Selection of predicted siRNA as potential antiviral therapeutic agent against influenza virus

Asif Raza¹, Hira Shareef¹, Hira Salim¹, Rashid Khushal², Habib Bokhari¹*¹

¹Department of Biosciences, COMSATS Institute of Information Technology, Chak Shazad Campus, Park Road, Islamabad, Pakistan; ²Astrazeneca, Nottingham, UK; Habib Bokhari - Email: habib@comsats.edu.pk; Phone: 44-07915191582; Fax: 0092-051-4442805; *Corresponding author

Received June 03, 2011; Accepted June 24, 2011; Published July 19, 2011

Abstract:
Influenza virus A (IVA) infection is responsible for recent death worldwide. Hence, there is a need to develop therapeutic agents against the virus. We describe the prediction of short interfering RNA (siRNA) as potential therapeutic molecules for the HA (Haemagglutinin) and NA (Neuraminidase) genes. We screened 90,522 siRNA candidates for HA and 13,576 for NA and selected 1006 and 1307 candidates for HA and NA, respectively based on the proportion of viral sequences that are targeted by the corresponding siRNA, with complete matches. Further short listing to select siRNA with no off-target hits, fulfilling all the guidelines mentioned in approach, provided us 13 siRNAs for haemagglutinin and 13 siRNAs for neuraminidase. The approach of finding siRNA using multiple sequence alignments of amino acid sequences has led to the identification of five conserved amino acid sequences, three in hemagglutinin i.e. RGLFGAIAGFIE, YNAELLV and AIAGFIE and two in neuraminidase i.e. RTQSEC and EECSYP which on reverse translation provided siRNA sequences as potential therapeutic candidates. The approaches used during this study have enabled us to identify potentially therapeutic siRNAs against divergent IVA strains.

Keywords: Influenza virus A, Sequence analysis, siRNA, Hemagglutinin, Neuraminidase.

Background:
Influenza viruses, a cause of significant morbidity and mortality in human population, possess a segmented genome of negative sense RNA encapsulated by a viral specified nucleoprotein. There are three types of influenza viruses that cause illness in humans: type A, B and C, classified according to the serological reactivity of internal proteins. The most common infectious agent mentioned in approach, provided us 13 siRNAs for haemagglutinin and 13 siRNAs for neuraminidase. The approach of finding siRNA using multiple sequence alignments of amino acid sequences has led to the identification of five conserved amino acid sequences, three in hemagglutinin i.e. RGLFGAIAGFIE, YNAELLV and AIAGFIE and two in neuraminidase i.e. RTQSEC and EECSYP which on reverse translation provided siRNA sequences as potential therapeutic candidates. The approaches used during this study have enabled us to identify potentially therapeutic siRNAs against divergent IVA strains.

Influenza A viruses in birds increases the number of viral subtypes available for reassortment, when two different viruses from two different host species co-infect a single individual animal, that have the potential to cause major global pandemics [3]. Due to such processes, human influenza is considered by some as a non eradicable disease, since the range of antigenically diverse subtypes of influenza A viruses are present in waterfowl and can be introduced into humans and may lead to millions of deaths [6-9]. Existing drugs and vaccines against Influenza A virus have limited value and the threat of pandemic persists. Broad spectrum anti-virals have been developed in the past that target an enzyme found in all viruses that have RNA genomes (except retroviruses) i.e. viral RNA-dependent RNA polymerases [10, 11].

Apart from viral proteases and polymerases, some other viral proteins can act as potential drug targets. For instance, Inhibitors of virus entry is another new promising class of drugs that show clinical efficacy. In this regard, inhibitors that block virus entry into cells, by targeting the virus envelope glycoproteins have yielded a clinically approved compound for HIV [12] as well as a promising lead to treat respiratory syncytial virus infections [13]. Similarly, inhibitors of the neuraminidase (glycoprotein) of influenza A and B viruses, two of which are zanamivir [14] and oseltamivir [15], are now approved for therapeutic use in humans. By inhibiting specific cell proteins that are required for virus replication it should be possible to interrupt the virus life cycle [16]. However, possible cross reactivity with host proteins can cause adverse side-effects inhibiting essential host cell functions [17]. On contrary to agents typically inhibiting the function of crucial viral proteins, RNAi is the relatively newly described natural biological phenomenon that achieves the identical goal by targeting the viral mRNA instead of the proteins they encode. RNA interference (RNAi) is now being widely used to knockdown gene expression,
sequences from corresponding countries as mentioned above were downloaded, for multiple alignments, from NCBi Influenza Virus Sequence Database [http://www.ncbi.nlm.nih.gov/ genomes/FLU/FLU.html]. Sequences for siRNA were retrieved from siVirus based on their degree of conservation, defined as the proportion of viral sequences that are targeted by the corresponding siRNA, with complete matches (i.e. 21/21 matches).

