Bioinformation
www.bioinformation.net
Volume 5
Issue 1
open access
Hypothesis

Structure function studies on different structural domains of nucleoprotein of H1N1 subtype

Parveen Salahuddin[1], Asad U. Khan[1, 2*]

[1]Distributed Information Sub-Centre; [2]Interdisciplinary Biotechnology Unit; A.M.U. Aligarh, 202002, India; Asad U Khan Email: asad.k@rediffmail.com; Phone: +91 571 2723088; Fax: +91 571 2721776 *Corresponding Author

Received April 04, 2010; accepted June 08, 2010, published June 24, 2010

Abstract:
Recent 2009 flu pandemic is a global outbreak of a new strain of influenza A virus subtype H1N1. The H1N1 virus has crossed species barrier to human and apparently acquired the capability to transmit this disease from human to human. The NP is a multifunctional protein that not only encapsidates viral RNA (vRNA), but also forms homo-oligomer and thereby maintains RNP structure. It is also thought to be the key adaptor for virus and host cell interaction. Thus, it is one of the factor that play a key role in the pathogenesis of influenza A virus infection. Therefore, to understand the cause of pathogenicity of H1N1 virus, we have studied the structure-function relationship of different domains of NP. Our results showed that conservative mutation in NP of various strains were pathogenic in nature. However, non-conservative mutation slightly abrogated oligomerization and was therefore less pathogenic. Our results also suggest that beside tail and body domain, head domain may also participate in an oligomerization process.

Keywords: antigenic site; nucleoprotein; swine flu; H1N1; pathogenicity.

Background:
Influenza A virus is a contagious respiratory illness causing annual epidemics and occasional pandemics killing million of people world-wide. Therefore, it possesses significant challenge to both human and animal health. Recent 2009 flu pandemic is a global outbreak of a new strain of influenza A virus subtype H1N1 also commonly known as "swine flu" or "mexican flu". The virus has emerged because of genetic reassortment of endemic strains of human, avian flu and swine flu **[1-2]**. The H1N1 virus has crossed species barrier to human and apparently acquired the capability to transmit this disease from human to human **[3-4]**. Influenza A virus belongs to the family of *orthomyxoviridae*. It is a lipid enveloped virus containing a negative-single stranded sense RNA genome that are organized into eight separate segments which code for eleven proteins **[5-6].**

The nucleoprotein (NP) is located on segment 5 of the segmented negative-strand RNA genome. The NP is the major component of the RNP, which is complex comprising of RNA, NP and RNA polymerase. The NP is a multifunctional protein that encapsidates viral RNA (vRNA) **[7]** as well as forms homo-oligomer and thereby maintains RNP structure **[8]**. It is also believed to be the key adaptor for virus and host cell interaction **[9]**. Recent study suggests a direct interaction of NP with polymerase proteins may be involved in regulating the switch of viral RNA synthesis from transcription to replication [**10].** The nucleoprotein is an entirely an alpha helical protein that has head domain, body domain and tail. Oligomerization of the influenza virus nucleoprotein is mediated by a tail loop that is inserted into body domain of neighbouring NP molecule **[11].** Thus, both tail and body domain participates in an oligomerization process. Proper self-association is important for the biological activities of nucleoprotein, because the R416A mutant was shown to be defective in RNA binding and also in viral RNA synthesis **[12]**.The core protein nucleoprotein do not mutate as fast as surface protein HA, therefore this protein can serve as a suitable target for both drugs and antiviral vaccines developments. In view of above, we have identified mutations in different structural domains of NP and we have deduced the effect of mutation on oligomerization. Further, we have also analyzed mutations in the antigenic site. These studies revealed the structure-function relationship of NP and cause of the antigenic variations and pathogenicity.

Methodology:
The sequence of NP of influenza A virus A/New Mexico/04/2009(H1N1), from human was analyzed in this study. The NP is available at NCBI **[13]** with accession number [GenBank: ACR08594]. Protein protein blast **[14]** was performed at NCBI server. This sequence was aligned with other homologous sequence using Clustal W **[15]** multiple alignment tool using weight matrix GONNET. The multiple aligned sequences were analyzed for mutations in the amino acids at specific positions in different domains of NP. Hydrophobicity values were obtained from the program ProtScale at ExPASy choosing Kyte & Doolittle hydrophobicity scale **[16].** Antigenic sites in nucleoprotein were predicted by program "Antigenic" **[17].**

Discussion:
Structure-function studies of different domains of NP:
The NP is a multifunctional protein that encapsidates vRNA through its RNA binding domain (residues 1-181) and form homo-oligomers. The nucleoprotein has NPI (residues189-358) and NPII (residues 371-465) binding domains important for mediating NP-NP contacts and thereby undergoes oligomerization **[9].** It also has a PB2 binding domains [PB2-I (1-161), PB2-2(255-340) PB2-3(340-465][9]. These domains of NP interact with PB2 subunit of RNA polymerase.

