Retraction

Targeting pseudoknots in H5N1 hemagglutinin using designed aptamers

Priyanka Dhar[1],*, Sayak Ganguli[2], Abhijit Datta[2]

[1]Defence Institute of High Altitude Research, DRDO, C/o 56 APO, Pin- 901205, India; [2]Bioinformatics Centre, Presidency College, Kolkata 700073, India; Priyanka Dhar – Email: priyanko_8@yahoo.co.in; Phone/Fax: 01982-252096; *Corresponding author

Abstract:
Influenza A virus subtype H5N1 is highly contagious among birds, causing high mortality among domestic poultry. The viral genome is contained on eight single RNA strands of which HA encode the antigenic glycoprotein called hemagglutinin. Hemagglutinin found on the surface of the influenza viruses and is responsible for binding the virus to the cell that is being infected. Among the most prevalent RNA structures the pseudoknot motif represents an important piece of RNA architecture, as it provides a means for a single RNA strand to fold upon itself to produce a globular structure capable of performing important biological functions. In this analysis we have identified the pseudoknot motifs in the hemagglutinin gene of HPAI A (H5N1) Asian strains. Specific aptamers have been designed against these pseudoknots. These *in-silico* aptamers can be used to hinder the ability of pseudoknots to facilitate ribosomal frameshifting. This may ultimately lead to reduce the coding efficiency of the HA that encodes hemagglutinin and might be used as molecular medicine for H5N1.

Keywords: H5N1, hemagglutinin, pseudoknot, ribosomal frameshifting, RNA aptamers

Background:
Influenza A virus subtype H5N1, which is also known as HPAI A (H5N1) for "highly pathogenic avian influenza virus of type A of subtype H5N1" that is highly contagious among birds, causing high mortality among domestic poultry. H5N1 epizootic in Asia, Europe, the Near East, and Africa is not expected to diminish significantly in the short term. So far, the spread of H5N1 virus from person-to-person has been very rare and unsustained. However, this epizootic continues to pose an important public health threat. They are classified into subtypes based on antigenic differences in the two surface glycoproteins, the hemagglutinin (HA) and neuraminidase (NA). Genetic variation of the HA and/or NA genes resulting in the emergence of new influenza virus strains has frequently been recognized. The mechanism of variation can be either 'genetic drift', which occurs as a result of point mutations, or 'genetic shift', which occurs as a result of genetic reassortment between two strains of influenza A virus. Influenza A virus contains eight segments of a single stranded RNA genome with negative polarity. The segmentation of the genome thus makes genomic reassortment possible when two strains co-infect a permissive cell [1, 2]. Direct transmission of avian influenza virus from domestic poultry to humans was demonstrated for the first time by the outbreak of the highly pathogenic avian H5N1 virus in Hong Kong in 1997 [3, 4] and subsequently confirmed by human infection caused by avian H9N2 influenza viruses in Hong Kong in 1999 [5], by avian H7N7 viruses in UK in 1996 [6] and in Netherlands in 2003 [7], and again by avian H5N1 viruses in Hong Kong in February 2003 [8].

The reemergence of H5N1 influenza A viruses affected a large area of East and Southeast Asia for the first time from late 2003 to May 2004 and eight countries viz. South Korea, Vietnam, Japan, Thailand, Cambodia, China, Laos and Indonesia [9] were involved. However, only Vietnam and Thailand reported the disease in humans [10, 11]. The viral genome of H5N1 is contained on eight single (non-paired) RNA strands or segments that code for eleven proteins. RNA molecules fulfill a diverse set of biological functions within cells, from the transfer of genetic information from DNA to protein, to enzymatic catalysis. Just as proteins form distinct structural motifs such as zinc fingers and beta barrels, certain structures are also commonly adopted by RNA molecules. Among the most prevalent RNA structures is a motif known as the pseudoknot, which are tertiary RNA structures that are formed by Watson-Crick base pairing between a secondary loop structure and compliment bases outside the loop. Pseudoknots are of 3 main types: I-type (interior loop), B-type (bulge loop) and H-type (hairpin loop) pseudoknot [12-14].

