on the cell and viral surfaces). The H of the influenza viruses is a trimer (similar to the spike proteins in SARS-CoVs) that recognizes and binds to the host cell-surface receptor, N-acetylneuraminic (sialic) acid. (Sialic acids are nine-carbon acidic monosaccharides commonly found at the termini of many glycoconjugates), which is followed by the release of the virions into the human cells by the sialidase activity of the N. On the other hand, the types C and D have only one surface glycoprotein, known as Hemagglutinin–Esterase Fusion (HEF) protein, which binds to the host cell-surface receptor (9-O-acetylneuraminic acid) and promotes viral membrane and cell membrane fusion. In C and D types, the neuraminate-O-acetyl esterase functions as neuraminidase in viral release from the infected cells. There are at least 18 haemagglutinins (H1 to H18) and 11 neuraminidases (N1 to N11) reported in influenza A virus, whereas the influenza B and C viruses do not have such subtypes. Although vaccines against influenza A and B viruses are available, the protection that they offer is limited by the antigenic variations mainly in the two glycoproteins, viz. H and N, which are found on the envelope of influenza A viral strains at 4:1 ratio [7].

1.2. Genomes of Human Influenza Viruses

Human influenza A, B and C viruses belong to the Orthomyxoviridae family, which are a group of enveloped, negative-sense RNA viruses, characterized by their segmented genome. Influenza viruses A and B are made up of 8 RNA segments and B type showed no antigenic relationship to A type virus. RNA segments are joined by the RNA polymerase subunits (PB1, PB2 and PA) and the nuclearcapsid protein, NP. Influenza A viral genome length is 13,588 nts, and that of B is 14,639 nts. This is due to the 5'- and 3'-UTRs are longer in B viral genome. In influenza viruses A and B, the 8 RNA segments encode at least 10 proteins. Influenza viruses C and D genome consists of only 7 segments and encode only 9 proteins [8]. In all the four genera, the viral RNA (vRNA) genome segments are bound by a heterotrimeric RNA-dependent RNA polymerase (RdRp), forming a viral ribonucleoprotein (vRNP) complex [9,10]. In the vRNP, the 5'- and 3'-termini of the vRNA are bound to the RNA polymerase, while the rest of the vRNA associates with oligomeric NPs. Among them, only influenza A virus has many subtypes with different combinations of the two surface glycoproteins.

1.2.1. Transcription and Replication Processes in Influenza Viruses

As they are negative-strand viruses, the viral transcription and replication processes are performed in the host cell nucleus and then transported to the cytoplasm for translation and viral assembly. The crucial enzyme for the multiplication of the viruses in the host cells, is the RNA-dependent RNA polymerases (RdRp), which is a component of the Ribonucleoprotein (RNP) that is imported into the host cell nucleus during the infection process. The viral polymerase is a heterotrimer and made up of three different subunits, viz. two polymerase basic protein subunits (PB1 and PB2) and a polymerase acidic protein subunit (PA) (Fig. 1a). Unlike other RNA viruses, the influenza virus polymerases do not have an inherent capping enzyme to cap its mRNAs for efficient translation in the host cells and therefore, it relies upon the capped pre-mRNAs of the host cells as cap-donors [11]. Thus, during multiplication, the enzyme uses the negative-sense viral RNA (vRNA) as a template to synthesize two positive-sense RNA species, viz. mRNA templates for viral protein synthesis, which are after polyadenylation and capping exported to cytoplasm and translated like other host mRNAs, and complementary RNA (cRNA) intermediates from which the same enzyme

transcribes more copies of negative-sense, genomic vRNAs. In contrast to transcription, the genome replication is cap-independent and proceeds via a cRNA replicative intermediate. Thus, genome replication is unprimed and generates exact full-length copies of the template.

The X-ray crystallographic structures of the complete heterotrimeric polymerase (M_r ~255 kDA) revealed that the polymerase forms a compact particle with PB1 at its centre, capped on one face by PB2 and clamped between the two globular domains of PA. Like other DNA and RNA polymerases, the PB1 had the canonical right-hand-like fold, possessing fingers, palm, and thumb subdomains [12, 13].

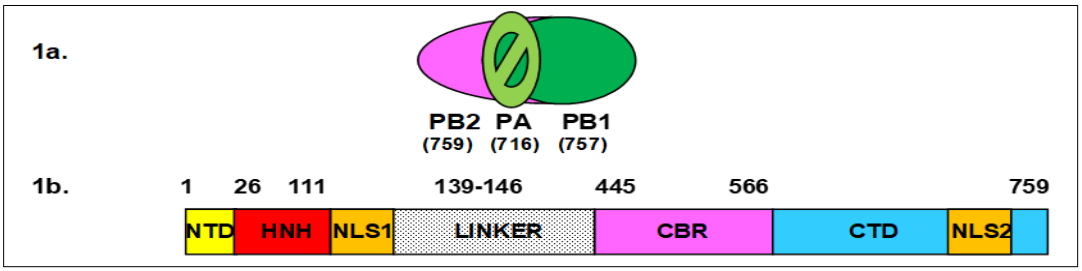


Figure 1a. A schematic diagramme showing various subunits of the heterotrimeric influenza viral polymerase. (The number of amino acids of the subunits is given in brackets from influenza A virus)

The influenza viral RNA polymerase (EC 2.7.7.48) is a primer-dependent enzyme. The enzyme cannot copy the (-) strand RNA template without a small piece of RNA that aligns on the template RNA and provides a starting point for mRNA synthesis. They generate the primers by a unique 'cap-snatching' mechanism. As RdRps are not found in mammalian cells, they are an excellent target for designing antiviral compounds.

Figure 1b. A schematic diagram showing various domains of the polymerase PB2 subunit of influenza virus A (CBR, Cap-binding region)

As the virus's initial stages of multiplication are performed in the nucleus, they possess nuclear localization signals (NLSs). The NLSs are generally short peptides rich in basic amino acids that act as a signal to mediate the transport of molecules from the cytoplasm into the nucleus. The NLSs are recognized by the corresponding nuclear transporters, which then interact with nucleoporins to help NLS-containing proteins reach the nucleus through Nuclear Pore Complexes.

A K/R rich NLS is found at the C-terminal region and another possible one at the N-terminal region (highlighted in orange). The first NLS to be discovered was the sequence –^{126}PKKKRKV^{132}- in the SV-40 large T-antigen (a monopartite NLS) [14]. Willis et al. [15] identified a slightly different putative NLS (-^{640}PKLKRQ^{646}-) in vasopressin-activated calcium-mobilizing protein which is similar to the N-terminal one in PB2. In recent years, NLSs are also widely used as targets in cancer treatment and prevention of viral infections [16].

Most of the cellular and eukaryotic viral mRNAs have a cap structure at their 5'-end that is critical for efficient translation. Cap structures also help in mRNA transport from nucleus to cytoplasm and, in addition, protect the mRNAs from degradation by 5'-exonucleases. A common aromatic-rich cap-binding pocket is conserved by convergent evolution. In this pocket the positively charged N(7)-methylated guanine ring of the cap structure is stacked between two aromatic amino acid residues. In the process called 'cap-snatching', the viral polymerase uses its PB2 cap-binding domain to capture the 5′-cap of nascently capped host mRNAs and cleaves and the 5'-cap of ~15 nucleotides which is used as the primer for viral mRNA synthesis. Most attempts to develop anti-influenza drugs against the influenza viruses are for the RdRp and are focused on the highly conserved PB1 polymerase catalytic subunit's active site, the PB2 cap-binding and PA endonuclease domains [17].

Transcription is initiated by adding **GMP**, the nucleotide which is complementary to the second nucleotide at the 3'-termini of all eight RNA segments, to the primer. The cap structure (highlighted in blue), snatched from the host mRNAs, is used as the primer for viral mRNA transcription (Fig. 2).

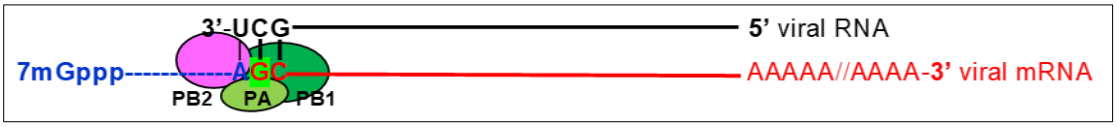


Figure 2 Cap structure of the host mRNAs with the viral polymerase subunits during viral mRNA transcription [11].

There are two different reports on the activities of the polymerase subunits, viz. the PB2 and PA. For example, some reports claim that the cap-binding and endonuclease activities reside on two different subunits, PB2 and PA, respectively [18-21]. Some reports claim that the cap-binding and endonuclease activities reside on the PB2 subunit itself. In this communication, evidences are presented to show that a HNH/N type endonuclease domain is identified in the PB2 subunits of all three human influenza viruses, which not only bind to the cap structure but could also make the endonucleolytic cleavage [21,11].

1.3. HNH Endonucleases

The HNH domains are ubiquitous and reported both in prokaryotes and eukaryotes. All of them invariably harbour an endonuclease site. The HNH domains are small nucleic acid binding motifs (~30 amino acids in length) with an associated DNA cleavage module. Such modules are commonly widespread in α-α-β-metal finger endonucleases. Proteins containing the HNH domain performs variety of functions in the organisms like site-specific group I and group II homing endonucleases, bacterial toxins with non-specific endonuclease activities, restriction enzymes, reverse transcriptase, etc. Their properties and mechanism of action are already discussed in detail by Palanivelu [22].

Based on the active site amino acid analysis, HNH-family of endonucleases are broadly classified into two major subgroups, depending upon the immediate amino acid adjacent to the invariant proton acceptor (His), highlighted in green, i.e., either as –HH- based or –DH- based enzymes (highlighted in red) [22]. (A few exceptions are also observed where the D is replaced by E and the second H is replaced by an N). A typical arrangement of active site amino acids in HNH endonucleases is shown in Fig. 3.

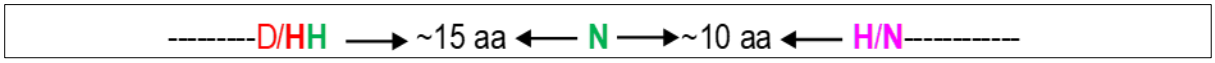


Figure 3 Arrangement and distance conservations active site amino acids in HNH/N endonucleases (Shows the invariant amino acids (green and bold) and the approximate distances between them**.**

The active site of HNH endonucleases consists of two highly conserved His and Asn, and a variable His (Asn replaces the second H and form another subfamily of HNH endonucleases). The HNH endonucleases can make a nick or a double-stranded break on specific/nonspecific regions on DNA molecules in the presence of a divalent metal ion. The invariant **His** residue in the conserved motif **H**NH/N serves as the general base that activates a metal-bound water molecule for a nucleophilic attack on the sugar-phosphate backbone of nucleic acids [23,24].

2. Material and methods

The protein sequence data of the polymerase basic protein subunit, PB2 of influenza viruses, A, B and C were obtained from PUBMED and SWISS-PROT databases. The advanced version of Clustal Omega was used for protein sequence analysis. Along with the conserved motifs identified by the bioinformatics analysis and from the data already available from biochemical, SDM and X-ray crystallographic analyses on the PB2 subunit and HNH endonucleases were used to arrive at the possible amino acids that make the active site of the enzyme. For pI calculations, the Expasy tool was used.

3. Results and discussion

3.1. MSA analysis of the PB2 subunit of the polymerase

Despite the availability of antivirals for influenza viruses, the emergence of resistant strains calls for antivirals with novel mechanisms of actions. The PB2 subunit of the influenza A viral polymerase is a promising drug target because of its vital role in the unique 'cap-snatching' mechanism. In fact, blocking the influenza virus "cap-snatching" activity was proved to be a new and efficient strategy for the treatment of influenza viral infections. However, all three subunits PB1 [25], PB2 [21] and PA [18, 20] are implicated in endonuclease activity of the polymerase heterotrimer.

