American Journal of Current Microbiology
http://ivyunion.org/index.php/ajcmicrob/

Research Article

Conservational Analysis of Influenza A Virus RNA-dependent RNA Polymerase

Vivek Darapaneni[1*]

[1] Department of virology and computational biochemistry, Saket Institute for Biomolecular Research, Visakhapatnam, India

Abstract

The RNA-dependent RNA polymerase of influenza A virus is important for catalyzing the replication and transcription of viral RNA. The polymerase of Influenza A Virus plays a significant role in the infectious virus life cycle. The objective of the present study was to identify the residue conservation in polymerase complex (PA, PB1 and PB2 proteins) of Influenza A Virus. The study was based on 11,966 amino acid sequences for the PA protein; 10,605 amino acid sequences for the PB1 protein and 11,331 sequences for the PB2 protein. The PA, PB1 and PB2 proteins exhibited similar level of sequence conservation. On the whole, this study exposed residues which are universally conserved among different influenza A virus subtypes. These universally conserved residues might be involved in either structure stabilizing or protein-protein interactions. Clusters of highly variable residues detected on the surface of the PA, PB1 and PB2 proteins might be linked to adaptation to various hosts or evolutionary pressure to evade the host immune system. The conserved residues identified in the present study could form a platform for designing universal anti-influenza drugs which are resistant to mutations arising in the future.

Keywords: Polymerase; Polymerase acidic; Polymerase basic 1; Polymerase basic 2; Conservation; Muta Resistance; Influenza virus

Peer Reviewer: Sihua Peng, Department of Biology, Shanghai Ocean University, China

Received: July 6, 2015; **Accepted**: August 19, 2015; **Published:** September 21, 2015

Competing Interests: The authors have declared that no competing interests exist.

*Correspondence to: Vivek Darapaneni, Department of virology and computational biochemistry, Saket Institute for Biomolecular Research, Visakhapatnam, India; **Email**: darapanenivivek.sibr@gmail.com

Introduction

Influenza virus belongs to orthomyxoviridae family, which is enveloped with negative sense RNA genome. A, B and C are major serotypes of influenza virus [1]. Influenza A Virus (IAV) with its global dominance is a significant pathogen among the three serotypes [1]. IAV is responsible for 250,000 to 500,000 human deaths worldwide [2] and also results in considerable losses among domesticated birds [3]. IAV is responsible for five devastating pandemics in the years 1918/1919, 1957, 1968, 1977 and 2009 [4-8]. Apart from these pandemics, IAV also causes seasonal epidemics with moderate severity. When a pathogenic IAV of animal or avian origin attains the potential of efficient human-human transmission, then influenza pandemics seems to arise [9, 10]. This potential can be accomplished either by an event of mutation or by reassortment of viral RNA segments between animal/avian and human viruses [9, 10]. In the recent past, pathogenic avian influenza viruses namely, H5N1, H6N1, H7N7, H9N2 and H10N7 have acquired the ability to infect humans [11]. Therefore, in the future it is highly likely that a novel avian IAV will acquire the potential to transmit efficiently among humans and results in a devastating pandemic. Vaccines and antivirals are available against annual epidemics caused by IAV. But these are rendered ineffective against antiviral resistant strains of IAV emerging in the future, which is a cause of concern [12]. While attempts are underway to develop antivirals and vaccines which are universally applicable and which does not lead to resistance.

Viral neucleocapside of IAV consists of eight separate segments and Segment 1 to segment 3 encode three proteins namely, polymerase basic 2 (PB2), polymerase basic 1 (PB1) and polymerase acidic (PA) respectively [1]. Collectively PA, PB1 and PB2 form RNA- dependent RNA polymerase (RdRp) with a molecular weight of 250 KDa [13]. RdRp of IAV is involved in (i) transcription and replication of viral RNA [14] (ii) polyadenylation of mRNA in vitro [15] and (iii) viral RNA cleavage and proofreading activity [16]. Among the three subunits, PB1 is accountable the entire process of RNA synthesis, while PB2 and PA play an important in RNA replication and transcription respectively [17].

The PA protein is 716 residues long and contains two domains namely, N-terminal domain (NTD; residues 1-256) and C-terminal domain (CTD; residues 257-716) [18-20]. Residues 239-716 of PA are implicated in binding to PB1 [19]. Regions between 124-139 and 186-247 were found to be important for nuclear localization of PA [21]. The NTD of PA was found to contain endonuclease activity with residues P107, D108 and E119 found to be critical for this activity [22, 23]. Residues 493-512 and 557-574 are implicated in binding of human CLE/C14orf166 protein (hCLE) and micro chromosomal maintenance complex [24, 25]. PA was also shown to contain protease activity with Ser624 as active site [26]. Residues 100-180 are implicated in vRNA and cRNA promoter binding [26-28].