The NP sequence of strain ACR08594 was compared with other strains of NP (**Table 1 see supplementary material**) showing 99.80% to 99.40% identity. Thirteen sequences of NP of different strains were selected. Structural-functional relationship of different domains of NP was deduced in this study. Our data of ACQ76390 strain shows that there was a mutation Val217 to Ile 217 in the head region of helix 4 (**see supplementary material Table 1 [18])**. Since, this is a conservative mutation that occurred at NP1 binding domain, it did not affect NP-NP contact and subsequent oligomerization **[9].** Therefore, it did not affect oligomerization of NP. Other strain ACQ76337 showed mutation in the

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 5(1): 28-30 (2010)

body region of helix 4 in RNA binding region, which also overlaps PB2 domain. This mutation causes in the substitution of Leu133 by an Ile133. This is also a conservative substitution. Hence, this strain binds RNA normally and interacted with PB2 normally. These strains were pathogenic in nature.

The strain ACQ76327 has a mutation in the body of helix 16 in overlapping NPII and PB2 domains. Mutation causes in the substitution of hydrophobic Val 343 by a polar Thr343. The mutation is non-conservative, where microenvironment of both NP II and PB2 domains was changed. Therefore, this strain formed oligomer weakly and showed reduces interaction with PB2. These observations is supported by the findings that oligomerization is important for NP function because R 416 A mutant was found defective in RNA binding and viral RNA synthesis **[12]**. There is also direct interaction of NP with PB2 **[10].** These two strains were weakly pathogenic.

In NP of ACS94528 strain, mutation occurred in the head region of helix 8 in the NPII binding domain. In this strain, there was a conservative substitution of Ile 189 by Val189. This leads to normal formation of NP-NP homo-oligomer **[9]** and hence this strain was pathogenic in nature. Similarly, the NP of strain ACQ76406 showed substitution Val 217 by Ile 217 in the head region of helix 9 and another substitution Val343 by Leu343 in the body region of helix 16. These mutations are conservative mutations. Therefore, this strain formed NP-NP oligomer normally and interacted with PB2 normally. The NP of other strain ACR67140 has a mutation located in the helix 10 of head region in the PB2 domain. This is a also a conservative substitution, Ala234 by Val234, and it did not affect oligomerization or interaction with PB2. These observations are supported by the findings that self-association or oligomerization is important for NP function **[12]** and there is also direct interaction of NP with PB2 **[10].**

The NP of strain ACQ63285 carried mutation in the head region of helix 11 in NPII and PB2 domains, where hydrophobic Ala 251 is substituted by polar Thr 251. This is a non-conservative substitution where hydrophobicity decreased in overlapping NPII and PB2 domains. Therefore, this leads to decrease in the formation of NP-NP homo-oligomer and poor interaction with PB2. Similarly, the NP of strains ACS14690 and ACQ76377 showed mutation in the head region of helix 19. The polar Ser450 is substituted by an amide derivative of acidic amino acid Asn450. These strains showed loss of NP-NP oligomer formation and interaction with PB2. All of these strains were weakly pathogenic. These results imply that self association is important for NP function **[12].** Furthermore, mutational study has shown that mutant NP has a reduced capability to interact with the polymerase complex and that this NP-polymerase interaction is responsible for making vRNPs switch from mRNA to cRNA synthesis **[19]**. A mutation in the body region of helix 12 in the NP and PB2 binding domains is seen in ACR67161 strain. The hydrophobic Leu283 is replaced by a cyclic imino acid Pro283. Since Leu 283 is a strong helix former, whereas Pro283 is a helix breaker. This leads to alteration of α-helix 12 and consequently overall conformation of nucleoprotein was slightly changed. Hence, it formed homo-oligomer weakly and showed poor interaction with PB2 **[12, 10].** This strain was weakly pathogenic.