Shi et al. [21] have shown that the endonuclease activity resides in the PB2 subunit of the polymerase in influenza A virus. Their results clearly showed that purified virion RNP-complexes cleaved the RNA specifically to generate a capped 14-nt RNA fragment (cap+13 nt) to be further used as primer to initiate viral mRNA synthesis (Fig. 2). Furthermore, they found that the purified anti-PB2 IgG, inhibited the endonuclease activity, but interestingly, anti-PB1 and anti-PA antibodies did not inhibit the cleavage. They have also further found that RNAs containing the 5'-terminal structure, the Gppp----------G/A, could not be cleaved to produce these specific fragments in the absence of one or two subunits of the polymerase. The presence of the endonucleolytic domain in the PB2 subunit was further supported by Plotch et al. [11].

They found that the purified trimeric complex, expressed by recombinant baculovirus in insect cells, cleaved the artificial substrate, but if one or two subunits were removed from the complex, the cleavage activity was totally lost. Therefore, they suggested that the viral PB2 is the endonuclease that cleaves the host cell mRNA to produce the primer to initiate viral transcription. Furthermore, they found the uncapped ribopolymer inhibitors of viral mRNA transcription inhibited the cleavage of capped RNAs [11]. UV cross-linking studies and photo-affinity labelling by cap analogues have shown that the PB2 subunit recognizes and binds to the cap structure at the 5'-end of the host cell hnRNAs [26]. By using temperature-sensitive mutants with defects in the PA gene, it was shown that the principal role of the PA subunit is not in the viral mRNA synthesis, but rather in viral RNA replication [27, 28]. However, it was shown that the PA subunit is required for successful assembly of an active polymerase complex with PB1 and PB2 [29].

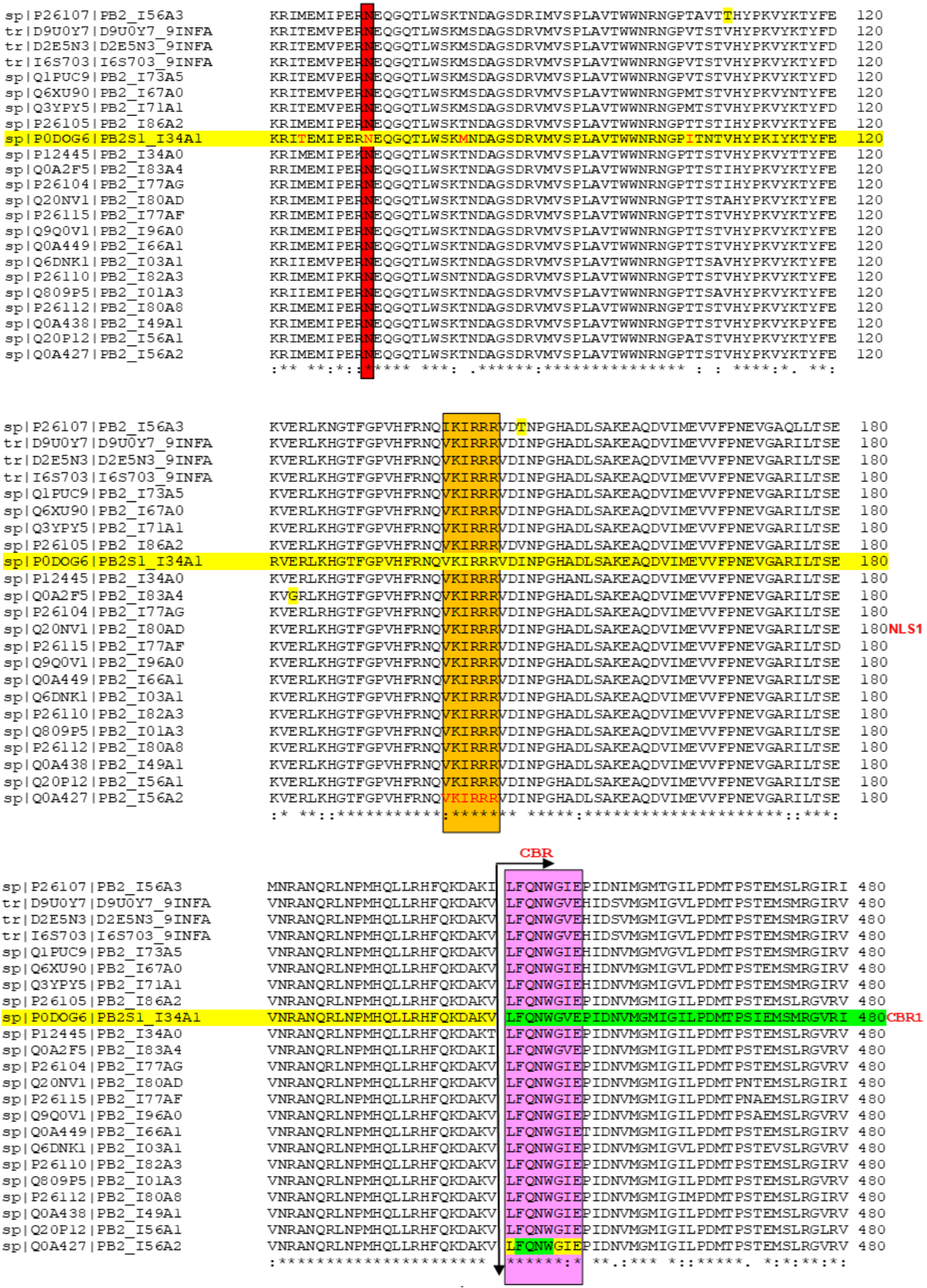
The present work found that the PB2 subunit not only possesses the cap-binding domain but also an endonuclease domain which possibly involves in cleaving the cap structure. For cleaving the cap structure from the host pre-mRNAs, a HNH type of endonuclease active site is identified in the PB2 subunits itself for the first time, supporting the observations by Plotch et al. [11] and Shi et al. (21].

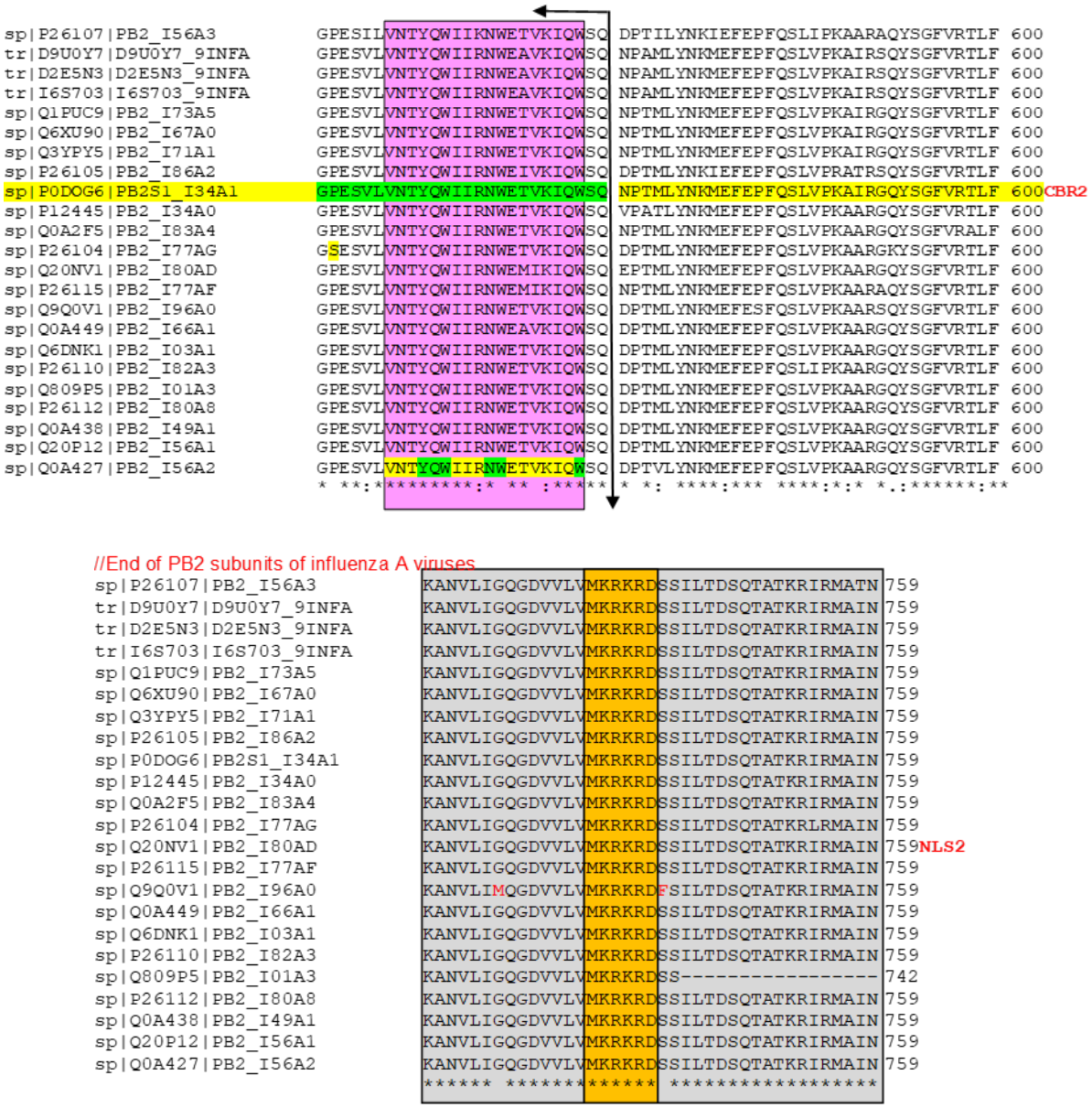
Figure 4 shows the MSA of PB2 subunits of the polymerase from different strains of influenza A virus. (only the required regions for the discussions are shown here). The influenza A virus strain (1934/ H1N1) is highlighted and it showed a theoretical pI of 9.73. The MSA shows that the entire sequence is highly conserved irrespective of the serotypes. Further analysis of the MSA found, a –DH- based HNH endonuclease domain in the PB2 subunit of polymerase at the N-terminal region. The –DH- dyad is followed by two completely conserved Ns, suggesting it belongs to HNN subfamily (Fig. 4). Two metal-binding –DxD- motifs are identified (data not shown).

Proposed NLSs and cap-binding regions (CBRs) are highlighted in orange and magenta, respectively. Several studies have shown that the CBR is located in the PB2 subunit. Honda et al. [30] identified a ^{32}P-labelled, V8 protease peptides of PB2 derived by UV cross-linking of the influenza ribonucleoprotein complex to a m7G^{32}ppp-labelled capped oligonucleotide. They suggested that residues 242–282 and a second region from 538–577, were involved in cap-binding. Li et al. [25] extended this approach by UV cross-linking a 4 thioU-containing, ^{32}P-labelled, capped oligonucleotide. A peptide, -SVLVNTYQWIIRNW- (residues 544–557) was identified after V8 protease digestion. Furthermore, mutation of W^{552}→Ala reduced cap binding to 25 % of wild-type levels. Given the proximity of the 4 thioU residue (at residue 2 of the oligonucleotide) to the labelled cap structure, it was proposed that the isolated peptide must be close to the aromatic sandwich. The authors concluded that one or other of the nearby aromatic residues, i.e., W537, Y550, W557 or W564, form the aromatic sandwich, similar to other cap-binding proteins (such as the eukaryotic initiation factor eIF4E or VP39) in which the 7-methylguanine moiety is sandwiched between two aromatic amino acid residues. Honda et al. [30] reported two regions of PB2 involved in cap binding, whereas Li et al. [25] reported one region only. MSA analysis shows at least two CBRs, one with the likely residues from 445 to 452 (-^{445}LFQNWGVE-) and the other one with the residues from 448 to 557 (-^{448}NTYQWIIRNW-). The second CBR is proved by SDM and other techniques [30, 25].