The PB1 protein is 757 residues long and consists of three domains namely, NTD (residues 1-15), RdRp catalytic domain (residues 286-483) and CTD (residues 678-757) [29]. Apart from the three domains, PB1 also contain six polymerase motifs: pre-A/F (residues 229-257), motif A (residues 296-314), motif B (residues 401-422), motif C (residues 436-449), motif D (residues 474-486) and motif E (residues 487-497) [30]. Residues 1-139 and 267-493 are implicated in complimentary RNA (cRNA) promoter binding [29]. Residues 1-83, 249-256 and 494-757 are implicated in viral RNA (vRNA) promoter binding [29]. Residues 179-297 and 458-519 are implicated in nucleotide binding [29]. PB1 also contain two nuclear localization signals (NLS) namely, NLS-1 (residues 187-195) and NLS-2 (residues 203-216) [31, 32]. Khurana et al have recently identified antigenic epitope (residues 586-599) on the PB1 protein [33].

Residues 1-37 are implicated in PA binding and residues 678-757 AA are implicated in PB2 binding [34, 35].

The PB2 protein is 759 residues long and consists of four domains namely, NTD (residues 1-247), cap-binding domain (CBD; residues 318-483), RNA binding domain (RBD; residues 535-684) and CTD (residues 686-759) [36-39]. Residues 1-37 are implicated in binding of C-terminal of PB1 [36]. Residues 363 and 404 were found to be critical for cap binding [37, 40]. Residues 701 to 753 are critical in PB2-importin-α interaction [39]. PB2 protein contains two NLS namely, NLS-1 (residues 449-495) and NLS-2 (residues 678-759) [37, 40]. Recently, Carr et al have identified mitochondrial-targeting signal (residues 1-120) on the PB2 protein with residues 7-10 were found to be critical [41].

Vaccines and antivirals are available against annual epidemics caused by IAV. But these are rendered ineffective against antiviral resistant strains of IAV emerging in the future, which is a cause of concern [12]. While attempts are underway to develop antivirals and vaccines which are universally applicable and which does not lead to resistance. To achieve this, we have to understand sequence conservation among all the IAV subtypes from different hosts. The aims of the present study are to identify the degree of conservation of RdRp subunits among all the IAV subtypes from all hosts to facilitate the identification of universally conserved sites. The mapping of the conservation scores onto the structures of the RdRp subunits proposes potential antiviral binding sites which are resistant to mutations.

Methods

Sequence analysis and protein structure

The protein sequences of the IAV PA, PB1 and PB2 were obtained from influenza virus resource of National Centre for Biotechnology Information (NCBI) [42]. Full-length sequences from all hosts and all IAV subtypes were chosen. The redundancy of the obtained sequences was removed by categorizing the clusters of sequences at 99% identity and replacing each cluster with representative sequence using the CD-HIT suite [43]. For alignment of the collected proteins sequences, MUSCLE version 3.8 [44] was used with default parameters. Multiple refinements of the obtained alignment were carried out resulting in 22-26 iterations, until no further improvement was attained. The experimental structure of the RdRp subunits were obtained from Protein Data Bank (PDB) [45] with the PDB-ID 4WSB [27].

Conservation analysis

By providing multiple sequence alignment (MSA) and experimental protein structure files as an input, conserved regions were identified using ConSurf server (http://consurf.tau.ac.il/) [46-49]. By taking evolutionary relationships between protein sequences into account, ConSurf algorithm generates resultant conservation scores. ConSurf algorithm emphasizes more on those protein sequences, which are evolutionarily distant, thus calculating conservation scores which are significant [46-49]. The resultant conservation scores are criterion scores (0, 1). The residues with score lesser than 0 indicate higher conservation, while those with score greater than 0 are variable residues [46-49]. The Bayesian algorithm is utilized to appraise the confidence intervals of calculated conservation scores [46-49]. The conservation scores given by ConSurf server are separated into scale of nine grades, which are given with the intention for visualization [46-49]. Most variable positions are placed in grade one (turquoise), intermediately conserved positions are placed in grade five (white), and most conserved positions are placed in grade

nine (maroon) [46-49].

Results

Multiple alignment of polymerase protein sequences

For the PA protein 11,966 sequences were obtained from the NCBI database, which were mainly from avian (52.6%), human (27.7%) and swine (15.2%) viruses. Following sequence clustering at 99% identity threshold, 1518 sequences remained which exhibited at least 1% sequence variance. For the PB1 protein 10,605 sequences were obtained from the NCBI database, which were mainly from avian (49.1%), human (30.1%) and swine (15.8%) viruses. Following sequence clustering at 99% identity threshold, 1148 sequences remained which exhibited at least 1% sequence variance. For the PB2 protein 11,331 sequences were obtained from the NCBI database, which were mainly from avian (52.7%), human (29.7%) and swine (15%) viruses. Following sequence clustering at 99% identity threshold, 1532 sequences remained which exhibited at least 1% sequence variance.

Conserved and variable residues

The variable and conserved residues in the IAV RdRp subunits were identified using ConSurf server [46-49] and the conservation scores were projected onto the spacefill model of the experimental structures which are illustrated in figure 1. The sequence conservation of PA, PB1 and PB2 proteins are shown in figure 2, figure 3 and figure 4 respectively.