First recognized in the turnip yellow mosaic virus [15], a pseudoknot is an RNA structure that is minimally composed of two helical segments connected by single-stranded regions or loops. Although several distinct folding topologies of pseudoknot exist, the best characterized is the H type. In the H-type fold, the bases in the loop of a hairpin form intramolecular pairs with bases outside of the stem. This causes the formation of a second stem and loop, resulting in a pseudoknot with two stems and two loops. The two stems are able to stack on top of each other to form a quasi-continuous helix with one continuous and one discontinuous strand. The single stranded loop regions often interact with the adjacent stems (loop 1–stem 2 or loop 2–stem 1) to form hydrogen bonds and to participate in the overall structure of the molecule. Hence, this relatively simple fold can yield very complex and stable RNA structures. It is fitting that they play a variety of diverse roles like forming the catalytic core of various ribozymes [16, 17], self-splicing introns [18], and telomerase [19]. Additionally, pseudoknots play critical roles in altering gene expression by inducing ribosomal frameshifting in many viruses [20-23]. The pseudoknot, therefore, represents an important piece of RNA architecture, as it provides a means for a single RNA strand to fold upon itself to produce a globular structure capable of performing important biological functions. Therefore, the pseudoknot structures involved in frameshifting are attractive targets for the development of antiviral drugs. Due to this significance of pseudoknots in viral perpetuation, it becomes necessary to take a look into the viral genome of H5N1 in order to locate presence of pseudoknots within it. Presence of pseudoknot within the genes might hold a great implication for future research into prevention of the deadly pandemic. In this analysis we have found the pseudoknot motifs in the hemagglutinin gene of HPAI A (H5N1) Asian strains and specific aptamers were designed against these pseudoknots.

Methodology:
Dataset: The work was performed absolutely *in-silico*. The materials are existing database sequences available at the GenBank of NCBI. Nucleotide sequences of hemagglutinin are obtained from the NCBI-GenBank [24] and the summary of sequence data given in **Table 1 (in supplementary material)**.

193
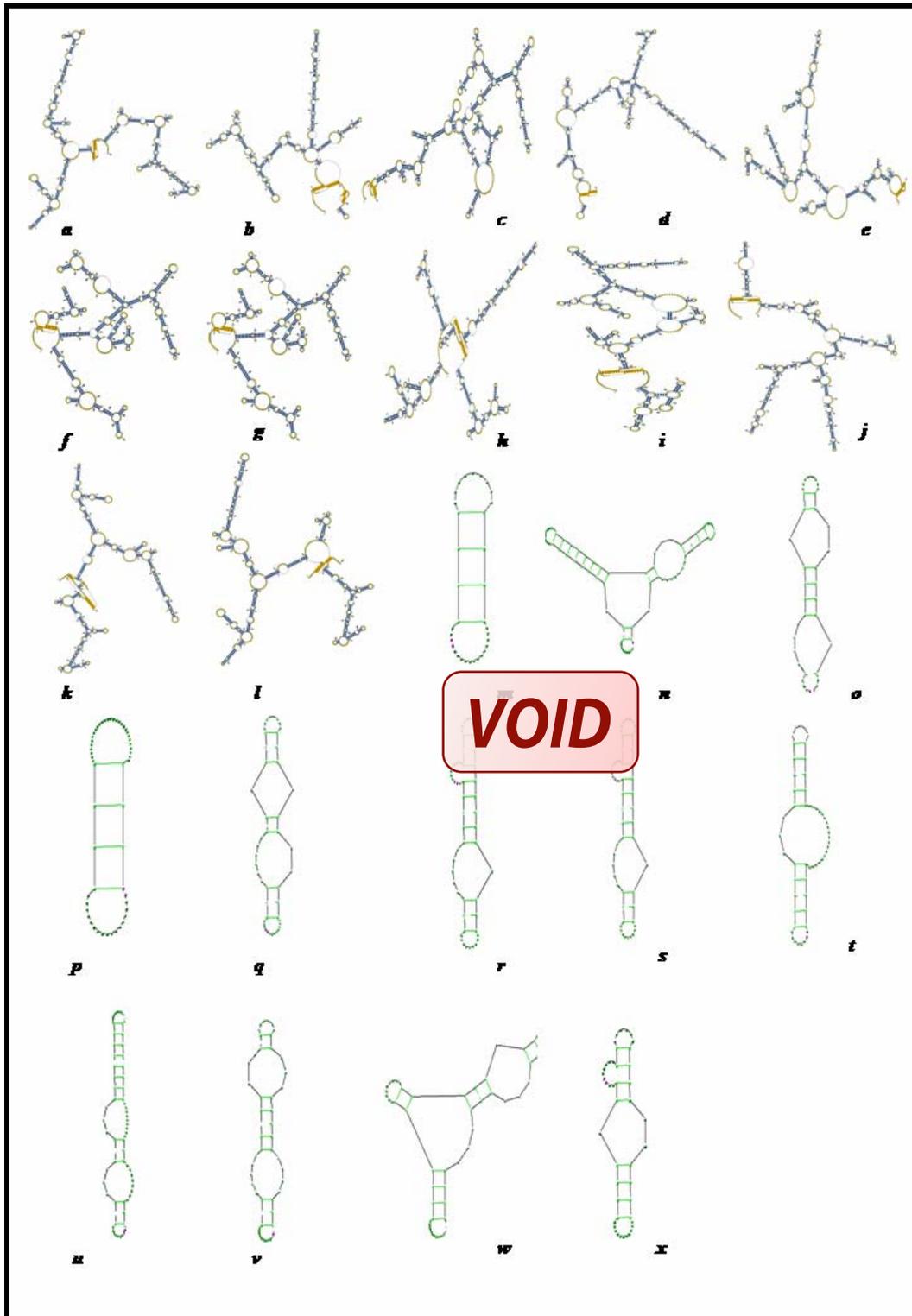
Figure 1: (*a* - *l*) Pseudoknots in the Asian strain of Influenza A virus hemagglutinin (HA) mRNA secondary structure. (*m* - *x*) Designed aptamers against the pseudoknots of HA mRNA

