

## Identification of a casein kinase II phosphorylation domain in NS1 protein of H5N1 influenza virus

Tamanna Anwar and Asad U Khan\*

Interdisciplinary Biotechnology Unit Aligarh Muslim University, Aligarh 202002 India;

\*Asad U Khan - Email: huzzi99@hotmail.com; Phone: 091 571 2723088; Fax: 091 571 2721776; \*Corresponding author

received June 19, 2007; revised September 05, 2007; accepted September 13, 2007; published online October 03, 2007

### Abstract:

Influenza virus causes febrile respiratory illness. The infection results in significant mortality, morbidity and economic disruption. In this bioinformatics study, we used the NS1 (the conserved nonstructural) protein of influenza A virus to demonstrate its role in infectivity. Our *in silico* study revealed a new Casein kinase II (CKII) phosphorylation domain at position 151-154. This domain was formed due to the mutation at position 151 (T151I). Moreover, considerable difference in the secondary structure of this protein due to mutation was also reported. It is also confirmed by contact residue analysis that the changes in secondary structure are due to mutations.

**Keywords:** nonstructural protein 1 (NS1); amino acid (AA); mutation; casein kinase II (CKII) phosphorylation

**Availability:** Prediction results are at <http://www.geocities.com/amubioinfo/InfluenzaAVirus.htm>.

### Background:

The importance of influenza viruses as worldwide pathogens in humans, domestic animals, and poultry is well known. Influenza A viruses belong to the family Orthomyxoviridae having a segmented genome of single stranded negative-sense RNA. [1] New strains of influenza viruses emerge each year that have been isolated from a variety of animals, including humans, pigs, horses, sea mammals, and birds. [2] Infection of poultry with highly pathogenic avian influenza virus can be devastating in terms of flock morbidity and mortality, economic loss and social disruption. [3] Influenza pandemics occur from genetic reassortment between avian and human influenza viruses or alternatively by the direct adaptation of an avian influenza viruses to human influenza virus which may cause human-to-human transmission. [2]

The NS gene is the smallest segment of the influenza A virus genome. It is one of the proteins which acts as virulent factor of influenza A virus. It regulates the splicing and translation of mRNA. [4] NS1 also regulates numerous cellular functions during influenza virus infection by binding to polyadenylated mRNAs, inhibiting nuclear export [5, 6, 7]; binding to small nuclear RNAs (snRNA), specifically to key components of the spliceosome, blocking pre-mRNA splicing [8, 9, 10, 11] and inhibiting the polyadenylation of host cell mRNA and interacting with several host cell proteins. [5, 12, 13] The NS1 protein exerts its proviral functions in part by binding free RNA (dsRNA/mRNA) thereby preventing the activation of cellular defense mechanisms and reducing the translation efficiency of cellular mRNAs for the benefit of viral

replication. [14] The NS1 protein has two functional domains which facilitate its binding with RNA. One of the domains is called RNA-binding domain, found at position 19 to 38 amino acids which binds to poly (A) sequences in mRNAs. The other domain, an effector domain found at position 134 to 161 amino acids which interacts with cellular proteins to inhibit nuclear mRNA. [7] These domains are highly conserved within the NS1 gene [11], suggesting that NS1 is evolutionarily conserved. It was recently shown that the NS1 protein plays a role in virulence of the highly pathogenic H5N1 influenza virus strains via a glutamic acid at position 92, which confers on the virus resistance to antiviral cytokines. [15]

Casein kinase II (CKII) is a ubiquitous, highly pleiotropic and constitutively active Ser/Thr protein kinase implicated in neoplasia, cell survival, and virus infection. [16, 17] CKII is critically involved in the budding process of influenza viruses. [18] It was demonstrated earlier that the phosphorylation of the P protein of vesicular stomatitis virus by cellular CKII is essential for its activity in viral transcription. [19] Recent *in vitro* studies have demonstrated that CKII converts the inactive unphosphorylated form of P (P0) to an active phosphorylated form P1, after phosphorylation at two serine residues, Ser-59 and Ser-61. [19]

At present, this strain (H5N1) of influenza virus does not have ability to cause human to human infection. Nevertheless, this ability can be acquired by this virus

due to antigenic shift. This may be, either by mutation or by recombination of genetic material with a human influenza virus. [20] Mutations in influenza virus proteins not only involve change in polarity or hydrophobicity but also the propensity of each amino acid residue to stabilize the secondary structure. [21] In view of the present background we have undertaken this study to analyze mutations and changes associated within the Non-structural protein (NS1) of Influenza A H5N1 virus.