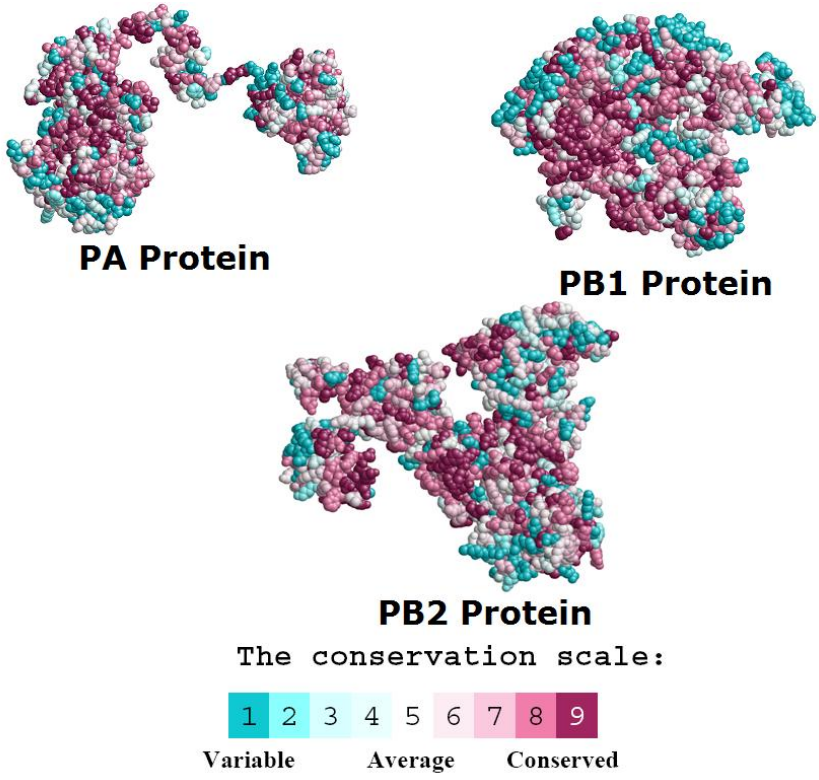

Figure 1 Spacefill representation of PA, PB1 and PB2 protein structures colour coded according to the conservation scores obtained from ConSurf server [46-49].

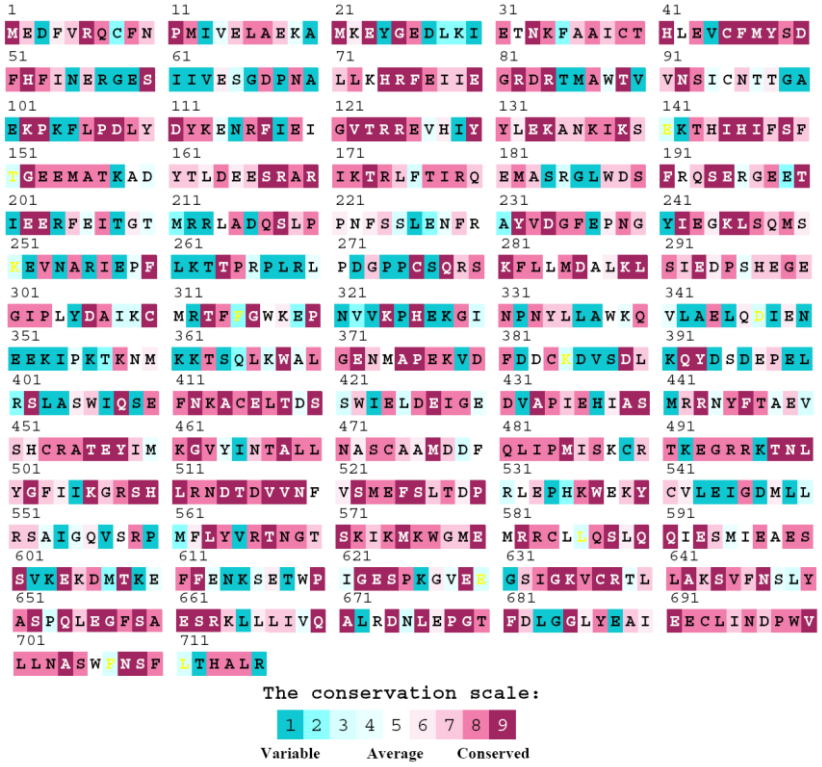

Figure 2 The amino acid sequence of the PA protein, colour coded according to the conservation scores obtained from ConSurf server [46-49].