CLUSTAL O (1.2.4) MSA of PB2 subunit of polymerase of influenza A viruses

```
sp|P26107|PB2_I56A3          MERIKELRDLMSQSRTREILTTTVDHMAIIKRYTSGRQEKNPALRMKWMMAMKYPITAD  60
tr|D9U0Y7|D9U0Y7_9INFA       MERIKELRNLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPSLRMKWMMAMKYPITAD  60
tr|D2E5N3|D2E5N3_9INFA       MERIKELRNLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPSLRMKWMMAMKYPITAD  60
tr|I6S703|I6S703_9INFA       MERIKELRNLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPSLRMKWMMAMKYPITAD  60
sp|Q1PUC9|PB2_I73A5          MERIKELRNLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPSLRMKWMMAMKYPITAD  60
sp|Q6XU90|PB2_I67A0          MERIKELRNLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPSLRMKWMMAMKYPITAD  60
sp|Q3YPY5|PB2_I71A1          MERIKELRNLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPSLRMKWMMAMKYPITAD  60
sp|P26105|PB2_I86A2          MERIKELRDLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPALRMKWMMAMKYPITAD  60
sp|P0DOG6|PB2S1_I34A1        MERIKELRNLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPALRMKWMMAMKYPITAD  60
sp|P12445|PB2_I34A0          MERIKELRDLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPALRMKWMMAMKYPITAD  60
sp|Q0A2F5|PB2_I83A4          MERIKELRDLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPALRMKWMMAMKYPITAD  60
sp|P26104|PB2_I77AG          MERIKGLRDLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPALRMKWMMAMKYPITAD  60
sp|Q20NV1|PB2_I80AD          MERIKELRDLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPALRMKWMMAMKYPITAD  60
sp|P26115|PB2_I77AF          MERIKELRDLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPALRMKWMMAMKYPITAD  60
sp|Q9Q0V1|PB2_I96A0          MERIKELRDLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPALRMKWMMAMKYPITAD  60
sp|Q0A449|PB2_I66A1          MERIKELRDLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPALRMKWMMAMKYPITAD  60
sp|Q6DNK1|PB2_I03A1          MERIKELRDLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPALRMKWMMAMKYPITAD  60
sp|P26110|PB2_I82A3          MERIKELRDLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPALRMKWMMAMKYPITAD  60
sp|Q809P5|PB2_I01A3          MERIKELRDLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPALRMKWMMAMKYPITAD  60
sp|P26112|PB2_I80A8          MERIKELRDLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPALRMKWMMAMKYPITAD  60
sp|Q0A438|PB2_I49A1          MERIKELRDLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPALRMKWMMAMKYPITAD  60
sp|Q20P12|PB2_I56A1          MERIKELRDLLSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPALRMKWMMAMKYPITAD  60
sp|Q0A427|PB2_I56A2          MERIKELRDLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPALRMKWMMAMKYPITAD  60
                             ***** **:*:**********.**********:**********:****************
```


sp|P26107|PB2_I56A3 KRIMEMIPERNEQGQTLWSKTNDAGSDRIMVSPLAVTWWNRNGPTAVTTHYPKVYKTYFE 120
tr|D9U0Y7|D9U0Y7_9INFA KRITEMVPERNEQGQTLWSKMSDAGSDRVMVSPLAVTWWNRNGPVTSTVHYPKVYKTYFD 120
tr|D2E5N3|D2E5N3_9INFA KRITEMVPERNEQGQTLWSKMSDAGSDRVMVSPLAVTWWNRNGPVTSTVHYPKVYKTYFD 120
tr|I6S703|I6S703_9INFA KRITEMVPERNEQGQTLWSKMSDAGSDRVMVSPLAVTWWNRNGPVTSTVHYPKVYKTYFD 120
sp|Q1PUC9|PB2_I73A5 KRITEMVPERNEQGQTLWSKMSDAGSDRVMVSPLAVTWWNRNGPVTSTVHYPKVYKTYFD 120
sp|Q6XU90|PB2_I67A0 KRITEMVPERNEQGQTLWSKMSDAGSDRVMVSPLAVTWWNRNGPMTSTVHYPKVYNTYFE 120
sp|Q3YPY5|PB2_I71A1 KRITEMVPERNEQGQTLWSKMSDAGSDRVMVSPLAVTWWNRNGPMTSTVHYPKVYKTYFD 120
sp|P26105|PB2_I86A2 KRIMEMIPERNEQGQTLWSKTNDAGSDRVMVSPLAVTWWNRNGPTTSTIHYPKVYKTYFE 120
sp|P0DOG6|PB2S1_I34A1 KRITEMIPERNEQGQTLWSKMNDAGSDRVMVSPLAVTWWNRNGPITNTVHYPKIYKTYFE 120
sp|P12445|PB2_I34A0 KRIMEMIPERNEQGQTLWSKTNDAGSDRVMVSPLAVTWWNRNGPTTSTVHYPKVYTTYFE 120
sp|Q0A2F5|PB2_I83A4 RRIMEMIPERNEQGQILWSKTNDAGSDRVMVSPLAVTWWNRNGPTTSTIHYPKVYKTYFE 120
sp|P26104|PB2_I77AG KRIMEMIPERNEQGQTLWSKTNDAGSDRVMVSPLAVTWWNRNGPTTSTIHYPKVYKTYFE 120
sp|Q20NV1|PB2_I80AD KRIMEMIPERNEQGQTLWSKTNDAGSDRVMVSPLAVTWWNRNGPTTSTAHYPKVYKTYFE 120
sp|P26115|PB2_I77AF KRIMEMIPERNEQGQTLWSKTNDAGSDRVMVSPLAVTWWNRNGPTTSTVHYPKVYKTYFE 120
sp|Q9Q0V1|PB2_I96A0 KRIMEMIPERNEQGQTLWSKTNDAGSDRVMVSPLAVTWWNRNGPTTSTVHYPKVYKTYFE 120
sp|Q0A449|PB2_I66A1 KRIMEMIPERNEQGQTLWSKTNDAGSDRVMVSPLAVTWWNRNGPTTSTVHYPKVYKTYFE 120
sp|Q6DNK1|PB2_I03A1 KRIIEMVPERNEQGQTLWSKTNDAGSDRVMVSPLAVTWWNRNGPTTSAVHYPKVYKTYFE 120
sp|P26110|PB2_I82A3 KRIMEMIPKRNEQGQTLWSNTNDAGSDRVMVSPLAVTWWNRNGPTTSTVHYPKVYKTYFE 120
sp|Q809P5|PB2_I01A3 KRIIEMIPERNEQGQTLWSKTNDAGSDRVMVSPLAVTWWNRNGPTTSAVHYPKVYKTYFE 120
sp|P26112|PB2_I80A8 KRIMEMIPERNEQGQTLWSKTNDAGSDRVMVSPLAVTWWNRNGPTTSTVHYPKVYKTYFE 120
sp|Q0A438|PB2_I49A1 KRIMEMIPERNEQGQTLWSKTNDAGSDRVMVSPLAVTWWNRNGPTTSTVHYPKVYKPYFE 120
sp|Q20P12|PB2_I56A1 KRIMEMIPERNEQGQTLWSKTNDAGSDRVMVSPLAVTWWNRNGPATSTVHYPKVYKTYFE 120
sp|Q0A427|PB2_I56A2 KRIMEMIPERNEQGQTLWSKTNDAGSDRVMVSPLAVTWWNRNGPTTSTVHYPKVYKTYFE 120
sp|P26107|PB2_I56A3 KVERLKNGTFGPVHFRNQIKIRRRVDTNPGHADLSAKEAQDVIMEVVFPNEVGAQLLTSE 180
tr|D9U0Y7|D9U0Y7_9INFA KVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSE 180
tr|D2E5N3|D2E5N3_9INFA KVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSE 180
tr|I6S703|I6S703_9INFA KVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSE 180
sp|Q1PUC9|PB2_I73A5 KVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSE 180
sp|Q6XU90|PB2_I67A0 KVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSE 180
sp|Q3YPY5|PB2_I71A1 KVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSE 180
sp|P26105|PB2_I86A2 KVERLKHGTFGPVHFRNQVKIRRRVDVNPGHADLSAKEAQDVIMEVVFPNEVGARILTSE 180
sp|P0DOG6|PB2S1_I34A1 RVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSE 180
sp|P12445|PB2_I34A0 KVERLKHGTFGPVHFRNQVKIRRRVDINPGHANLSAKEAQDVIMEVVFPNEVGARILTSE 180
sp|Q0A2F5|PB2_I83A4 KVGRLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSE 180
sp|P26104|PB2_I77AG KVERLRHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGAKILTSE 180
sp|Q20NV1|PB2_I80AD KVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSE 180
sp|P26115|PB2_I77AF KVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSD 180
sp|Q9Q0V1|PB2_I96A0 KVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSE 180
sp|Q0A449|PB2_I66A1 KVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSE 180
sp|Q6DNK1|PB2_I03A1 KVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSE 180
sp|P26110|PB2_I82A3 KVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSE 180
sp|Q809P5|PB2_I01A3 KVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSE 180
sp|P26112|PB2_I80A8 KVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSE 180
sp|Q0A438|PB2_I49A1 KVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSE 180
sp|Q20P12|PB2_I56A1 KVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSE 180
sp|Q0A427|PB2_I56A2 KVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSE 180
NLS1
CBR
sp|P26107|PB2_I56A3 MNRANQRLNPMHQLLRHFQKDAKILFQNWGIEPIDNIMGMTGILPDMTPSTEMSLRGIRI 480
tr|D9U0Y7|D9U0Y7_9INFA VNRANQRLNPMHQLLRHFQKDAKVLFQNWGVEHIDSVMGMIGVLPDMTPSTEMSMRGIRV 480
tr|D2E5N3|D2E5N3_9INFA VNRANQRLNPMHQLLRHFQKDAKVLFQNWGVEHIDSVMGMIGVLPDMTPSTEMSMRGIRV 480
tr|I6S703|I6S703_9INFA VNRANQRLNPMHQLLRHFQKDAKVLFQNWGVEHIDSVMGMIGVLPDMTPSTEMSMRGIRV 480
sp|Q1PUC9|PB2_I73A5 VNRANQRLNPMHQLLRHFQKDAKVLFQNWGIEHIDNVMGMVGVLPDMTPSTEMSMRGIRV 480
sp|Q6XU90|PB2_I67A0 VNRANQRLNPMHQLLRHFQKDAKVLFQNWGIEHIDNVMGMIGVLPDMTPSTEMSMRGIRV 480
sp|Q3YPY5|PB2_I71A1 VNRANQRLNPMHQLLRHFQKDAKVLFQNWGIEHIDNVMGMIGVLPDMTPSTEMSMRGIRV 480
sp|P26105|PB2_I86A2 VNRANQRLNPMHQLLRHFQKDAKVLFQNWGIEPIDNVMGMIGILPDMTPSTEMSLRGVRV 480
sp|P0DOG6|PB2S1_I34A1 VNRANQRLNPMHQLLRHFQKDAKVLFQNWGVEPIDNVMGMIGILPDMTPSIEMSMRGVRI 480
sp|P12445|PB2_I34A0 VNRANQRLNPMHQLLRHFQKDAKTLFQNWGIEPIDNVMGMIGILPDMTPSTEMSLRGVRV 480
sp|Q0A2F5|PB2_I83A4 VNRANQRLNPMHQLLRHFQKDAKILFQNWGVEPIDNVMGMIGILPDMTPSTEMSLRGVRV 480
sp|P26104|PB2_I77AG VNRANQRLNPMHQLLRHFQKDAKVLFQNWGIEPIDNVMGMIGILPDMTPSTEMSLRGVRV 480
sp|Q20NV1|PB2_I80AD VNRANQRLNPMHQLLRHFQKDAKVLFQNWGIEPIDNVMGMIGILPDMTPNTEMSLRGIRI 480
sp|P26115|PB2_I77AF VNRANQRLNPMHQLLRHFQKDAKVLFQNWGIEPIDNVMGMIGILPDMTPNAEMSLRGIRV 480
sp|Q9Q0V1|PB2_I96A0 VNRANQRLNPMHQLLRHFQKDAKVLFQNWGIEPIDNVMGMIGILPDMTPSAEMSLRGVRV 480
sp|Q0A449|PB2_I66A1 VNRANQRLNPMHQLLRHFQKDAKVLFQNWGIETIDNVMGMIGILPDMTPSTEMSLRGIRV 480
sp|Q6DNK1|PB2_I03A1 VNRANQRLNPMHQLLRHFQKDAKVLFQNWGIEPIDNVMGMIGVLPDMTPSTEVSLRGVRV 480
sp|P26110|PB2_I82A3 VNRANQRLNPMHQLLRHFQKDAKVLFQNWGIEPIDNVMGMIGILPDMTPSTEMSLRGIRV 480
sp|Q809P5|PB2_I01A3 VNRANQRLNPMHQLLRHFQKDAKVLFQNWGIEPIDNVMGMIGILPDMTPSTEMSLRGVRV 480
sp|P26112|PB2_I80A8 VNRANQRLNPMHQLLRHFQKDAKVLFQNWGIEPIDNVMGMIGIMPDMTPSTEMSLRGIRV 480
sp|Q0A438|PB2_I49A1 VNRANQRLNPMHQLLRHFQKDAKVLFQNWGIEPIDNVMGMIGILPDMTPSTEMSLRGVRV 480
sp|Q20P12|PB2_I56A1 VNRANQRLNPMHQLLRHFQKDAKVLFQNWGIEPIDNVMGMIGILPDMTPSTEMSLRGLRV 480
sp|Q0A427|PB2_I56A2 VNRANQRLNPMHQLLRHFQKDAKVLFQNWGIEPIDNVMGMIGILPDMTPSTEMSLRGVRV 480