Conversion of nucleotide sequences: The nucleotide sequences were converted to the complementary mRNA by the Complementary Sequence Conversion Tool [25].

Identification of pseudoknot motifs: The mRNA sequences were analyzed for the identification of possible pseudoknot motifs. The online pseudoknot detection tool from the Bielefeld University Bioinformatics Server [26] was used to identify pseudoknots in the mRNAs.

Designing of aptamers: The sequences that are involved in the pseudoknots were selected and complementary RNA sequences were used to design the aptamers against these pseudoknots using the Vienna RNA suite [27].

~~VOID~~

Discussion:
Pseudoknots induce ribosomes to slip into alternative reading frames, otherwise known as frameshifting. Ribosomes typically translate mRNA without shifting the translational reading frame [28]. However, a number of organisms have evolved mechanisms to cause site-specific or programmed frameshifting of the ribosome in either the +1 or −1 direction [29]. Programmed −1 ribosomal frameshifting is typically found in viruses and is required for the replication and proliferation of all retroviruses. Therefore, the pseudoknot structures involved in frameshifting are attractive targets for the development of antiviral drugs. The frameshifting event is induced by two RNA elements within the mRNA: (i) a heptanucleotide slippery sequence X XXY YYZ (spaced triplets represent preframeshift codons) and (ii) a downstream RNA structure, typically a pseudoknot [30]. The mechanism behind how these elements promote −1 frameshifting is not fully understood. The current model posits that the ribosome encounters the downstream pseudoknot while the slippery sequence is being decoded by the ribosome. The pseudoknot structure likely causes the ribosome to pause, which is necessary but not sufficient for frameshifting to occur [31]. While paused on the slippery sequence, the ribosome slips back one nucleotide and subsequently continues translation in the −1 reading frame. The analysis has revealed that there are several pseudoknots present in the hemagglutinin mRNA [**Figure 1** (*a - l*)]. The pseudoknot motif that is predominant among the sequences is the H-type pseudoknot. After the identification of these pseudoknots we have designed the aptamers which can be targeted to those pseudoknots [**Figure 1** (*m - x*)]. By this approach the pseudoknot motifs can be blocked and their ability to facilitate the ribosomal frameshifting can be hindered. This may ultimately lead to reduce the coding efficiency of the HA that encodes hemagglutinin.

Conclusion:
The pseudoknot represents an important piece of RNA architecture, as it provides a means for a single RNA strand to fold upon itself to produce a globular structure capable of performing significant biological functions. The pseudoknot structures involved in frameshifting are the targets for the development of antiviral drugs.

The present study focused on the identification of pseudoknots within the viral genome of H5N1. Presence of pseudoknot within the genes might hold a great implication for future research into prevention of the virulent disease. We have found the pseudoknot motifs in the hemagglutinin gene of HPAI A (H5N1) Asian strains and specific aptamers were designed against these pseudoknots which will play great role in development of drugs for this deadly epidemic.