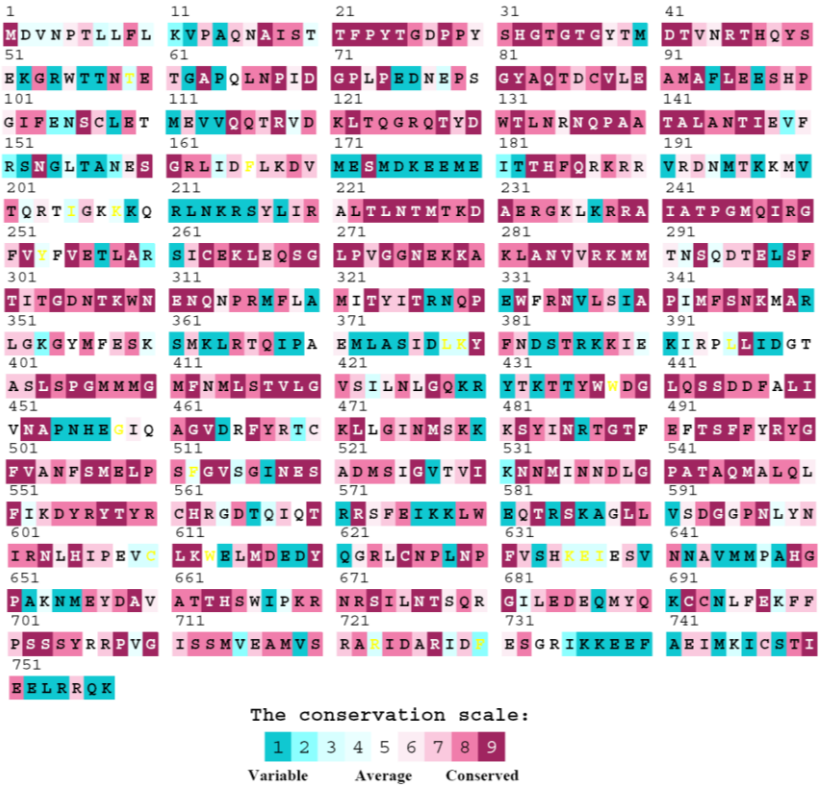

Figure 3 The amino acid sequence of the PB1 protein, colour coded according to the conservation scores obtained from ConSurf server [46-49].

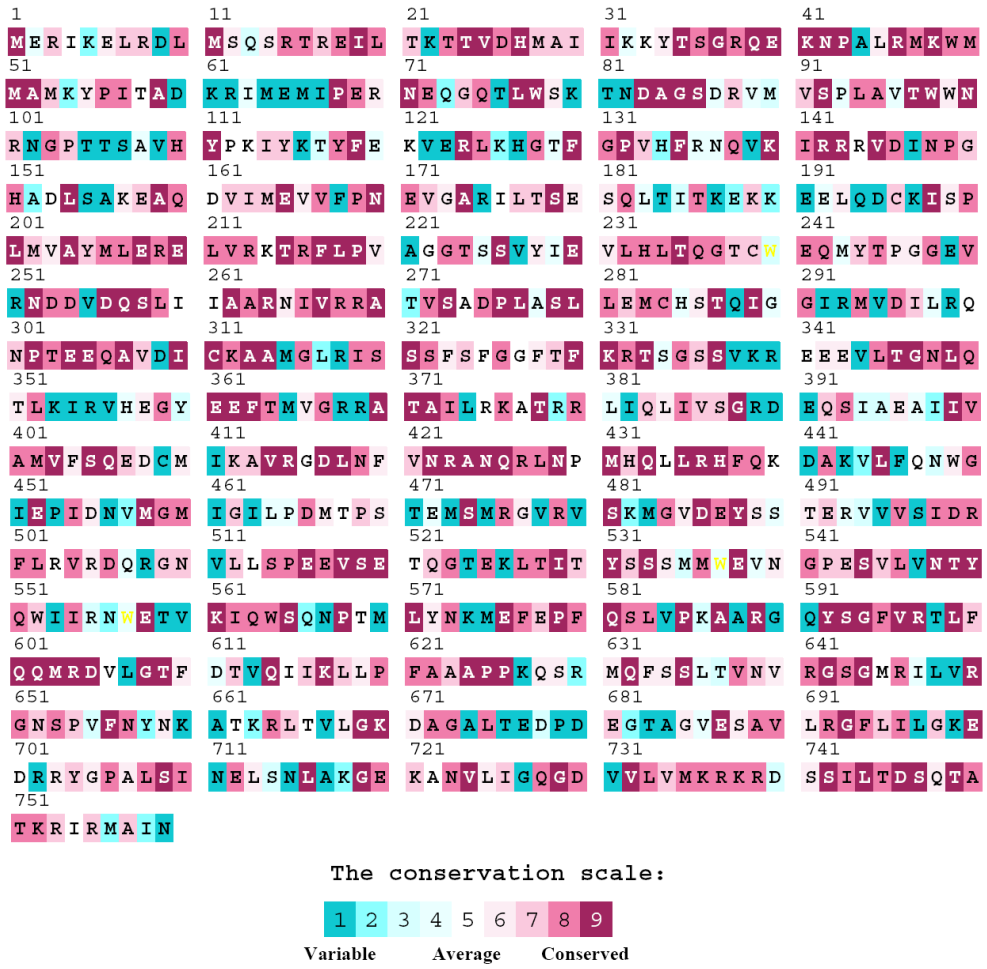

Figure 4 The amino acid sequence of the PB2 protein, colour coded according to the conservation scores obtained from ConSurf server [46-49].