P26107|PB2_I56A3 Influenza A virus (Equine/Prague/1956 H7N7)
D9U0Y7_9INFA Influenza A virus (Peru/2007 H3N2)
D2E5N3_9INFA Influenza A virus (Thailand/2007 H3N2)
I6S703_9INFA Influenza A virus (Nepal/2006 H3N2)
Q1PUC9|PB2_I73A5 Influenza A virus (Port Chalmers/1973 H3N2)
Q6XU90|PB2_I67A0 Influenza A virus (Tokyo/1967 H2N2)
Q3YPY5|PB2_I71A1 Influenza A virus (Memphis/1971 H3N2)
P26105|PB2_I86A2 Influenza A virus (Equine/Kentucky/1986 H3N8)
P0DOG6|PB2S1_I34A1 Influenza A virus (Puerto Rico/1934 H1N1) (pI = 9.73)
P12445|PB2_I34A0 Influenza A virus (Fowl plague virus/Rostock/1934 H7N1)
Q0A2F5|PB2_I83A4 Influenza A virus (/Turkey/Ireland/1983 H5N8)
P26104|PB2_I77AG Influenza A virus (Budgerigar/Hokkaido/1977 H4N6)
Q20NV1|PB2_I80AD Influenza A virus (Gull/Minnesota/1980 H13N6) (pI = 9.49)
P26115|PB2_I77AF Influenza A virus (Gull/Maryland/1977 H13N6)
Q9Q0V1|PB2_I96A0 Influenza A virus (Goose/Guangdong/1996 H5N1 genotype
Q0A449|PB2_I66A1 Influenza A virus (Turkey/Wisconsin/1966 H9N2)
Q0A438|PB2_I49A1 Influenza A virus (Duck/Germany/1949 H10N7)
Q6DNK1|PB2_I03A1 Influenza A virus (Chicken/Shantou/2003 H5N1)
P26110|PB2_I82A3 Influenza A virus (Seal/Massachusetts/1982 H4N5)
Q809P5|PB2_I01A3 Influenza A virus (Chicken/Hong Kong/2001 H5N1)
P26112|PB2_I80A8 Influenza A virus (strain A/Turkey/Minnesota/833/1980 H4N2)
Q0A427|PB2_I56A2 Influenza A virus (strain A/Duck/England/1/1956 H11N6)
Q20P12|PB2_I56A1 Influenza A virus (strain A/Duck/Czechoslovakia/1956 H4N6)

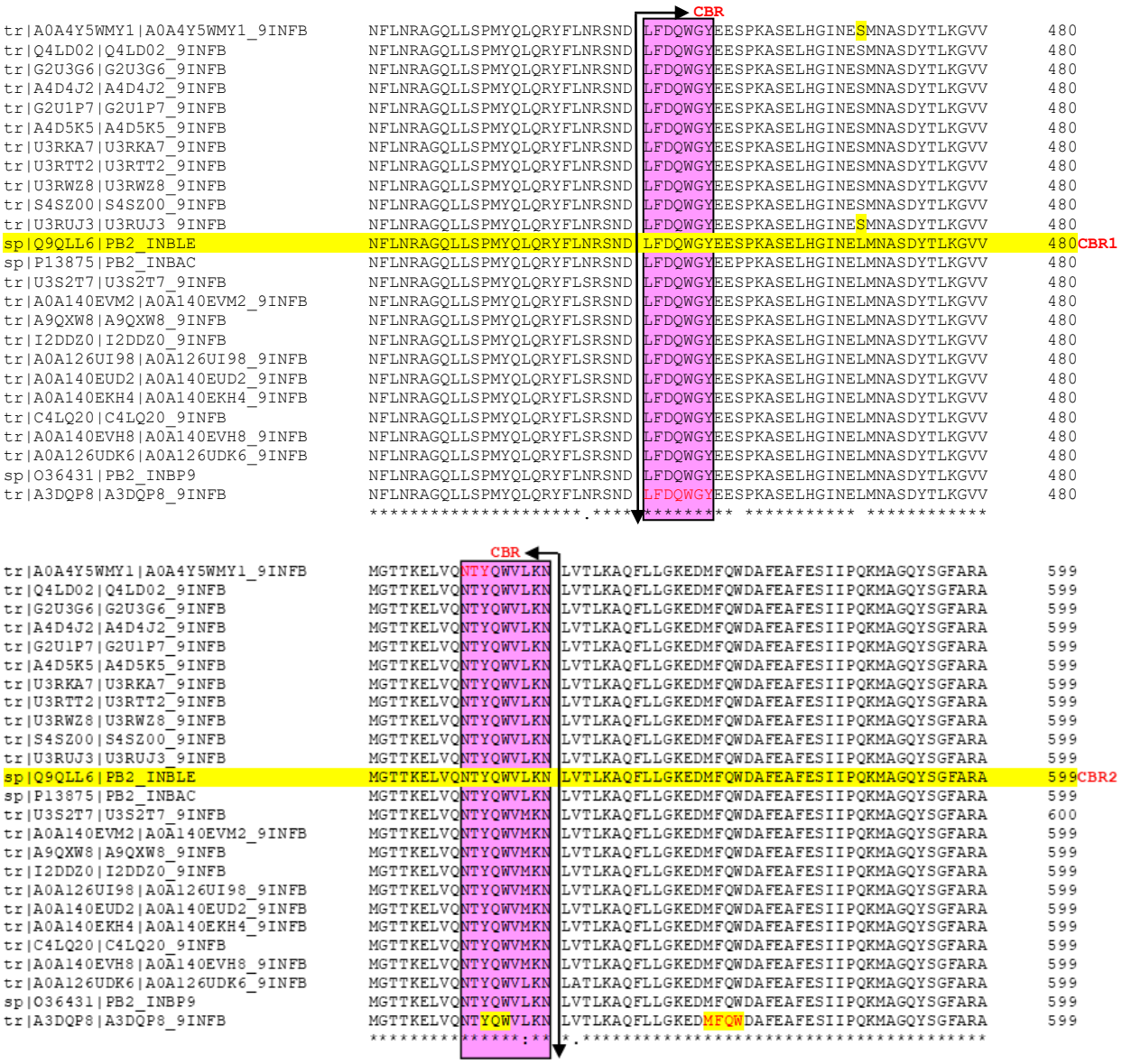
Figure 4 MSA of polymerase basic protein subunits PB2 of different strains of influenza A virus.
NLS, Nuclear localization signal; CBR, Cap-binding region

Figure 5 shows the MSA of the PB2 subunits from various strains of influenza B virus (only the required regions for the discussions are shown here). The Influenza B viral strain (1940) is highlighted and it showed a theoretical pI of 9.24. The PB2 of influenza B viral strains are almost completely conserved from N- to C-terminal than influenza A viral strains. The PB2 subunits of the influenza B viral strains also possess an HNH domain and belong to HNN subfamily. However, in a striking contrast to the influenza A viruses and other HNH endonucleases, in the influenza B viral strains, a –DQ- is identified as the likely first dyad of the HNN domain which is followed by two invariant Ns (highlighted in red). The likely NLSs and CBRs are highlighted in orange and magenta, respectively. DxD/E type of metal binding motifs are found (data not shown). A highly acidic peptide motif is also identified in the C-terminal region (data not shown)

CLUSTAL O (1.2.4) MSA of polymerase PB2 subunit of influenza B viruses

```
tr|A0A4Y5WMY1|A0A4Y5WMY1_9INFB      MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
tr|Q4LD02|Q4LD02_9INFB              MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
tr|G2U3G6|G2U3G6_9INFB              MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
tr|A4D4J2|A4D4J2_9INFB              MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
tr|G2U1P7|G2U1P7_9INFB              MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
tr|A4D5K5|A4D5K5_9INFB              MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
tr|U3RKA7|U3RKA7_9INFB              MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
tr|U3RTT2|U3RTT2_9INFB              MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
tr|U3RWZ8|U3RWZ8_9INFB              MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
tr|S4SZ00|S4SZ00_9INFB              MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
tr|U3RUJ3|U3RUJ3_9INFB              MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
sp|Q9QLL6|PB2_INBLE                 MTLAKIELLKQLLRDNEAKTVLRQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
sp|P13875|PB2_INBAC                 MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIERNPSLRMKWAMCSNFPLA      60
tr|U3S2T7|U3S2T7_9INFB              MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
tr|A0A140EVM2|A0A140EVM2_9INFB      MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
tr|A9QXW8|A9QXW8_9INFB              MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
tr|I2DDZ0|I2DDZ0_9INFB              MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
tr|A0A126UI98|A0A126UI98_9INFB      MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
tr|A0A140EUD2|A0A140EUD2_9INFB      MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
tr|A0A140EKH4|A0A140EKH4_9INFB      MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
tr|C4LQ20|C4LQ20_9INFB              MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
tr|A0A140EVH8|A0A140EVH8_9INFB      MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
tr|A0A126UDK6|A0A126UDK6_9INFB      MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
sp|O36431|PB2_INBP9                 MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
tr|A3DQP8|A3DQP8_9INFB              MTLAKIELLKQLLRDNEAKTVLKQTTVDQYNIIRKFNTSRIEKNPSLRMKWAMCSNFPLA      60
                                    ************************:*****************:*************** 

tr|A0A4Y5WMY1|A0A4Y5WMY1_9INFB      FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|Q4LD02|Q4LD02_9INFB              FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|G2U3G6|G2U3G6_9INFB              FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|A4D4J2|A4D4J2_9INFB              FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|G2U1P7|G2U1P7_9INFB              FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|A4D5K5|A4D5K5_9INFB              FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|U3RKA7|U3RKA7_9INFB              FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|U3RTT2|U3RTT2_9INFB              FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|U3RWZ8|U3RWZ8_9INFB              FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|S4SZ00|S4SZ00_9INFB              FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|U3RUJ3|U3RUJ3_9INFB              FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
sp|Q9QLL6|PB2_INBLE                 FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180NLS1
sp|P13875|PB2_INBAC                 FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|U3S2T7|U3S2T7_9INFB              FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|A0A140EVM2|A0A140EVM2_9INFB      FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|A9QXW8|A9QXW8_9INFB              FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|I2DDZ0|I2DDZ0_9INFB              FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|A0A126UI98|A0A126UI98_9INFB      FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|A0A140EUD2|A0A140EUD2_9INFB      FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|A0A140EKH4|A0A140EKH4_9INFB      FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|C4LQ20|C4LQ20_9INFB              FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|A0A140EVH8|A0A140EVH8_9INFB      FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|A0A126UDK6|A0A126UDK6_9INFB      FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
sp|O36431|PB2_INBP9                 FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
tr|A3DQP8|A3DQP8_9INFB              FLRKMRLDNATWGRITFGPVERVRKRVLLNPLTKEMPPDEASNVIMEILFPKEAGIPRES      180
                                    ************************************************************
```