### Methodology:

Protein NS1 of Influenza A Virus (Bird flu virus) sub-type A/Hatay/2004/(H5N1) from chicken was analyzed in the present study. This protein sequence is available at NCBI with accession number [GenBank: CAJ01906].

Protein protein Blast [22] was performed at the NCBI server and the sequences of NS1 with at least 97% similarity of the sub-type H5N1 only was sorted out. The sequences that were selected for the study are:

- (1) A/chicken/VietNam/KG-076/2004(H5N1), [AAV54500]
- (2) A/chicken/Viet Nam/LD-080/2004(H5N1), [AAV54502]
- (3) A/chicken/Viet Nam/TG-023/2004(H5N1), [AAV54490]
- (4) A/chicken/Vietnam/27/2003(H5N1), [ABC68585].

The above sequences of the data set were taken to compare with A/Hatay/2004/(H5N1) for the analysis of mutations, using our tool, MAP MUTATION. [23] Hydrophobicity values were taken out using the tool protscale at expasy server [24] choosing the Kyte & Doolittle hydrophobicity scale. The secondary structures of the proteins were predicted using secondary structure prediction program, NNPREDICT. [25] Moreover, domains or motifs among all protein sequences including A/Hatay/2004/(H5N1) were searched, using ScanProsite at ExPASy Server. [26] Motifs with high probability of occurrence were also included in the search. A program, Residue Contacts [27] was used to look for the contacts of residue 33 in RNA-binding domain structure (PDBID: 1NS1) at distance threshold set to 3.4 Angstroms. The Rasmol program [28] was used to draw the modifications in structure of NS1.

### Result:

#### Sequence analysis and Secondary Structure Prediction

On comparing sequence of NS1 protein of A/Hatay/2004/(H5N1) with the other homologous sequences described in method section, by using the program MAP MUTATION, it was observed that mutations at positions V81A, I151T, I171N and P223S were common in all the strains used in this study (Table 1 in supplementary material).

Mutation V81A is responsible for change in secondary structure from coil to helix at this position. T151I involves

the change of hydrophobic to hydrophilic AA. Secondary structure was also changed from strand to coil at this position. This mutation (I151T) was reported at the Casein kinase II (CKII) phosphorylation site. The mutation at I171N changed the amino acid from hydrophilic to hydrophobic which changed the secondary structure change from helix to strand. P223S was a change from an uncharged to hydrophilic AA. In strains A/chicken/VietNam/KG-076/2004, A/chicken/Viet Nam/LD-080/2004 and A/chicken/VietNam/TG-023/2004 mutation A132T was a change from hydrophilic to hydrophobic AA that altered the secondary structure from coil to helix. Mutation L33I in A/chicken/Viet Nam/LD-080/2004 involved change of hydrophobic residues. In A/chicken/VietNam/TG-023/2004 mutation E147K involved change of hydrophilic AA, which altered the secondary structure from coil to helix. In A/chicken/Vietnam/27/2003, G148E mutation, glutamic acid that is hydrophilic AA was replaced by hydrophobic glycine and the secondary structure was altered from helix to coil. Secondary structure prediction results are given in the additional file NS1 at <http://www.geocities.com/amubioinfo/InfluenzaAVirus.htm>.