Conservation scores for the PA protein were obtained between the values -0.814 (maximum conservation) and 3.746 (maximum variability) by the ConSurf server [46-49]. The sequence conservation of PA protein is shown in figure 3. In general the PA protein is intermediary conserved with 48.4% of the residues belong to grades 8-9 (conserved), while 20.1% of the residues belong to grades 1-2 (variable). Altogether, hundred and sixty seven residue positions (23.3% of total residues) were found to be highly conserved (grade 9). Residue positions 11, 21, 23, 47, 48, 84, 117, 130, 152, 191, 310, 313, 368, 392, 462, 500, 525, 530, 537, 593, 622, 635, 637, 654, 656, 657, 661, 671, 676, 679, 685 and 691 showed no amino acid variation at all among 11,966 sequences analyzed. In total, hundred and twenty seven residues (17.7% of total residues) were found to be highly variable (grade 1). The residue positions 57, 61, 66, 99, 142, 226, 269, 272, 321, 388, 394, 400 and 716 showed highest variations among all the sequences analyzed.

Conservation scores for the PB1 protein were obtained between the values -0.730 (maximum conservation) and 3.977 (maximum variability) by the ConSurf server [46-49]. The sequence conservation of PB1 protein is shown in figure 4. In general the PB1 protein is intermediately conserved

with 51% of the residues belong to grades 8-9 (conserved), while 19.2% of the residues belong to grades 1-2 (variable). Altogether, two hundred and five residue positions (27% of total residues) were found to be highly conserved (grade 9). Residues positions 1, 22-24, 33, 71, 72, 81, 82, 122, 128, 137, 138, 147, 153, 224, 239, 242, 250, 263, 285, 290, 311, 312, 326, 332, 341, 347, 380, 403, 405, 406, 410, 415, 420, 437, 441, 443, 450, 472, 482, 489, 501, 519, 526, 540, 541, 545, 551, 557, 585, 594, 597, 620, 625, 629, 649, 651, 659, 692, 697, 708 and 710 showed no amino acid variation at all among 10,605 sequences analyzed. In total, hundred and thirty residues (17.2% of total residues) were found to be highly variable (grade 1). The residue positions 52, 54, 113, 154, 179, 200, 211, 212, 317, 375, 386, 430 and 667 showed highest variations among all the sequences analyzed.

Conservation scores for the PB2 protein were obtained between the values -0.795 (maximum conservation) and 3.840 (maximum variability) by the ConSurf server [46-49]. The sequence conservation of PB2 protein is shown in figure 1. In general the PB2 protein is intermediately conserved with 48.2% of the residues belong to grades 8-9 (conserved), while 19.4% of the residues belong to grades 1-2 (variable). Altogether, one hundred and eighty six residue positions (24.5% of total residues) were found to be highly conserved (grade 9). Residues positions 1, 37, 40, 78, 83, 85, 98, 124, 132, 135, 140, 218, 219, 264, 276, 333, 337, 361, 388, 417, 538, 571, 581, 602, 604, 605, 608, 625, 642, 644 and 693 showed no amino acid variation at all among 11,331 sequences analyzed. In total, one hundred and twenty three residues (16.6% of total residues) were found to be highly variable (grade 1). The residue positions 184, 195, 249, 292, 338, 340, 355, 389, 451, 453, 508, 559, 588, 661 and 676 showed highest variations among all the sequences analyzed.

Discussion

The objective of the present study was to determine the degree of conservation of the PA, PB1 and PB2 proteins among all the isolates of influenza A viruses. The PA, PB1 and PB2 proteins from all hosts were analyzed together to facilitate the identification of universally conserved residues of potential pandemic viruses that might arise in future due to either a event of mutation or reassortment. The conserved residues detected on the PA, PB1 and PB2 proteins may have either structural importance or functional importance [50]. On the contrary, variable sites arise as a result of either adaptation or evolutionary pressure to evade the host immune system.

In the established functionally important regions of the PA protein high conservation was found, for example in the PA-PB1 binding region (residues 239-716), several highly conserved residues were found, of which residues C310, T313, W368, Q392, G462, L500, F525, P530, W537, E593, G622, K635, C637, Q654, E656, G657, E661, A671, L676, G679, G685 and E691 showed highest conservation. In the regions identified as important for nuclear localization of PA (residues 124-139 and 186-247), residues R124, R125, Y130, E133, K134, K139, F191, S194, E195, T200, E202, E203, S218, Y232, D234, I242, K245 and L246 were found to be highly conserved (grade 9). The amino acid residues critical for the endonuclease activity of PA protein (residues P107, D108 and E119), were found to be conserved. Residues P107 and D108 were found to be highly conserved (grade 9) while the residue E119 was found to be conserved (grade 8). In the regions which are implicated in PA-hCLE and PA-micro chromosomal maintenance complex binding (residues 493-512 and 557-574), residues T498-L500, G502, K506, S509-L511, L563, T567 and S571, were found to be highly conserved (grade 9) while residues

E493-R496, R508, R512, R559, Y564, R566, N568, G569, K572 and K574 were found to be conserved (grade 8). The active site of PA protein protease activity (residue Ser624) was found to be highly conserved (grade 9). In the region which is implicated in vRNA and cRNA promoter binding (residues 100-180), residues K102, P103, P107, D108, D111, F117, V122-R125, Y130, E133, K134, K139, I145-I147, S149, G152, S167, R168, R170 and T173 were found to be highly conserved (grade 9).