//End of PB2 subunits of influenza B viruses

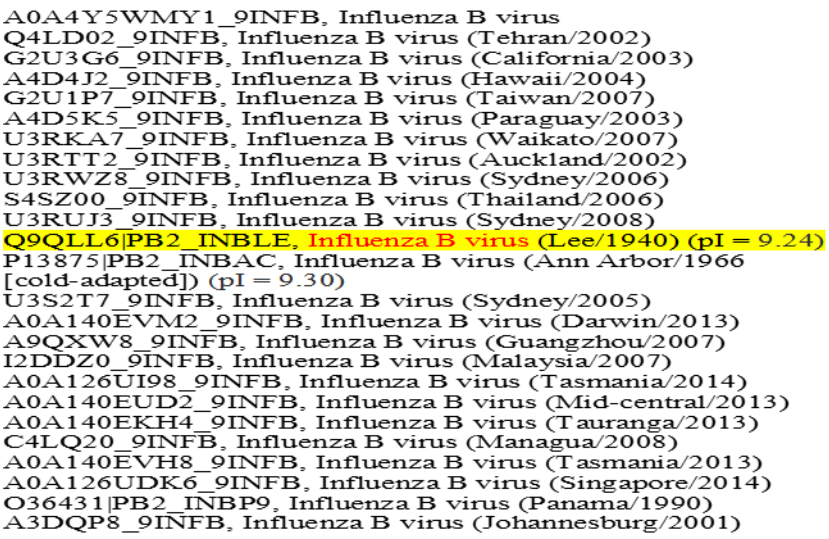
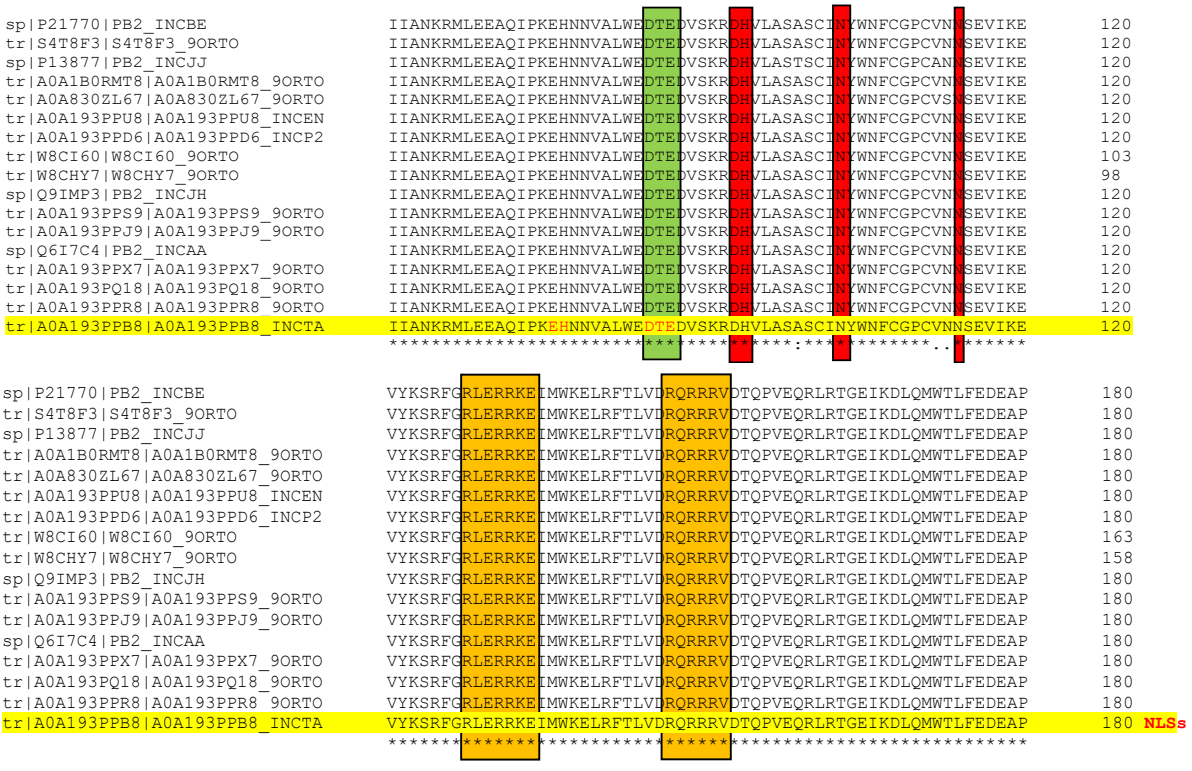


Figure 5 MSA of the polymerases basic protein subunits PB2 of from influenza B viruses

Figure 6 shows the MSA of the PB2 subunits from different strains of influenza C virus (only the required regions for the discussions are shown here). The influenza C virus strain (1947) is highlighted and it showed a theoretical pI of 9.25. But for a small region in the N- and C-terminals, the entire sequence is almost completely conserved in all the strains. The PB2 subunits of the influenza C viral strains are more completely conserved than influenza A viral strains. A –DH- based HNN motif, as in influenza A virus, is indented in a smaller peptide region at the N-terminal, which is followed by two invariant Ns. The likely NLSs, CBRs and metal-binding motifs are highlighted in orange, magenta and green, respectively. The second consecutive basic amino acids which represents the NLS was identified in the N-terminal itself, suggesting a significant divergence in their evolutions.

CLUSTAL O (1.2.4) MSA of the basic protein subunits PB2 of from influenza C viruses.



```
sp|P21770|PB2_INCBE                VRAVQFEYWSEQEEFYGEYKSATALFSRKERSLEWITIGGGINEDRKRLLAMCMIFCRDG       420
tr|S4T8F3|S4T8F3_9ORTO             VRAVQFEYWSEQEEFYGEYKSATALFSRKERSLEWITIGGGINEDRQRLLAMCMIFCRDG       420
sp|P13877|PB2_INCJJ                VRAVQFEYWSEQEEFYGEYKSATALFSRKERSLEWITIGGGINEDRKRLLAMCMIFCRDG       420
tr|A0A1B0RMT8|A0A1B0RMT8_9ORTO     VRAVQFEYWSEQEEFYGEYKSATALFNRKERSLEWITIGGGINEDRKRLLAMCMIFCRDG       420
tr|A0A830ZL67|A0A830ZL67_9ORTO     VRAVQFEYWSEQEEFYGEYKSATALFNRKERSLEWITIGGGINEDRKRLLAMCMIFCRDG       420
tr|A0A193PPU8|A0A193PPU8_INCEN     VRAVQFEYWSEQEEFYGEYKSATALFNRKERSLEWITIGGGINEDRKRLLAMCMIFCRDG       420
tr|A0A193PPD6|A0A193PPD6_INCP2     VRAVQFEYWSEQEEFYGEYKSATALFNRKERSLEWITIGGGINEDRKRLLAMCMIFCRDG       420
tr|W8CI60|W8CI60_9ORTO             VRAVQFEYWSEQEEFYGEYKSATALFNRKERSLEWITIGGGINEDRKRLLAMCMIFCRDG       403
sp|Q9IMP3|PB2_INCJH                VRAVQFEYWSEQEEFYGEYKSATALFSRKERSLEWITIGGGINEDRKRLLAMCMIFCRDG       420
tr|A0A193PPS9|A0A193PPS9_9ORTO     VRAVQFEYWSEQEEFYGEYKSATALFSRKERSLEWITIGGGINEDRKRLLAMCMIFCRDG       420
tr|A0A193PPJ9|A0A193PPJ9_9ORTO     VRAVQFEYWSEQEEFYGEYKSATALFSRKERSLEWITIGGGINEDRKRLLAMCMIFCRDG       420
sp|Q6I7C4|PB2_INCAA                VRAVQFEYWSEQEEFYGEYKSATALFSRKERSLEWITIGGGINEDRKRLLAMCMIFCRDG       420
tr|A0A193PPX7|A0A193PPX7_9ORTO     VRAVQFEYWSEQEEFYGEYKSATALFSRKERSLEWITIGGGINEDRKRLLAMCMIFCRDG       420
tr|A0A193PQ18|A0A193PQ18_9ORTO     VRAVQFEYWSEQEEFYGEYKSATALFSRKERSLEWITIGGGINEDRKRLLAMCMIFCRDG       420
tr|A0A193PPR8|A0A193PPR8_9ORTO     VRAVQFEYWSEQEEFYGEYKSATALFSRKERSLEWITIGGGINEDRKRLLAMCMIFCRDG       420
tr|A0A193PPB8|A0A193PPB8_INCTA     VRAVQFEYWSEQEEFYGEYKSATALFSRKERSLEWITIGGGINEDRRRLLAMCMIFCRDG       420
                                   *********************.******************:*************

sp|P21770|PB2_INCBE                DYFKDAPATITMADLTTKLGREIPYQYVMMNWIQKSEDNLEALLYSRGIVETNPGKMGSS       480
tr|S4T8F3|S4T8F3_9ORTO             DYFKDAPATITMADLTTKLGREIPYQYVMMNWIQKSEDNLEALLYSRGIVETNPGKMGSS       480
sp|P13877|PB2_INCJJ                DYFKDAPATITMADLSTKLGREIPYQYVMMNWIQKSEDNLEALLYSRGIVETNPGKMGSS       480
tr|A0A1B0RMT8|A0A1B0RMT8_9ORTO     DYFKDAPATITMADLTTKLGREIPYQYVMMNWIQKSEDNLEALLYSRGIVETNPGKMGSS       480
tr|A0A830ZL67|A0A830ZL67_9ORTO     DYFKDAPATITMADLTTKLGREIPYQYVMMNWIQKSEDNLEALLYSRGIVETNPGKMGSS       480
tr|A0A193PPU8|A0A193PPU8_INCEN     DYFKDAPATITMADLTTKLGREIPYQYVMMNWIQKSEDNLEALLYSRGIVEPNPGKMGSS       480
tr|A0A193PPD6|A0A193PPD6_INCP2     DYFKDAPATITMADLTTKLGREIPYQYVMMNWIQKSEDNLEALLYSRGIVETNPGKMGSS       480
tr|W8CI60|W8CI60_9ORTO             DYFKDAPATITMADLTTKLGREIPYQYVMMNWIQKSEDNLEALLYSRGIVETNPGKMGSS       463
sp|Q9IMP3|PB2_INCJH                DYFKDAPATITMADLSTKLGREIPYQYVMMNWIQKSEDNLEALLYSRGIVETNPGKMGSS       480
tr|A0A193PPS9|A0A193PPS9_9ORTO     DYFKDAPATITMADLSTKLGREIPYQYVMMNWIQKSEDNLEALLYSRGIVETNPGKMGSS       480
tr|A0A193PPJ9|A0A193PPJ9_9ORTO     DYFKDAPATITMADLSTKLGREIPYQYVMMNWIQKSEDNLEALLYSRGIVETNPGKMGSS       480
sp|Q6I7C4|PB2_INCAA                DYFKDAPATITMADLSTKLGREIPYQYVMMNWIQKSEDNLEALLYSRGIVETNPGKMGSS       480
tr|A0A193PPX7|A0A193PPX7_9ORTO     DYFKDAPATITMADLSTKLGREIPYQYVMMNWIQKSEDNLEALLYSRGIVETNPGKMGSS       480
tr|A0A193PQ18|A0A193PQ18_9ORTO     DYFKDAPATITMADLSTKLGREIPYQYVMMNWIQKSEDNLEALLYSRGIVETNPGKMGSS       480
tr|A0A193PPR8|A0A193PPR8_9ORTO     DYFKDAPATITMADLSTKLGREIPYQYVMMNWIQKSEDNLEALLYSRGIVETNPGKMGSS       480
tr|A0A193PPB8|A0A193PPB8_INCTA     DYFKDAPATITMADLSTKLGREIPYQYVMMNWIQKSEDNLEALLYSRGIVETNPGKMGSS       480
                                   ***************:*****************************************  ********
```