#### Residue Contact Analysis

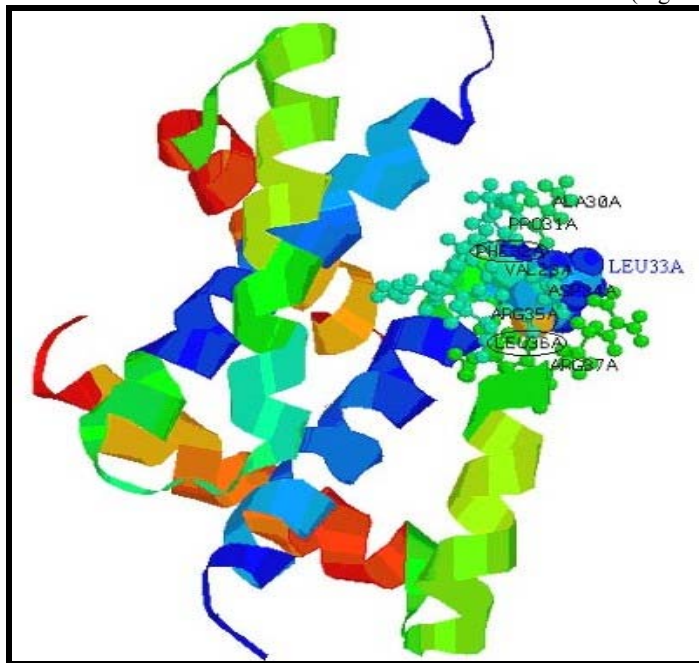
Analysis of residue contacts in rna-binding domain structure (1NS1) at position 33 where the mutation I33L in A/chicken/Viet Nam/LD-080/2004 is observed, shows that the AAs VAL23, ALA30, PRO31, PHE32, ASP34, ARG35, LEU36 and ARG37 (Additional file RC at <http://www.geocities.com/amubioinfo/InfluenzaAVirus.htm>) are binding with LEU33 (Figure 1).

#### Domain/Motif Search

Different domains that were found in NS1 are given in Table 2 (in supplementary material). Mutation T151I is reported at the Casein kinase II phosphorylation site. Residues from 151 to 154 code for CKII phosphorylation domain (TvgE). The most important finding that we observed was the identification of a new CKII phosphorylation domain. This domain was formed in A/Hatay/2004 due to the mutation from ILE151 in all other strains of the dataset to THR151 in A/Hatay/2004. Thus, there is an introduction of new CKII phosphorylation site in A/Hatay/2004.

Further mutational analysis revealed that these mutations were accompanied with changes in hydrophobicity and secondary structure (Table 1 in supplementary material). Almost all the mutations in NS1 showed the change in secondary structure conformation of the mutated AA and those in contact with mutated AA, except for P223S mutation in A/chicken/Viet Nam/KG-076/2004 and L33I in A/chicken/Viet Nam/LD-080/2004, where only the conformation of residues in contact with these AAs was changed and not of that particular AA. It was confirmed

from contact residue analysis that the changes in secondary structure of LEU33 to ILE as these AAs were binding with LEU33 structure at PHE32 and LEU36 were due to mutation of (Figure1).



**Figure 1:** Representation of mutation in RNA binding domain (INS1) of NS1 protein showing the mutated amino acid site LEU33 in space fill, residues in contact with LEU33 are represented in ball and stick and the residues that have gone change in conformation are labeled inside oval

### Discussion:

The most important finding that we observed was the identification of a new CKII phosphorylation domain. This domain was formed in A/Hatay/2004 due to the mutation from ILE151 to THR151. The mutation at I171N changed amino acid from hydrophilic to hydrophobic and hence secondary structure was changed from helix to strand. Mutation at P223S might play a role in folding properties of the protein since proline is an uncharged amino acid that helps in protein folding while serine is a hydrophilic, polar amino acid and it has no role in protein folding. [29]

These changes are significant in studying secondary structure of protein as it is fact that Glycine being a smallest amino acid fits into tight places inside a folded protein; on the other hand glutamic acid helps in the formation of  $\alpha$ -helix whereas glycine disrupts  $\alpha$ -helix formation. [30] As it was found in A/chicken/Vietnam/27/2003, G148E mutation, glutamic acid that is hydrophilic AA was replaced by hydrophobic glycine and the secondary structure was altered from helix to coil (Table 1 in supplementary material).