Figure 1: Result showing conserved regions in haemagglutinin using multiple sequence alignment

siRNA analysis:

siVirus interface was used to determine anti-siRNA sequence results. This interface especially focuses on anti-siRNA design and provides (a) highly conserved target sites for designing anti-siRNA that would resist viral mutational escapes (b) selection of effective siRNAs based on guidelines of Ui-Tei et al., Reynolds et al., Amarzguioui et al. [18-20] (c) Off-target minimized siRNAs within two mismatches against human genes. Off-target hits are those which have dissimilarity to the target sequence. It is desirable to select siRNA that has less off-target hits.

Sequence analysis:

TCOFFEE [http://www.ebi.ac.uk/t-coffee/html] was used for multiple alignments to observe conservation and functional sites in the HA and NA datasets at amino acid level. The consensus in at least 7 amino acids corresponding to 21 nucleotides was noted. BLAST program was used for similar sequence searches.

Results and Discussion:

To design siRNA against influenza virus A, we analyzed all the sequences for gene segment 4 (haemagglutinin) and 6 (neuraminidase) of Influenza Virus A from some of the Asian countries (Tables 3 & 4 see Supplementary material) The selection of effective siRNAs sequences were based on published guidelines [18-20]. Those guidelines were: (a) A/U at the 5' end of the antisense strand, (b) G/C at the 5' end of the sense strand, (c) At least 5 A/U residues at the 5' terminal, one-third of the antisense strand (d) The absence of any GC stretch of more than 9 nucleotides in length. In case any of the above mentioned guideline is not fulfilled, then no gene silencing by a particular siRNA is expected. Furthermore, consensus in amino acid sequences in the HA (H1, H3, H5 and H9) and NA (N1-N2) of different strains was noted and only those strings of at least 7 amino acids were retained which were present in all types studied.

siRNA prediction using siVirus:

Figure 1 shows a typical output from siVirus for designing anti-influenza Virus A siRNAs. The sequences from Influenza virus A subtypes A and B, which are the most prevalent genotypes circulating in Asia, were selected. The results were sorted by the degree of conservation and filtered to display siRNAs that satisfy at least one efficacy guideline mentioned above. Out of 90522 siRNA candidates screened, 1006 sequences targeting HA segment, whereas out of 13576 siRNA candidates screened, 1307 sequences targeting NA segment of influenza virus A genome were short listed respectively. Analysis of these siRNA candidates revealed that no conserved siRNAs were
obtained when using 90% conservation stringency parameter. However, on decreasing the stringency parameter close to 90%, 100% conservation of siRNAs pertaining to gene segment 4 (haemagglutinin) (Figure 1 & 2) and segment 6 (neuraminidase) (Data Not shown) was noted. Further short listing to select siRNA with no off-target hits, fulfilling all the guidelines mentioned in approach, provided us 13 siRNAs for haemagglutinin and 13 siRNAs for neuraminidase (Tables 1 & 2 see Supplementary material). We believe that the approach of developing siRNA based on combination of both approaches i.e. Direct search for siRNA against appropriate targets and finding the consensus of amino acids of majority of the circulating virus variants would provide a broad range protection molecule to control the infection.

Acknowledgements:
We thank Higher Education Commission of Pakistan and COMSATS Institute of Information Technology, Islamabad for all the support and facilities provided to carry out this work.

References:
[1] Hatz AJ et al. J Med Microbiol. 2009 58: 408 [PMID: 19273634]
[2] Fitch WM et al. Proc Natl Acad Sci U S A. 1997 94: 7712 [PMID: 9223253]
[3] Zambon MC. Rev Med Virol. 2001 11: 227 [PMID: 11479929]
[4] Rohn C et al. Virology. 1996 217: 508 [PMID: 8610442]
[5] Webster RG et al. Microbial Rev. 1992 56: 152 [PMID: 1579108]
[6] Subbarao K & Shaw MW. Rev Med Virol. 2000 10: 337 [PMID: 11015744]
[7] Paterson KD & Pyle GF. Bull Hist Med. 1991 65: 4 [PMID: 2021692]
[8] Malik JSP et al. Clin Microbiol Rev. 2007 20: 243
[9] Silva JM et al. Trends Mol Med. 2002 8: 5050
[10] Ge Q et al. Proc Natl Acad Sci U S A. 2003 100: 2718 [PMID: 12594334]
[11] Nicholson KG et al. Lancet. 2000 355: 1845 [PMID: 10866439]
[12] Chen KY et al. Expert Opin Investig Drugs. 2002 11: 1837
[13] Cianci C et al. J Antimicrob Chemother. 2005 55: 289 [PMID: 15681582]
[14] Itzstein MV et al. Nature 1993 363: 418
[15] Kim CU et al. J Am Chem Soc. 1997 119: 681 [PMID: 16526129]
[16] Schumann D et al. AIDS. 2007 21: 1293
[17] Esser S et al. J Dtsch Dermatol Ges. 2007 5: 745
[18] Fire A et al. Nature 1998 391: 806 [PMID: 9486653]
[19] Mello CC & Conte D. Nature 2004 431: 338 [PMID: 15372040]
[20] Ui-Tei K et al. Nucleic Acids Res. 2004 32: 936 [PMID: 14769950]
[21] Reynolds A et al. Nat Biotechnol. 2004 22: 326 [PMID: 14758366]
[22] Amarzguioui M & Prydz H. Biochem Biophys Res Commun. 2004 316: 1050 [PMID: 15044091]
[23] Dave RS & Pomereitz RJ. Rev Med Virol. 2003 13: 373 [PMID: 14625885]
[24] Hamada M et al. Antisense Nucleic Acid Drug Dev. 2002 12: 301 [PMID: 12477280]
[25] Jaque JM et al. Nature 2002 418: 435
[26] Lee MT et al. J Virol. 2003 77: 11964
[27] Randall G et al. Proc Natl Acad Sci U S A. 2003 100: 235
[28] McCown M et al. Virology. 2003 313: 514