//End of PB2 subunits of influenza C viruses

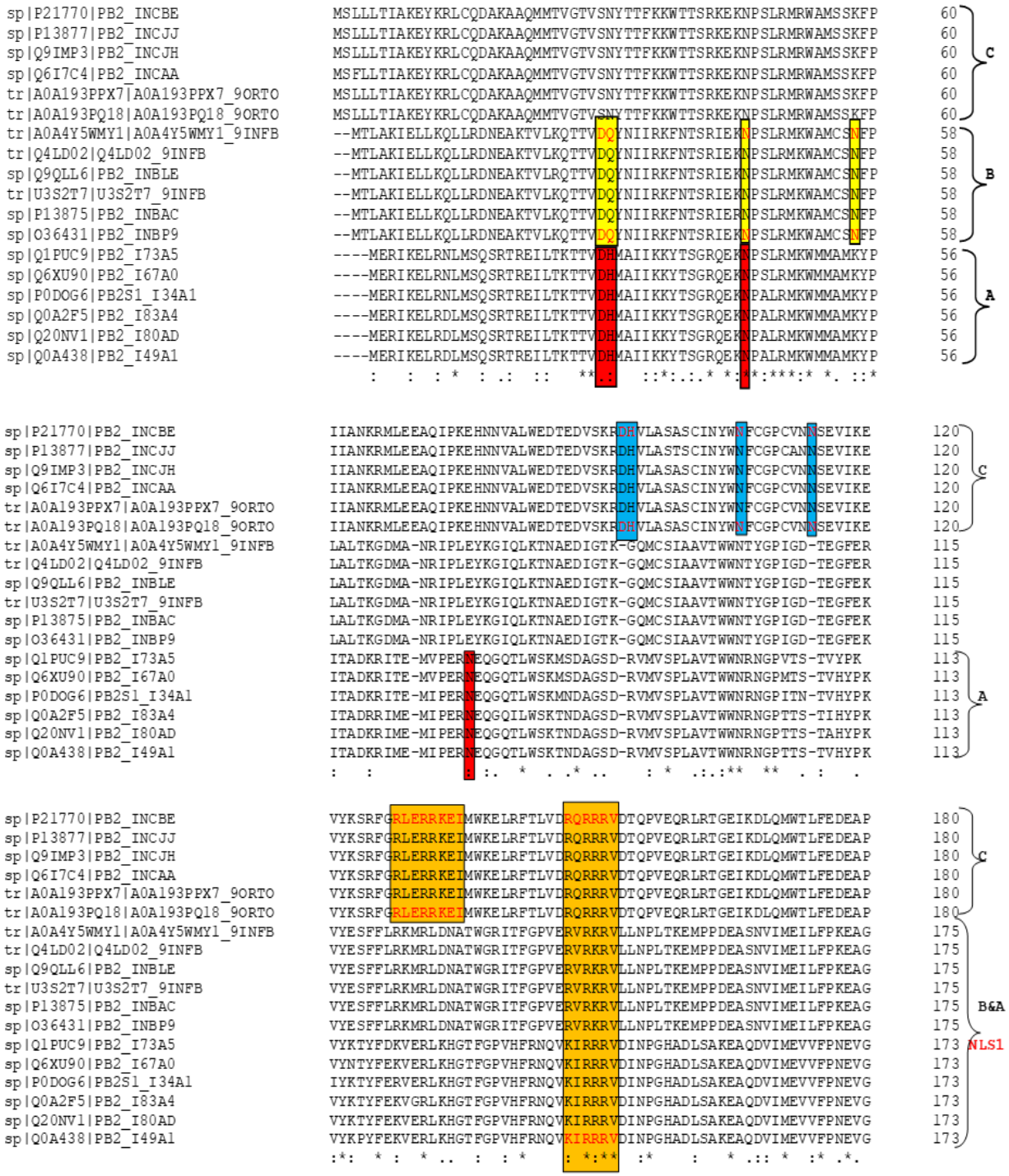
```
sp|P21770|PB2_INCBE                APMVTGQDLIDVGFGQKVRLFVGQGSVRTFKRTASQRAASSDVNKNVKKIKMSN     774
tr|S4T8F3|S4T8F3_9ORTO             APMVTGQDLIDVGFGQNVRLFVGQGSVRTFKRTASQRAASSDVNKNVKKIKMS-     773
sp|P13877|PB2_INCJJ                APMVTRQDLIDVGFGQKVRLFIGQGSVRTFKRTASQRAASSDVNKNVKKIKMSN     774
tr|A0A1B0RMT8|A0A1B0RMT8_9ORTO     APMVTRQDLIDVGLGQKVRLFVGQGSVRTFKRTASQRAASSDVNKNVKKIKMS-     773
tr|A0A830ZL67|A0A830ZL67_9ORTO     APMVTRQDLIDVGFGQKVRLFVGQGSVRTFKRTASQRAASSDVNKNVKKIKMSN     774
tr|A0A193PPU8|A0A193PPU8_INCEN     APMVTRQDLIDVGFGQKVRLFVGQGSVRTFKRTASQRAASSDVNKNVKKIKMSN     774
tr|A0A193PPD6|A0A193PPD6_INCP2     APMVTRQDLIDVGFGQKVRLFVGQGSVRTFKRTASQRAASSDVNKNVKKIKMSN     774
tr|W8CI60|W8CI60_9ORTO             APMVTRQDLIDVGFGQKVRLFVGQGSVRTFKRTASQRAASSDVNKN--------     749
sp|Q9IMP3|PB2_INCJH                APMVTRQDLIDVGFGQKVRLFVGQGSVRTFKRTASQRAASSDVNKNVKKIKMSN     774
tr|A0A193PPS9|A0A193PPS9_9ORTO     APMVTRQDLIDVGFGQKVRLFVGQGSVRTFKRTASQRAASSDVNKNVKKIKMSN     774
tr|A0A193PPJ9|A0A193PPJ9_9ORTO     APMVTRQDLIDVGFGQKVRLFVGQGSVRTFKRTASQRAASSDVNKNVKKIKMSN     774
sp|Q6I7C4|PB2_INCAA                APMVTRQDLIDVGFGQKVRLFVGQGSVRTFKRTASQRAASSDVNKNVKKIKMSN     774
tr|A0A193PPX7|A0A193PPX7_9ORTO     APMVTRQDLIDVGFGQKVRLFVGQGSVRTFKRTASQRAASSDVNKNVKKIKMSN     774
tr|A0A193PQ18|A0A193PQ18_9ORTO     APMVTRQDLIDVGFGQKVRLFVGQGSVRTFKRTASQRAASSDVNKNVKKIKMSN     774
tr|A0A193PPR8|A0A193PPR8_9ORTO     APMVTRQDLIDVGFGQKVRLFVGQGSVRTFKRTASQRAASSDVNKNVKKIKMSN     774
tr|A0A193PPB8|A0A193PPB8_INCTA     APMVTRQDLIDVGFGQKVRLFVGQGSVRTFKRTASQRAASSDVNKNVKKIKMSN     774
                                   ***** *******:**:****:************************
```

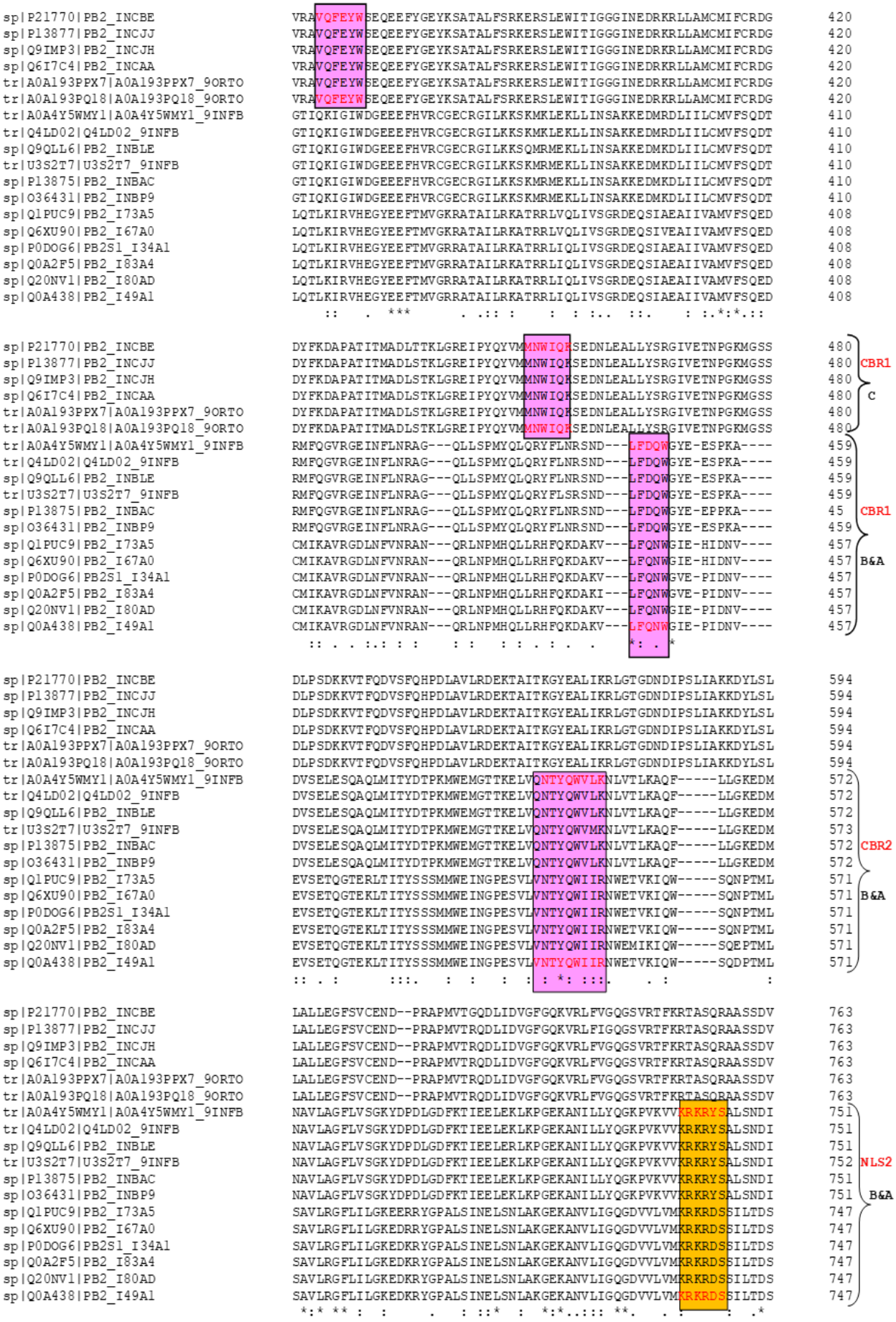
P21770|PB2_INCBE, Influenza C virus (strain C/Berlin/1985)
S4T8F3_9ORTO, Influenza C virus (C/Eastern India/2011)
P13877|PB2_INCJJ, Influenza C virus (strain C/1950)
A0A1B0RMT8_9ORTO, Influenza C virus (C/India/2011) (pI = 9.24)
A0A830ZL67_9ORTO, Influenza C virus
A0A193PPU8_INCEN, Influenza C virus (strain C/England/1983)
A0A193PPD6_INCP2, Influenza C virus (strain C/Pig/Beijing/1981) (pI = 9.24)
W8CI60_9ORTO, Influenza C virus (C/Singapore/2006)
Q9IMP3|PB2_INCJH, Influenza C virus (strain C/Johannesburg/1/1966) (pI = 9.24)
A0A193PPS9_9ORTO, Influenza C virus (C/Paris/1967) (pI = 9.24)
A0A193PPJ9_9ORTO, Influenza C virus (C/Tokyo/2010
Q6I7C4|PB2_INCAA, Influenza C virus (strain C/Ann Arbor/1950)
A0A193PPX7_9ORTO, Influenza C virus (C/Yamagata/2008)
A0A193PQ18_9ORTO, Influenza C virus (C/Fukuoka/2012)
A0A193PPR8_9ORTO, Influenza C virus (C/Greece/1/1979)
A0A193PPB8_INCTA, Influenza C virus (strain C/Taylor/1947) (pI = 9.25)

Figure 6 MSA of BP2 subunit of the polymerase from different strains of Influenza C viruses

Figure 7 shows the mix and match analysis of all the three human influenza viruses (only the required regions for the discussions are shown here). It is interesting to note that the human influenza viruses A and B align in the proposed conserved motifs, suggesting their possible similar origins. The first dyad (-DH/DQ-)- which is shown to involve in Mg^{2+}-binding and as the proton acceptor are aligned in the influenza viruses A and B, but not in C. The dipeptide alignment suggests that the B virus could possibly use the Q as the proton acceptor instead of H, as reported in other HNH/N endonucleases. Though one of the proposed NLSs align in all three viruses, the CBRs align only in A and B viruses, further suggesting their similar origins (Fig. 6). From the MSA analysis, it is clear that the A and B influenza viruses are close to each other, whereas the influenza virus C has diverged significantly during evolution or may have a different origin.