References:
[1] KG Nicholson *et al., Lancet* **362**: 1733 (2003) [PMID: 14643124]
[2] M Zambon, *Rev Med Virol.* **11**: 227 (2001) [PMID: 11479929]
[3] EC Claas *et al., Lancet* 351: 472 (1998) [PMID: 9482438]
[4] KY Yuen *et al., Lancet* **351**: 467 (1998) [PMID: 9482437]
[5] M Peiris *et al., Lancet* **354**: 916 (1999) [PMID: 10489954]
[6] J Kurtz *et al., Lancet* **348**: 901 (1996) [PMID: 8826845]
[7] M Koopmans *et al., Lancet* **363**: 587 (2004) [PMID: 14987882]
[8] JSM Peiris *et al., Lancet* **363**: 617 (2004) [PMID: 14987888]
[9] Centers for Disease Control and Prevention. *Morb Mortal Wkly Rep.* (2004) **53**: 97
[10] T Chotpitayasunondh *et al., Morb Mortal Wkly Rep.* **53**: 100–102 (2004)
[11] TT Hien *et al., N Engl J Med.* **350**: 1179 (2004) [PMID: 14985470]
[12] PJ Michiels *et al., J Mol Biol.* **310**: 1109 (2001) [PMID: 11501999]
[13] MH Kolk *et al., Science* **280**: 434 (1998) [PMID: 9545221]
[14] DP Giedroc *et al., J Mol Biol.* **298**: 167 (2000) [PMID: 10764589]
[15] K Rietveld *et al., Nucleic Acids Res.* **10**: 1929 (1982) [PMID: 7079175]
[16] T Rastogi *et al., EMBO J.* **15**: 2820 (1996) [PMID: 8654379]
[17] A Ke *et al., Nature* **429**: 201 (2004) [PMID: 15141216]
[18] PL Adams *et al., Nature* **430**: 45 (2004) [PMID: 15175762]
[19] CA Theimer *et al., Mol Cell.* **17**: 671 (2005) [PMID: 15749017]
[20] LX Shen *et al., J Mol Biol.* **247**: 963 (1995) [PMID: 7723043]
[21] PL Nixon *et al., J Mol Biol.* **322**: 621 (2002) [PMID: 12225754]
[22] PJ Michiels *et al., J Mol Biol.* **310**: 1109 (2001) [PMID: 11501999]
[23] M Egli *et al., Proc Natl Acad Sci U S A.* **99**: 4302 (2002) [PMID: 11904368]
[24] http://www.ncbi.nlm.nih.gov/
[25] http://clasher.myweb.uga.edu/testpages/seqconv.html
[26] http://bibiserv.techfak.uni-bielefeld.de/
[27] http://www.tbi.univie.ac.at/RNA/
[28] PJ Farabaugh, GR Bjork, *EMBO J.* **18**: 1427 (1999) [PMID: 10075915]
[29] RF Gesteland, JF Atkins, *Annu Rev Biochem.* **65**: 741 (1996) [PMID: 8811194]
[30] EB ten Dam *et al., Virus Genes* **4**: 121 (1990) [PMID: 2402881]
[31] P Somogyi *et al., Mol Cell Bio.* **13**: 6931 (1993) [PMID: 8413285]

Supplementary material:

Table 1: Influenza A virus strains selected for analysis.

GenBank accession	Strain	Subtype	Source
gi|193795159|gb|EU497921.2|	A/chicken/Thailand/ICRC-213/2007	H5N1	Chicken
gi|167716900|gb|EU430496.1|	A/domestic mallard/Hunan/67/2005	H5N1	Domestic mallard
gi|83940767|gb|DQ323672.1|	A/chicken/Kurgan/3/2005	H5N1	Chicken
gi|54126527|gb|AY747617.1|	A/swine/Fujian/F1/2001	H5N1	Swine
gi|76800619|gb|DQ211925.1|	A/chicken/luohuo/3/03/	H5N1	Chicken
gi|76800615|gb|DQ211923.1|	A/chicken/zheng	H5N1	Chicken
gi|76800613|gb|DQ211922.1|	A/chicken/jiyuan/0	H5N1	Chicken
gi|56418513|gb|AY830774.1|	A/chick/Macheng/2004	H5N1	Chick
gi|50843949|gb|AY679514.1|	A/Thailand/LFPN-2004/2004	H5N1	Chicken
gi|76800617|gb|DQ211924.1|	A/chicken/zhoukou/2/02/	H5N1	Chicken
gi|169143561|gb|EU497920.1|	A/chicken/Thailand/ICRC-195/2007	H5N1	Chicken
gi|169143559|gb|EU497919.1|	A/chicken/Thailand/ICRC-V586/2008	H5N1	Chicken