In NP of eight strains, Ile 373→Thr 373 mutations were found conserved (**see supplementary material Table 2**). In all of these strains, hydrophobic Ile is substituted by polar Thr. This is a non-conservative mutation. Besides this, six strains also showed a new mutation Val100→Ile100 which was found conserved in several strains (**see supplementary material Table 2**).These results imply that due to the presence of multiple mutations in the NP of these strains (**see supplementary material Table 2**), the conformation of NP was altered significantly. Hence these strains formed homo-oligomer weakly **[12]** and interacted with PB2 poorly **[10].** This leads to some loss of RNP structure as well. Consequently, it inhibited the switch of genomic vRNA synthesis from transcription to replication. These strains showed weak pathogenicity. None of the mutations we observed occurred in tail-loop, which play an important role in oligomerization by inserting inside a body domain of neighbouring molecule. Our results showed non-conservative mutation in the head domain causes in the inhibition of oligomerization. This implies that head domain may also participate in an oligomerization process. Strains carrying conservative mutations showed that NP domains have native like structure, therefore these strains were virulent. Contrary to this, strains possessing non-conservative mutations in domains showed a slight loss of structure. Therefore, these strains were weakly pathogenic.

Antigenicity:
The computer program "Antigenic" predicted 14 antigenic segments in NP of ACR08594 strain (See **supplementary material - Table 3**). This antigenic profile was found conserved in all strains except ACR77442, ACQ76384 and ACR67140. The NP of these strains namely ACR77442 and ACQ76384 showed mutation Val183→Met183 and Val352→Met352 respectively. Therefore, corresponding antigenic segments 179-189 and 349-356 were missing. Conversly, ACR67140 strain showed mutation Ala 234→Val234 and consequently there was a gain of an antigenic site 231-236. These are major source of antigenic variations in NP.

Conclusion:
These results suggest that head domain of NP may also participate in an oligomerization process. These results also conclusively showed that conservative mutation in NP is more pathogenic than non-conservative mutation. Thus, our hypothesis confirms the structure-function relationship of NP. Loss or gain of an antigenic site represents a source of antigenic variations in NP.

Acknowledgements:
This work was supported by the DBT grant sanction no. BT/PR7507/BID/07/201/2006 to AUK. Authors are also grateful to Prof. M Saleemuddin for providing the internal fund for this project.

References:
[1] AU Khan, *Bioinformation* **4**:52 (2009) [PMID: 20011154]
[2] P Salahuddin & AU Khan, *Bioinformation* **4**:30 (2009) [PMID: 20011150]
[3] D Butler, *Nature* **458**:1082 (2009) [PMID: 19407756]
[4] I Cohen & M Enserink, *Science* **324**:572 (2009) [PMID: 19407164]
[5] D Kobasa & Y Kawaoka, *Curr Mol Med* **5**:791 (2009) [PMID: 16375713]
[6] JC Obenauer *et al*. *Science* **311**:1576 (2006) [PMID: 16439620]
[7] M Kobayashi *et al*. *J Virol*. **68**:8433 (1994) [PMID: 7966640]
[8] EN Prokudina-Kantorovich NP Semenova,*Virology* **223**:51(1996)[PMID: 8806539]
[9] A Portela & P Digard, *J Gen Virol* **83**:723 (2002) [PMID: 11907320]
[10] SK Biswas *et al*. *J Virol* **72**:5493 (1998) [PMID: 9621005]
[11] Q Ye *et al*. *Nature* **444**:1078(2006) [PMID: 17151603]
[12] D Elton *et al*. *J Virol* **73**: 7357 (1999) [PMID: 10438825]
[13] http://www.ncbi.nlm.nih.gov/
[14] http://www.ncbi.nlm.nih.gov/BLAST/
[15] http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_clustalw.html
[16] http://www.expasy.org/tools/protscale.html
[17] http://liv.bmc.uu.se/cgi-bin/emboss/antigenic
[18] K Ng *et al*. *Sci China C Life Sci*. **52**:439(2009) [PMID: 19471866]
[19] I Mena *et al*. *J Virol*.**73**:1186(1999) [PMID: 9882320]

Edited by P. Kangueane
Salahuddin & Khan, Bioinformation, 5 (1) 25-30, 2010

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 5(1): 28-30 (2010)