CLUSTAL O (1.2.4) Mix and Match Analysis of the PB2 subunits from Influenza A, B & C viruses




CBR1
C
CBR1
B&A
CBR2
B&A
NLS2
B&A

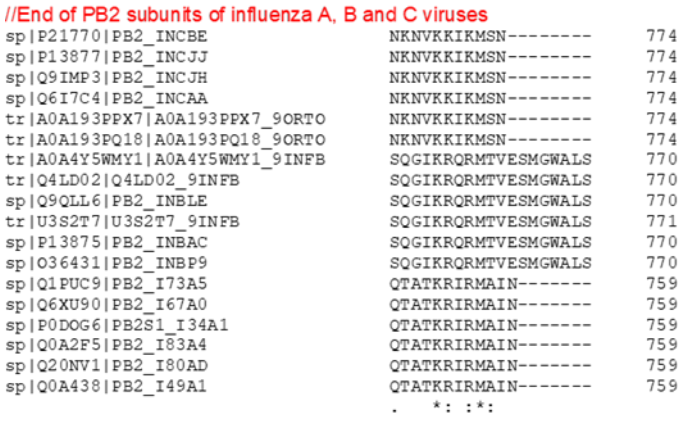


Figure 7 A mix and match analysis of human influenza viruses A, B and C (For legends, see Figs. 4, 5 and 6)

Table 1 shows the active site regions of different HNH/N endonucleases which makes both double-stranded breaks or nicks on DNAs and single–stranded cleavage on RNAs (CRIPSR-Cas13a and CRIPSR-Cas12a) suggesting the proposed regions on the influenza viral polymerase PB2 subunits could also make a cleavage on the host cell mRNAs in the cap-snatching process. Proposed HNH/N endonuclease domains, by sequence similarity, are highlighted in yellow and needs further experimental validation.

Table 1 Active site regions in different HNH/N family of endonucleases

HNH/N Type (Organism)	Active Site Region
HH-Homing endonuclease domain (Bacteriophage Bp7 I-*Tev*II)	-YEI**HH**KDGNRENNDLD**N**LMCLSIQE**H**Y^{49}-
HH-based (HNH/N) group II introns (*S. cerevisiae*)	-LEV**HH**VRTLNNAA**N**KIKDDYLLGRMIKM**N**RKQITICKTC**H**F^{842}-
HH-based mcr A (HNH) restriction endonuclease (*E. coli*)	-LEV**HH**VIPLSSGGADTTD**N**CVALCPNC**H**RELHYS^{258}-
DH-based HNN endonucleases (*E. coli* plasmids)	-WYA**DH**VQAV// PEADCPE**N**LVPACAPC**N**LLK^{85}-
DH-based HNN endonucleases (*E. proavitum*)	-MEA**DH**ITPWHEGGKTTSV**N**CQMLCKDC**N**RRK^{355}-
DH-based HNH Endonuclease VII (Resolvase) (T4 Phage)	-LDH**DH**ELNGPKAGKVRGLLC**N**LCNAAEGQMK**H**KFNR^{74}-
HH based HNH Colicin endonuclease (Type 9*) (*E. coli*)	-YEL**HH**DKPISQGGEVYDMD**N**IRVTTPKRHIDI**H**RGK^{592}-
HH- HNH Pyocins endonuclease (Type-S1) (*P. aeruginosa*)	-IEI**HH**KVRVADGGGVYNMG**N**LVAVTPKRHIE**I**HKGGK^{618}-
HH- HNH Pyocins endonuclease (Type-S2) (*P. aeruginosa*)	-IEI**HH**KVRIADGGGVYNMG**N**LVAVTPKRHIEI**H**KGGK^{689}-
DH-based Influenza virus A (HNN endonuclease domain)*	-TTV**DH**^{27}MAIIKKYTSGRQEK**N**^{42}PALRMKWM//PITADKRITEMIPER**N**^{71}E-
DQ-based Influenza virus B (HNN endonuclease domain)*	-TTV**DQ**^{29}YNIIRKFNTSRIEK**N**^{44}PSLRMKWAMCS**N**^{56}F-
DH-based Influenza virus C (HNN endonuclease domain) *	-SKR**DH**^{94}VLASASC**N**^{103}YWNFCGPCV**N**^{113}NS-
HH-based CRISPR-Cas13a (HNH endonucleases domains)* >*sp|P0DPB8|CS13A), Listeria seeligeri*	-TLI**HH**^{111}LGVLFFCDYMY**N**^{73}RRE//DRKKVLISRD**K**NG^{56}- -QIM**HH**^{174}KKSSFYKSV**N**^{883}KTICRPEQKQMKKL**V**IG^{102}-
HH-based CRISPR-Cas13b (HNN endonuclease domain) >*tr|E6K398|Prevotella buccae*	-DYM**HH**^{139}ENIDMQRDFTHL**N**^{212}RKKQVGRT**K**N^{222}II-
DH-based CRISPR-Cas12a (HNN endonuclease domain)* >*sp|U2UMQ6|CS12A, Acidaminococcus* sp.	-AKG**H**H^{730}GKPNLHTLYWTGLFSPE**N**^{778}LAKTSIKL**N**G^{792}-
DH-based CRISPR-Cas9 (HNN endonuclease domain)* >*sp|J3F2B0|CRISPR-CAS9, Actinomyces naeslundii*	-CQL**DH**^{581}IVPQAGPGSN**N**^{593}RRGNLVAVCERC**N**^{606}RSKS^{610}-
DH-based CRISPR-Cas9 (HNN endonuclease domain) >*sp|Q99ZW2|CRIPRS-Cas9, Streptococcus pyogenes*	-YDV**DH**^{840}IVPQSFLKDDSID**N**^{854}KVLTRSDK**N**^{863}RGKS^{867}-
SDM & X-ray data	
Type II restriction endonuclease **R.KpnI**	-LTP**D**^{148}**H**^{149}//DV**N**D//RH**C**^{175}VMKK-**Mg**$^{2+}$ [31]
ColE7 HNH endonuclease (*E. coli*) (**Colicin**)	-FEL**H**^{545}EKPISQNGGVYDMD**N**^{560}ISVVTPKRHIDI**H**^{573}RGK- **Zn**$^{2+}$ [32]
Pyocin S8 (*P. aeruginosa*)	-Y**E**I**HH**^{789}VVQISQGGAVYDID**N**^{804}LRVMTPKMHIQVHS**N**^{819}KGK-**Zn**$^{2+}$ [33]
Vvn (*Vibrio vulnicus* HNN nuclease)	-EW**E**H^{80}VV----------H**N**^{118}LTPAIGEV**N**^{127}GDR- **Mg**$^{2+}$ [34]

Adapted from Palanivelu (22) *By sequence similarity. Experimentally validated active site amino acids of the HNH/N domains are highlighted in dark blue (SDM) and light blue (X-ray crystallography).

CRISPR-Cas13a is a single-molecule effector of the Class II, Type VI family of CRISPR-Cas systems that is part of the bacterial and archaeal defense systems. All so far characterized Cas13-family members possess two distinct active sites: one for pre-crRNA processing and the other for target ssRNA cleavage. Recently, Kick et al. [35] found that the Cas13a from the purple bacteria, Rhodobacter capsulatus, possessed two HEPN (higher eukaryotes and prokaryotes nucleotide binding) domains which exhibited nuclease activity. Interestingly, they found that RcCas13a did not rely on the catalytic HEPN-domains for pre-crRNA processing since RcCas13a mutants, where the active-site residues in one or both of the HEPN1 or HEPN2 domains were mutated to Ala, were not affected in their pre-crRNA processing activity, suggesting the active site is elsewhere and possibly in the HNH domain.

The CRISPR-associated protein Cas12a (Cpf1) possesses two distinct nuclease activities: endoribonuclease activity for processing its own guide RNAs and RNA-guided DNase activity for target DNA cleavage. Like CRISPR-Cas9, CRISPR-Cas12a has a conserved RuvC nuclease domain at its C-terminal region.

In CRISPR-Cas9 enzymes, the HNH domain binds to the target DNA and cleaves it.

3.2. HNH/N Active Site Analyses of Human Influenza Viruses A, B and C

Saravanan et al. [31] have analyzed the HNH motif in the Type II restriction endonuclease R.KpnI, a member of the HNH nuclease superfamily, which possess the HNH domain with the conserved amino acids D148, H149 and Q175. By SDM analysis they have shown that the D148→G/A and H149→L/A led to complete loss of activity and the mutant Q175→E, failed to bind DNA at the standard conditions, although the DNA binding and cleavage was rescued at pH 6.0.

Furthermore, the nuclease domain of ColE7 (nuclease-ColE7) purified from *Escherichia coli* contained a one-to-one stoichiometry of zinc ion and that this zinc-containing enzyme hydrolyzed DNA without externally added divalent metal ions. The apo-enzyme, in which the indigenous zinc ion was removed from the nuclease-ColE7, had no detectable DNase activity [36]. SDM experiments followed by fluorescence resonance energy transfer (FRET) assays, by Huang and Yuan [32] to decipher the role of conserved Asn and His residues in the H545, N560, H573 motif of the Colicin, ColE7, found that in the H545 mutants, the activity was completely abolished while activities of N560 and H573 mutants varied from 6.9% to 83.2% of the wild-type activity. Both N560→A and N560→D mutants contained a disordered loop in the HNH motif due to the disruption of the hydrogen bond network surrounding the side-chain of residue of N560 [32].

The active site amino acids of Vvn HNH endonuclease were analyzed by both SDM and X-ray crystallography by Li et al. [34]. They found that the mutation of the invariant His80→Ala abolished the endonuclease activity of Vvn, demonstrating the critical importance of this residue in DNA hydrolysis. The X-ray crystallographic study has further suggested that the Vvn HNH endonuclease hydrolyzed DNA by a general single-metal ion mechanism. The metal ion located in the ββα-metal motif was assigned as a magnesium ion in Vvn [34]. The X-ray crystallographic data have further shown that the magnesium ion was bound to Glu79, Asn127 and four water molecules. A water molecule, W1, bridges the invariant His80 and a Mg^{2+}-bound water (Table 1).

3.3. Proposed Mechanism of Action of HNN of Human Influenza Viruses A, B and C

The invariant D/H is always found to be engaged with a metal ion and the first invariant **His** of D/H**H**NH/N domain acts as a general base to activate a metal-bound water molecule for the nucleophilic attack on the scissile phosphate bond of nucleic acids, and the first Asn interacts with base to be cleaved and thus, generating 3'-OH for subsequent viral mRNA synthesis (in the nucleus) by the PB1 subunit of the polymerase in influenza viruses. The second His/Asn stabilizes the leaving group (Fig. 7) [22, 23].