In the established functionally important regions of the PB1 protein, high conservation was found, for example in the cRNA promoter binding regions (residues 1-139 and 267-493) several highly conserved (grade 9) residues were identified, out of which residues , M1, F22-Y24, G33, G71, P72, G81, Y82, L122, T128, Q137, P138, V285, M290, E311, N312, T326, W332, P341, K347, Y380, L403, P405, G406, G410, L415, G420, W437, L441, S443, I450, L472, S482 and T489 showed highest conservation. In the region implicated in the vRNA promoter binding (residues 1-83, 249-256 and 494-757), there are several highly conserved residues (grade 9), of which residues M1, F22-Y24, G33, G71, P72, G81, Y82, G250, F501, E519, G526, G540, P541, Q545, F551, Y557, S585, G594, N597, Y620, C625, N629, H649, P651, A659, C692, E697, P708 and G710 were found with highest conservation. Residues K188 (grade 8), R189 (grade 7) and D193 (grade 7) were found to be conserved in the NLS1 (187-195 AA), while a lone residue G206 (grade 7) was found to be conserved in NLS-2 (residues 203-216) of PB1 protein. Residues S585, L589, L590, D593, G594, P596 and N597 were found to be conserved in the antigenic epitope (586-599 AA) on the PB1 protein [33]. In the PB1-PA binding region (residues 1-37) M1, A17, T21-T25, P28, H32, H33 and G35-G37 were found to be highly conserved residues (grade 9). In the PB1-PB2 binding region (residues 678-757), residues G681, C692, E697, S702, S703, P708, G710, R727 and I750 were found to be highly conserved residues (grade 9).

In the established functionally important regions of the PB2 protein, high conservation was found, for example in the region implicated in PB2-PB1 binding (residues 1-37), residues M1, M11, I19, T24, V25, H27, T35 and G37 were found to be highly conserved. Of the two residues (363 and 404) which were found to be critical for cap binding, only residue F363 was found to be highly conserved, while F404 was found to be intermediately conserved (grade 6). In the region which was found to be critical for PB2-importin-α interaction (residues 701-753), residues S709, L716, G719, V724, G729, V732, S742, L744, D746, S747 and T749 were found to be highly conserved (grade 9). In the NLS-1 (residues 449-495) of PB2 protein, residues E452, M458, S474, S481, D486 and Y488 were found to be highly conserved (grade 9), while G450, I454, G459, M460, D466, R476, R479, G484 and E487 were found to be conserved residues (grade 8). In the NLS-2 (residues 678-759), residues E687, G693, E700, S709, L716, G719, V724, G729, V732, S742, L744, D746, S747 and T749 were found to be highly conserved (grade 9), while A689, V690, R692, F694, I696, P706, L708, I710, E712, E720, N723, Q728, D730, L733, M735, K736, R737, K738, R739, I743, T745, A750, T751 and K752 were found to be conserved residues (grade 8). The critical residues in the region identified as mitochondrial-targeting signal (residues 7-10) showed varied degree of conservation. Residues L7 and R8 were found to be conserved (grade 7); residue L10 was found to be conserved at grade 8, while residue D9 was found to be highly variable (grade 1).

Conclusion

In culmination, this study has identified that influenza polymerase complex revealed a pattern of conserved and variable residues among all IAV hosts and subtypes. By identifying drug binding sites in

close proximity to the conserved residues in the PA, PB1 and PB2 proteins will help in designing anti-influenza drugs which are improbable to get inefficient in case of a mutation in IAV leading to drug resistant form. Moreover these anti-influenza drugs will be universally efficient against all strains of IAV from various hosts. The functions of the previously unknown highly conserved residues identified in this study should be experimentally characterized.

Abbreviations

IAV, Influenza A Virus; RdRp, RNA- dependent RNA polymerase; PA, Polymerase acidic; PB1, Polymerase basic 1; PB2, Polymerase basic 2; NCBI, National Centre for Biotechnology Information; PDB, Protein Data Bank

Acknowledgement

The author received no funding for this work.