The signature for CKII phosphorylation domain is [ST] - x(2) - [DE] where [S or T] is the phosphorylation site, here at first position of the domain only S or T is allowed (17). CKII is a multifunctional protein kinase that has been

implicated in a variety of cellular processes and functions, including mitosis and cellular transformation. Several nuclear proteins, enzymes and transcription factors serve as substrates for CKII. Phosphorylation by CKII in vitro alters the DNA binding or catalytic activity of some of these substrates. [31, 17, 32] It was reported earlier that amino acids 134 to 161 of NS1 gene is an effector domain which interacts with cellular proteins to inhibit mRNA nuclear export [6] and it is highly conserved within the NS1 gene. [11]

Thus, it can be said that the changes observed in the predicted secondary structure of NS1 are due to the mutations at various positions. We have also analyzed in our previous work that the main reason of changes in secondary structure conformation is mutation. [21]

### Conclusion:

Influenza is capable of exerting significant mortality, morbidity and economic disruption. Vaccine production is complicated due to antigenic diversity created by 'shift' and interpandemic 'drift'. In this work we have tried to provide insight into NS1 protein of Influenza H5N1 to demonstrate its role in virulence. We noticed a new CKII phosphorylation domain at position 151-154. CKII is critically involved in viral budding and transcription. [33, 34] We also conclude that secondary structure is greatly

tampered due to these mutations, along with changes in hydrophobicity and polarity.

### Acknowledgement:

The authors are grateful to Prof. M Saleemuddin for providing facilities to carryout this work. We also thanks to the Staff of the Distribution information sub-center for technical help. Department of Biotechnology, Ministry of Science and Technology, Government of India is acknowledged for the financial support.

### References:

- [01] D. E. Swayne & D. L. Suarez, *Rev Sci Tech.*, 19:463 (2000) [PMID: 10935274]
- [02] R. G. Webster, *et al.*, *Microbiol Rev.*, 56:152 (1992) [PMID: 1579108]
- [03] L. T. Lau, *et al.*, *Biochem. Biophys. Res. Commun.*, 313:336 (2004) [PMID: 14684165]
- [04] S. Ludwig, *et al.*, *Viral Immunol.*, 12:175 (1999) [PMID: 10532647]
- [05] Y. Lu, *et al.*, *Virology*, 214:222 (1995) [PMID: 8525619]
- [06] X. Y. Qian, *et al.*, *J. Virol.*, 68:2433 (1994) [PMID: 8139028]
- [07] Y. Qiu & R. M. Krug, *J. Virol.*, 68:2425 (1994) [PMID: 7908060]
- [08] F. V. Alonso-Caplen, *et al.*, *Genes Dev.*, 6:255 (1992) [PMID: 1531330]
- [09] Y. Lu, *et al.*, *Genes Dev.*, 8:1817 (1994) [PMID: 7958859]
- [10] Y. Qiu, *et al.*, *RNA*, 1:304 (1995) [PMID: 7489502]
- [11] W. Wang & R. M. Krug, *RNA*, 4:55 (1998) [PMID: 9436908]
- [12] T. Wolff, *et al.*, *J. Virol.*, 70:5363 (1996) [PMID: 8764047]
- [13] T. Wolff *et al.*, *J. Virol.*, 72:7170 (1998) [PMID: 9696811]
- [14] <http://deposit.ddb.de/>
- [15] S. H. Seo, *et al.*, *Virus Res.*, 103:107 (2004) [PMID: 15163498]
- [16] F. Meggio & L. A. Pinna, *The FASEB Journal*, 17:349 (2003) [PMID: 12631575]
- [17] L. A. Pinna, *Biochim. Biophys. Acta.*, 1054:267 (1990) [PMID: 2207178]
- [18] E. K. Hui & D. P. Nayak, *Journal of General Virology*, 83:3055 (2002) [PMID: 12466482]
- [19] T. Das, *et al.*, *The Journal Of Biological Chemistry*, 270:24100 (1995) [PMID: 7592611]
- [20] <http://www.nature.com/nsu/040112/040112-3.html>
- [21] T. Anwar, *et al.*, *In Silico Biology* 6:0015 (2006) [PMID: 16922680]
- [22] <http://www.ncbi.nlm.nih.gov/BLAST/>
- [23] T. Anwar & A. U. Khan, *Bioinformatics*, 1:92 (2006) [PMID: 17597864]
- [24] <http://www.expasy.org/tools/protscale-ref.html>
- [25] <http://www.cmpharm.ucsf.edu/~nomi/nnpredict.html>
- [26] <http://ca.expasy.org/tools/scanprosite/>
- [27] <http://i.moltalk.org/iMolTalk.cgi>
- [28] [www.rcsb.org/pdb](http://www.rcsb.org/pdb)
- [29] R. S. Prajapati, *et al.*, *Proteins*, 66:480 (2007) [PMID: 17034035]
- [30] T. J. Hyde, *et al.*, *J Biol Chem.*, 281:36937 (2006) [PMID: 16998200]
- [31] A. M. Edelman, *et al.*, *Ann. Rev. Biochem.*, 56:567 (1987) [PMID: 2956925]
- [32] P. T. Tuazon & J. A. Traugh, *Adv. Second Messenger Phosphoprotein Res.*, 23:123 (1991) [PMID: 1997039]
- [33] R. Kaushik & M. S. Shaila, *J Gen Virol.*, 85:687 (2004) [PMID: 14993654]
- [34] Y. Gao & J. Lenard, *EMBO J.*, 14:1240 (1995) [PMID: 7720714]