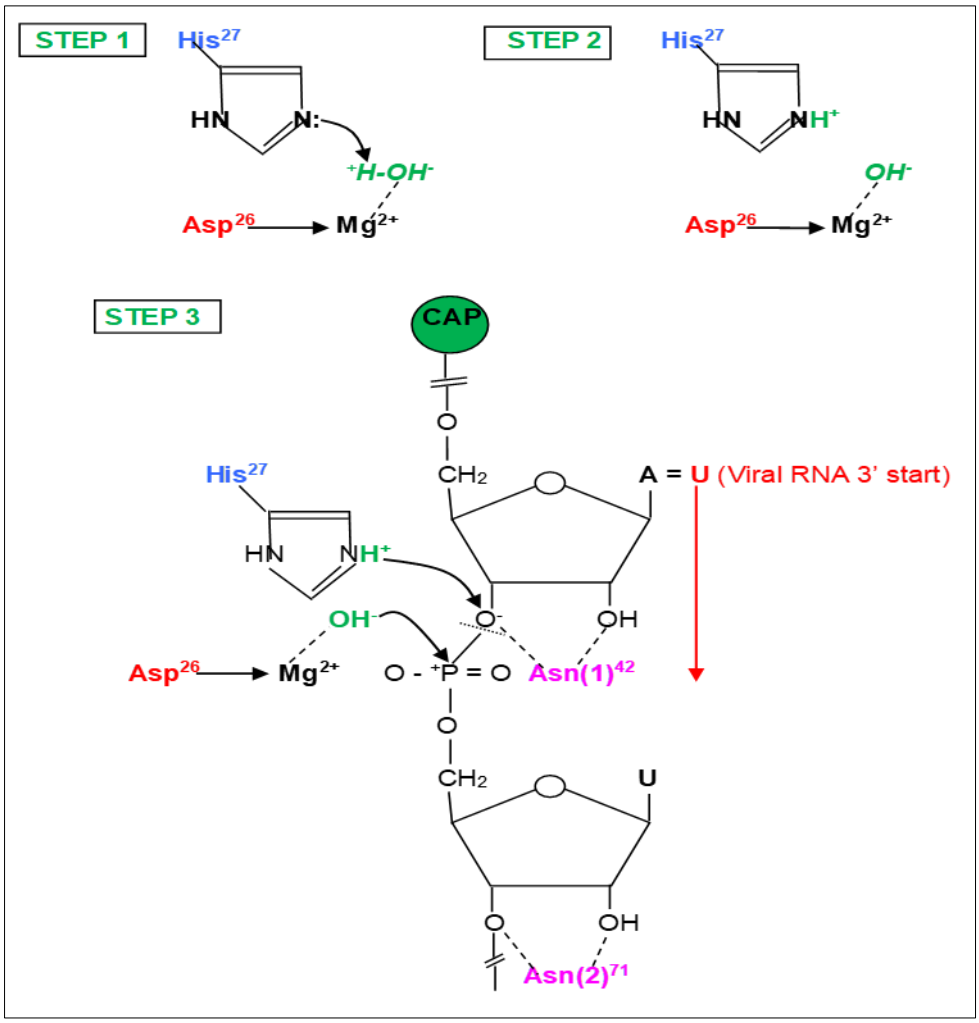


Figure 8 Proposed mechanism of 'Cap-binding and cleavage by PB2 subunits of influenza polymerase (Numbering from influenza virus A, H1N1). CAP, Host mRNA cap structure

4. Conclusions

The endonuclease is suggested to be present in the PB1 and PA subunits of the heterotrimeric human influenza polymerases. However, by sequence similarity, a typical HNH/N motif is identified in the cap-snatching PB2 subunit itself, suggesting that it could not only bind the cap structures of the host mRNAs but also could cleave it. The PB2 subunits of the human influenza viruses A and B are very close to each other, whereas the C has significantly diverged from them. MSA analyses have also shown that even though these viruses are genetically diverged, but share a common genetic ancestry. These results will facilitate the optimization of endonuclease inhibitors as potential new anti-influenza drugs, and could also help in developing new antiviral drugs for the treatments of flu in the future.

Compliance with ethical standards

Acknowledgments

The author wishes to thank Dr. H. Shakila, Professor and Head, Department of Molecular Microbiology, School of Biotechnology, Madurai Kamaraj University, Madurai, for useful suggestions on the manuscript.

Disclosure of conflict of interest

The author has declared that no competing interests exist.

References

[1] Palese P. Influenza: Old and new threats. Nat Medicine. 2004; 10: S82–S87.

[2] Dangi T, Jain A. Influenza Virus: A Brief Overview. Proc. Natl. Acad. Sci. (USA). 2012; 82(1):111–121. DOI 10.1007/s40011-011-0009-6.

[3] Su S, Fu X, Li G, Kerlin F, Veit M. Novel Influenza D virus: Epidemiology, pathology, evolution and biological characteristics, Virulence. 2017; 8:1580-1591.

[4] Guo Y, Jin F, Wang P, Wang M, Zhu JM. Isolation of Influenza virus C from pigs and experimental infection of pigs with influenza C virus. J Gen Virol. 1983; 64:177-182.

[5] Molinari NA, Ortega-Sanchez IR, Messonnier ML, Thompson WW, Wortley PM, Weintraub E, Bridges CB. The annual impact of seasonal influenza in the US: measuring disease burden and costs. Vaccine. 2007; 25:5086–5096. [PubMed: 17544181]

[6] Muraki Y, Hongo S. The molecular virology and reverse genetics of influenza C virus. Jpn J Infect Dis. 2010; 63:157–65. [PubMed: 20495266]

[7] Krammer F, Palese P. Advances in the development of influenza virus vaccines. Nat Rev Drug Discov. 2015; 14:167–182.

[8] Krug R, Fodor E. Textbook of Influenza. Webster R, Monto A, Braciale T, Lamb R, editors. Wiley-Blackwell; 2013. pp. 57-66.

[9] Fodor E. The RNA polymerase of influenza a virus: mechanisms of viral transcription and replication. Acta Virol. 2013; 57:113–22. [PubMed: 23600869]

[10] Resa-Infante P, Jorba N, Coloma R, Ortin J. The influenza virus RNA synthesis machine: advances in its structure and function. RNA Biol. 2011; 8:207–15. [PubMed: 21358279]

[11] Plotch SJ, Bouloy M, Ulmanen L, Krug RM. A unique cap(m7GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. Cell. 1981;23:847-858.

[12] Pflug A, Guilligay D, Reich S, Cusack S. Structure of influenza A polymerase bound to the viral RNA promoter. Nature. 2014;516:355–360.

[13] Fan H, Walker AP, Carrique L, Keown JR, Martin IS, Karia D, Sharps J, Hengrung N, Pardon E, Steyaert J, Grimes JM, Fodor E. Structures of influenza A virus RNA polymerase offer insight into viral genome replication. Nature. 2019; 573:287-295.

[14] Lu J, Wu T, Zhang B, Liu S, Song W, Qiao J, Ruan H. Types of nuclear localization signals and mechanisms of protein import into the nucleus. Cell Commun Signal. 2021;19, 1-10

[15] Willis AN, Dean SE, Habbouche JA, Kempers BT, Ludwig ML, Sayfie AD, Lewis SP, Harrier S, DeBruine ZJ, Garrett R, et al. Nuclear localization signal sequence is required for VACM-1/CUL5-dependent regulation of cellular growth. Cell Tissue Res. 2017;368(1):105–114.

[16] Yan C, Shi W, Gu J, Lee RJ, Zhang Y. Design of a novel nucleus-targeted NLS-KALA-SA nanocarrier to delivery poorly water-soluble anti-tumor drug for lung cancer treatment. J Pharm Sci. 2021;S0022–3549(21):00001.

[17] Monod A, Swale C, Tarus B, Tissot A, Delmas B, Ruigrok RW, et al. Learning from structure-based drug design and new antivirals targeting the ribonucleoprotein complex for the treatment of influenza. Expert Opin Drug Discov. 2015; 10:345–371.

[18] Dias A, et al. The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. Nature. 2009; 458:914–8. [PubMed: 19194459]

[19] Guilligay D, et al. The structural basis for cap binding by influenza virus polymerase subunit PB2. Nat Struct Mol Biol. 2008; 15:500–6. [PubMed: 18454157]

[20] Yuan P, Bartlam M, Lou Z, Chen S, Zhou J, He X, et al. Crystal structure of an avian influenza polymerase PA(N) reveals an endonuclease active site. Nature. 2009; 458:909–13.

[21] Shi L., Summers DF, Peng Q, Galarza JM. Influenza A virus RNA polymerase subunit PB2 is the endonuclease which cleaves host cell mRNA and functions only as the trimeric enzyme. Virology. 1995; 208:38–47.

[22] Palanivelu P. Assessment and Analyses of Homing Endonucleases and Mechanism of Action of CRISPR-Cas9 HNH Endonucleases, In Current Advances in Chemistry and Biochemistry Vol. 1, pp 20-48. SCIENCEDOMAIN International Book Publishers, UK. Print ISBN: 978-93-90516-10-0, eBook ISBN: 978-93-90516-18-6. DOI: 10.9734/bpi/cacb/v1/2273E, 2021

[23] Xu SY, and Gupta YK. Natural zinc ribbon HNH endonucleases and engineered zinc finger nicking endonuclease. Nucleic Acids Res. 2013, Vol. 41, 378–390.

[24] Xu SY, Sequence-specific DNA nicking endonucleases. Bio Mol Concepts. 2015; 6(4): 253–267

[25] Li ML, Rao P, Krug RM. The active sites of the influenza cap-dependent endonuclease are on different polymerase subunits. EMBO J. 2001; 20, 2078–2086.

[26] Blaas D, Patzelt E, Kuechler E. Identification of the cap binding protein of influenza virus. Nucleic Acids Res. 1982;10:4803–4812.

[27] Krug RM, Ueda M, Palese P. Temperature-sensitive mutants of influenza WSN virus defective in virus- specific RNA synthesis. J Virol. 1974; 64: 790-796.

[28] Mahy BWJ, Barret T, Nichol ST, Penn CR, Wolstenholme AJ. Analysis of the functions of influenza virus genome RNA segments by use of temperature-sensitive mutants of fowl plague virus. In "Replication of negative strsnd viruses" (DHL Bishop and RW Compans, Eds). Pp 379-387, Elsevier/North Holland, New York, 1981.

[29] Kobayashi M, Tuchiya K, Nagata K, Ishihama A Reconstitution of influenza virus RNA polymerase from three subunits expressed using recombinant baculovirus system. Virus Res. 1992; 22: 235–245

[30] Honda A, Mizumoto K, Ishihama A. Two separate sequences of PB2 subunit constitute the RNA cap-binding site of influenza virus RNA polymerase. Genes Cells. 1999; 4, 475–485.

[31] Saravanan M, Bujnicki JM, Cymerman IA, Rao DN, Nagaraja V. Type II restriction endonuclease R.KpnI is a member of the HNH nuclease superfamily. Nucleic Acids Res. 2004; 23:6129-6135.

[32] Huang H, HS Yuan. The Conserved Asparagine in the HNH Motif Serves an Important Structural Role in Metal Finger Endonucleases. J Mol Biol. 2007; 368:812-821

[33] Turano H, Gomes H, Domingos RM, Degenhardt MFS, Oliveira CLP, Garratt RC, Lincopan N Netto LES. Molecular Structure and Functional Analysis of Pyocin S8 from Pseudomonas aeruginosa Reveals the Essential Requirement of a Glutamate Residue in the H-N-H Motif for DNase Activity. J Bacteriol. 2020; 202:1-17, e00346-20

[34] Li CL, Hor LI, Chang ZF, Tsai LC, Yang WZ, Yuan HS. DNA binding and cleavage by the periplasmic nuclease Vvn: a novel structure with a known active site. EMBO J. 2003; 22:4014-4025.

[35] Kick LM, von Wrisberg MK, Runtsch LS, Schneider S. Structure and mechanism of the RNA dependent RNase Cas13a from Rhodobacter capsulatus. Commun Biol. 2022; 5:71, https://doi.org/10.1038/s42003-022-03025-4.

[36] Ku WY, Liu YW, Hsu YC, Liao CC, Liang PH, Yuan HS, Chak FU. The zinc ion in the HNH motif of the endonuclease domain of colicin E7 is not required for DNA binding but is essential for DNA hydrolysis. Nucleic Acids Res. 2002; 30:1670-1678.