References

1. Palese P, Shaw ML. Orthomyxoviridae: the viruses and their replication. In: Fields Virology, ed. Knipe DM and Howley PM. Philadelphia: ***Lippincott Williams & Wilkins***. 2007, 1647-1689.
2. WHO. Influenza (seasonal). ***Fact sheet.*** 2009, No. 211
3. Leibler JH, Otte J, Roland-Holst D, Pfeiffer DU, Magalhaes RS, Rushton J, Graham JP, Silbergeld EK. Industrial Food Animal Production and Global Health Risks: Exploring the Ecosystems and Economics of Avian Influenza. ***Ecohealth.*** 2009, 6(1): 58-70
4. Johnson NP, Mueller J. Updating the accounts: global mortality of the 1918–1920 ‘‘Spanish’’ influenza pandemic. ***Bull Hist Med.*** 2002, 76: 105-115
5. Simonsen L, Clarke MJ, Williamson GD, Stroup DF, Arden NH, Schonberger LB. The impact of Influenza epidemics on mortality: introducing a severity index. ***Am J Public Health.*** 1997, 87: 1944-1950
6. Cox NJ, Subbarao K. Global epidemiology of influenza: past and present. ***Annu Rev Med.*** 2000, 51: 407-421
7. Garten RJ, Davis CT, Russell CA, Shu B, Lindstrom S, Balish A, Sessions WM, Xu X et al. Antigenic and genetic characteristics of swine-origin 2009 A (H1N1) influenza viruses circulating in humans. ***Science.*** 2009, 325(5937): 197-201
8. Smith GJD, Dhanasekaran VK, Justin B, Samantha JL, Michael W, Oliver GP, Siu KM et al. Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. ***Nature.*** 2009, 459(7250): 1122-1125
9. Horimoto T, Kawaoka Y. Influenza: Lessons from past pandemics, warnings from current incidents. ***Nature Reviews Microbiology.*** 2005, 3(8): 591-600
10. Lin YP, Shaw M, Gregory V, Cameron K, Lim W, Klimov A, Subbarao K, Guan Y, Krauss S, Shortridge K, Webster R, Cox N, Hay, A. Avian-to-human transmission of H9N2 subtype influenza A viruses: Relationship between H9N2 and H5N1 human isolates. ***Proc Natl Acad Sci***. 2000, 97(17):

9654-9658
11. Schrauwen EJA, Fouchier RAM. Host adaptation and transmission of Influenza A viruses in mammals. ***Emerging Microbes and Infections.*** 2014, 3: e9
12. Das K. Antivirals targeting influenza A virus. ***J Med Chem.*** 2012, 55: 6263-6277
13. Stubbs TM, Velthuis AJWT. The RNA-dependent RNA polymerase of the influenza A virus. ***Future Virol***. 2014, 9(9): 863-876
14. Ishihama A, Nagata K. Viral RNA polymerases. ***CRC Crit. Rev. Biochem.*** 1988, 23:27-76
15. Pritlove DC, Poon LL, Fodor E, Sharps J, Brownlee GG. Polyadenylation of influenza virus mRNA transcribed in vitro from model virion RNA templates: requirement for 50 conserved sequences. ***J Virol.*** 1998, 72: 1280-1286
16. Ishihama A, Mizumoto K, Kawakami K, Kato A, Honda A. Proofreading function associated with the RNA-dependent RNA polymerase from influenza virus. ***J Bio. Chem***. 1986, 261: 10417-10421
17. Mahy BWJ. Mutants of influenza virus. In: Genetics of influenza viruses, ed. Palese P and Kingsbury DW. Wien: ***Springer Verlag.*** 1983, 192-253
18. He X, et al. (2008) Crystal structure of the polymerase PA(C)–PB1(N) complex from An avian influenza H5N1 virus. ***Nature.*** 454, 1123-1126
19. Obayashi E, et al. The structural basis for an essential subunit interaction in influenza virus RNA polymerase. Nature 2008, 454: 1127-1131
20. Zhao C, et al. Nucleoside monophosphate complex structures of the endonuclease domain from the influenza virus polymerase PA subunit reveal the substrate binding site inside the catalytic center. ***J Virol.*** 2009, 83: 9024-9030
21. Nieto A, de la Luna S, Barcena J, Portela A, Ortin J. Complex structure of the nuclear translocation signal of influenza virus polymerase PA subunit. ***J Gen Virol*** 1994, 75(Pt1): 29-36
22. Yuan P, Bartlam M, Lou Z, et al. Crystal structure of an avian influenza polymerase PA(N) reveals an endonuclease active site. ***Nature.*** 2009, 458: 909-913
23. Dias A, Bouvier D, Crepin T, et al. The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. ***Nature.*** 2009, 458: 914-918
24. Huarte M, Sanz-Ezquerro J J, Roncal F, et al. PA subunit from influenza virus polymerase complex interacts with a cellular protein with homology to a family of transcriptional activators. ***J Virol.*** 2001, 75: 8597-8604
25. Kawaguchi A, Nagata K. *De novo* replication of the influenza virus RNA genome is regulated by DNA replicative helicase, MCM. ***Embo J.*** 2007, 26: 4566-4575
26. Hara K, Schmidt FI, Crow M, Brownlee GG. Amino acid residues in the N-terminal region of the PA subunit of influenza A virus RNA polymerase play a critical role in protein stability, Endonuclease activity, cap binding, and virion RNA promoter binding. ***J Virol.*** 2006, 80:7789-7798
27. Deng T, Vreede FT, Brownlee GG. Different De Novo Initiation Strategies Are Used by Influenza Virus RNA Polymerase on Its cRNA and Viral RNA Promoters during Viral RNA Replication. ***J Virol.*** 2006, 80: 2337-2348
28. Maier H, Kashiwagi T, Hara K, Brownlee GG. Differential role of the influenza A virus polymerase PA subunit for vRNA and cRNA promoter binding. ***Virology.*** 2008, 370: 194-204