Bioinformation **Volume 5**
www.bioinformation.net Issue 1 **Hypothesis**
open access

Supplementary material:

Table 1: Nucleoprotein sequence dataset from human source

Strain	Gene Accession No	Country
A/New Mexico/04/2009(H1N1)	ACR08594	Mexico
A/Nebraska/02/2009(H1N1)	ACQ76390	U.S.A.
A/Kansas/03/2009(H1N1)	ACQ76337	U.S.A.
A/Indiana/09/2009(H1N1)	ACQ76377	U.S.A.
A/South Carolina/09/2009(H1N1)	ACQ76327	U.S.A.
A/New York/58/2009(H1N1)	ACS94528	U.S.A.
A/Colorado/03/2009(H1N1)	ACQ76406	U.S.A.
A/New York/3189/2009(H1N1)	ACR77442	U.S.A.
A/Missouri/01/2009(H1N1)	ACR67140	U.S.A.
A/Ohio/07/2009(H1N1	ACQ76384	U.S.A.
A/Ohio/07/2009(H1N1)	ACQ63285	U.S.A.
A/New York/3210/2009(H1N1)	ACS14690	U.S.A.
A/South Dakota/05/2009(H1N1)	ACR67161	U.S.A.

Table 2: Amino acid mutations in Nucleoprotein of different strains at specific positions and their location in different domains

Gene Acession	Mutation at specific position	Location of mutation in domain	Hydrophobicity change (Kyte & Doolittle)
ACQ76390	V217→I217	Head	-1.944 →-1.911
ACQ76337	L133→ I133	Body	+0.978→+1.056
ACQ76377	S450 →N450	Head	-1356→-1.656
	V363→I363	-	-0.022→+0.011
ACQ76327	V343 →I343	Body	+0.224→+0.278
	I 373→T373	-	-0.856→-1.433
ACS94528	V100→I100	-	1.144→-1.111
	I373→t373	-	-0.856→-1.433
	I189→V189	Head	+1.224→+1.211
ACQ76406	V217→I217	Head	-1.914→-1.911
	V343→I343	Body	+0.244→+0.278
	I373→T373	-	-0.856→-1.433
ACR77442	V100→I100	-	-1.144→-1.111
	V183→M183	-	+0.967→+0.711
	I373→T373	-	-0.856→-1.433
ACR67140	A234→V234	Head	-0.233→+0.033
	I373→T373	-	-0.856→-1.433
ACQ76384	V100→I100	-	-1.144→-1.111
	I373→T373	-	-0.856→-1.433
	V352→M352	-	-1.167→-1.422
ACQ63285	V100 →I100	-	-1.144→-1.111
	A251→T251	Head	-1.467→-1.744
	I373→T373	-	-0.856→-1.433
ACS14690	V100→I100	-	-1.144→-1.111
	I373→T373	-	-0.856→-1.433
	S450→N450	Head	-1.356→-1.656
ACR67161	L283→P283	Body	+1.756→+1.156
	I373→T373	-	-0.856→-1.433

Table 3: Predicted antigenic sites in human Nucleoprotein (A/New Mexico/04/2009(H1N1)) and in other strains

Amino acid position	Antigenic Segment	Total no of amino acid residues	Score
254->288	EDLIFLARSALILRGSVAHKSCLPACVYGLAVASG	35	1.249
325->346	KSQLVWMACHSAAFEDLRVSSF	22	1.133
296->316	YSLVGIDPFKLLQNSQVVSLM	21	1.126
403->417	AGQISVQPTFSVQRN	15	1.122
473->479	NPIVPSF	7	1.090
38->50	RFYIQMCTELKLS	13	1.088
179->189	AGAAVKGVGTI	11	1.082
349->356	GKKVIPRG	8	1.075
64->71	ERMVLSAF	8	1.074
107->113	ELILYDK	7	1.066
360->366	TRGVQIA	7	1.053
151->156	TRALVR	6	1.052
222->229	MCNILKGK	8	1.046
132->138	GLTHIMI	7	1.041
ACR77442	AGAAVKGVGTI	-	-
179->189	(Antigenic Segment missing)		
ACQ76384	GKKVIPRG	-	-
349->356	(Antigenic Segment missing)		
ACR67140	QTAVQR	-	-
231->236	(Gain of Antigenic Segment)		