Edited by M. Gollery

Citation: Anwar & Khan, *Bioinformatics* 2(2): 57-61 (2007)

**License statement:** This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.

### Supplementary material

Base Position of Mutation	Change of AA in Hatay/2004	Change in Properties	Change in Secondary Structure	Change in Hydrophobicity (Kyte & Doolittle)
NS1 protein				
A/chicken/Viet Nam/KG-076/2004				
81	A → V	Hydrophobic	C → H	-0.089 → 0.178
132	T → A	Hydrophilic → Hydrophobic	C → H	0.067 → 0.344
151	I → T	Hydrophobic → Hydrophilic	E → C	0.756 → 0.178
171	N → I	Hydrophilic → Hydrophobic	H → E	-0.011 → 0.878
223	S → P	Hydrophilic → Uncharged	C	-0.587 → -0.676
A/chicken/Viet Nam/LD-080/2004				
33	I → L	Hydrophobic	H	-0.522 → -0.6
132	T → A	Hydrophilic → Hydrophobic	C → H	0.067 → 0.344
A/chicken/Viet Nam/TG-023/2004				
132	T → A	Hydrophilic → Hydrophobic	C → H	0.067 → 0.344
147	K → E	Hydrophilic	C → H	0.111 → -0.422
A/chicken/Vietnam/27/2003				
132	I → A	Hydrophobic	E → H	0.644 → 0.344
148	E → G	Hydrophilic → Hydrophobic	H → C	0.778 → 0.544

**Table 1:** Comparison of mutations in Non Structural protein (NS1) of A/Hatay/2004/(H5N1) with other strains of the data set giving its specific position with changes in secondary structure and their properties. \* H - Alpha helix, E - Extended strand and C - Random coil.

NS1		
Site	Position	Domain
Protein kinase C	42 - 44	SIR
phosphorylation site	73 - 75	SdK
	186 - 188	TvR
N-myristoylation site	47 - 52	GNtlGL
	149 - 154	GAtvGE
	174 - 179	GVliGG
	179 - 184	GLewND
Casein kinase II phosphorylation site	89 - 92	TleE
	151 - 154	TvgE
	165 - 168	TgeD
	200 - 203	SsdE
	201 - 204	SdeD
N-glycosylation site	128 - 131	NFSV

**Table 2:** Domains/Motifs in Non-structural protein (NS1) of A/Hatay/2004/(H5N1) representing the site name, its position on the sequence and the domain directing the site