29. Binh NT, Wakai C, Kawaguchi A, Nagata K. The N-terminal region of influenza virus polymerase PB1 adjacent to the PA binding site is involved in replication but not transcription of the viral genome. ***Front Microbiol.*** 2013, doi: 10.3389/fmicb.2013. 00398
30. Pflug A, Guilligay D, Reich S, Cusack S. Structure of influenza A polymerase Bound to the viral RNA promoter. ***Nature.*** 2014, 516: 355-360
31. Hutchinson E, Orr O, Man LS, Engelhardt O, Fodor, E. Characterisation of the interaction between the influenza A virus polymerase subunit PB1 and the host nuclear import factor Ran Binding Protein 5. ***J Gen Virol.*** 2011, 92: 1859-1869
32. Nath ST, Nayak DP. Function of two discrete regions is required for nuclear localization of polymerase basic protein 1 of A/WSN/33 influenza virus (H1N1). ***Mol Cell Biol.*** 1990, 10: 4139-4145
33. Khurana S, Suguitan ALJ, Rivera Y, Simmons CP, Lanzavecchia A, et al. Antigenic fingerprinting of H5N1 avian influenza using convalescent sera and monoclonal antibodies reveals potential vaccine and diagnostic targets. ***PLoS Med.*** 2009, 6(4):e1000049
34. Sugiyama K et al. Structural insight into the essential PB1-PB2 subunit contact of the influenza virus RNA polymerase. ***EMBO J.*** 2009, 28: 1803-1811
35. Guilligay D et al. The structural basis for cap binding by influenza virus polymerase subunit PB2. ***Nat Struct Mol Biol.*** 2008, 15: 500-506
36. Kuzuhara T et al. Structural basis of the influenza A virus RNA polymerase PB2 RNA-binding domain containing the pathogenicity-determinant lysine 627 residue. ***J Biol Chem.*** 2009, 284: 6855-6860
37. Tarendeau F et al. Structure and nuclear import function of the C-terminal domain of influenza virus polymerase PB2 subunit. ***Nat Struct Mol Biol.*** 2007, 14: 229-233
38. Fechter P, Mingay L, Sharps J, Chambers A, Fodor E, Brownlee GG. Two aromatic residues in the PB2 subunit of influenza A RNA polymerase are crucial for cap binding. ***J Biol Chem.*** 2003, 278: 20381-20388
39. Fontes MRM, Jan TTD, Brinkworth RI, Kobe B. Structural basis for the specificity of bipartite nuclear localization sequence binding by importin-α. ***Journal of Biological Chemistry.*** 2003, 278(30): 27981-27987
40. Mukaigawa, J, Nayak, DP. Two signals mediate nuclear localization of influenza virus (A/WSN/33) polymerase basic protein 2. ***J Virol.*** 1991, 65: 245-53
41. Carr SM, Carnero E, Garcia-Sastre A, Brownlee GG, Fodor E. Characterization of a mitochondrial-targeting signal in the PB2 protein of influenza viruses. ***Virology.*** 2006, 344: 492-508
42. Bao Y, Bolotov P, Dernovoy D, Kiryutin B, Zaslavsky L, Tatusova T, Ostell J, Lipman, D. The influenza virus resource at the National Centre for Biotechnology Information. ***J Virol.*** 2008, 82: 596-601
43. Huang Y, Niu BF, Gao Y, Fu LM, Li WZ. CD-HIT Suite: a web server for clustering and comparing biological sequences. ***Bioinformatics***. 2010, 26(5): 680-682.
44. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high through-put. ***Nucleic Acids Res***. 2004, 32: 1792-1797
45. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. ***Nucleic Acids Res***. 2000, 28(1): 235-242

46. Landau M, Mayrose I, Rosenberg Y, Glaser F, Martz E, Pupko T, Ben-Tal N. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. ***Nucl Acids Res***. 2005, 33: W299-W302
47. Glaser F, Pupko T, Paz I, Bell RE, Bechor D, Martz E, Ben-Tal N. ConSurf: Identification of Functional Regions in Proteins by Surface-Mapping of Phylogenetic information. ***Bioinformatics***. 2003, 19:163-164
48. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal, N. ConSurf 2010: calculating Evolutionary conservation in sequence and structure of proteins and nucleic Acids. ***Nucleic Acids Res***. 2010 38: W529-W533
49. Celniker G, Nimrod G, Ashkenazy H, Glaser F, Martz E, Mayrose I, Pupko T, Ben-Tal, N. ConSurf: Using Evolutionary Data to Raise Testable Hypotheses about Protein Function. ***Israel Journal of Chemistry***. 2013, 53(3-4): 199-206
50. Schueler-Furman O, Baker D. Conserved residue clustering and protein structure prediction. ***Proteins***. 2003, 52: 225-235