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.

Bioinformation 6(9): 340-343 (2011)
Supplementary material:

Table 1: siRNA result for Haemagglutinin for Asia

| S# | SiRNA target sites + 2 Nucleotides overhang | % conservation |
|----|---------------------------------------------|----------------|
| 1  | CAGCAACTGTTCACCCTTTATGAg                  | 82.9 (834/1006) |
| 2  | AAGCATCTATTGGACAATGTTaa                   | 81.6 (821/1006) |
| 3  | TACGGTTTCAGGCACTCAAATAtc                  | 81.2 (817/1006) |
| 4  | ACAGTTTCAGGCATCAAATTct                    | 81.2 (817/1006) |
| 5  | ATGGAAGCATTCCCAATGACAa                    | 80.8 (813/1006) |
| 6  | TGGAAAGCATTCCCAATGACAa                    | 80.8 (813/1006) |
| 7  | CAGCAACTGTTCACCCTTTATGAg                  | 72.2 (86/119)   |
| 8  | TACGGTTTCAGGCACTCAAATAtc                  | 72.2 (86/119)   |
| 9  | ACAGTTTCAGGCATCAAATTct                    | 72.2 (86/119)   |
| 10 | GGCAACATTAGGTGAACCATTtg                   | 69.7 (83/119)   |
| 11 | AGGTGCAACATTTCAGTGACCAtg                 | 68 (81/119)     |
| 12 | ATGGAAGCATTCCCAATGACAa                    | 67.2 (80/119)   |
| 13 | TGGAAAGCATTCCCAATGACAa                    | 67.2 (80/119)   |

Table 2: siRNA result for Neuraminidase for Asia

| S# | SiRNA target sites + 2 Nucleotides overhang | % conservation |
|----|---------------------------------------------|----------------|
| 1  | TAGCATGGTCACCTCAAGTgt                      | 85 (1112/1307) |
| 2  | CTCAAAGTGTCAGGATGTGAAa                     | 84.1 (1100/1307) |
| 3  | TGCAACTGTGCTAGCTAATTTAtg                  | 83.3 (1090/1307) |
| 4  | ATGCAACGTGTCAGCTATTatg                    | 83.3 (1089/1307) |
| 5  | GTGGTTTTTAGTGGATGTGAAa                     | 82 (1073/1307)  |
| 6  | ATCCCACTCAAAAGATTAAa                       | 81.4 (1064/1307) |
| 7  | ATGAAATCCAAATCCCAAAAGTAa                   | 74 (120/162)    |
| 8  | ATCCCACTCAAAAGATTAAa                       | 74 (120/162)    |
| 9  | TAGCATGGTCAGCTGTTGtt                      | 72.8 (118/162)  |
| 10 | TCTAAGTTGTCAGATGGTTGAAa                    | 72.2 (117/162)  |
| 11 | TGCAACTGCTAGCTAATTTAtg                    | 71.6 (116/162)  |
| 12 | ATGCAACGTGTCAGCTATTatg                    | 70.9 (115/162)  |
| 13 | TTGCACCTTTTCTAAGGAAa                       | 70.3 (114/162)  |

Table 3: Haemagglutinin sequences used for BLAST and multiple sequence alignment.

| S# | Influenza A | Accession Number | Country of Origin |
|----|-------------|------------------|-------------------|
| 1  | H1          | 33622380         | Beijing           |
| 2  | H1          | 33622384         | Shenzhen          |
| 3  | H1          | 221340           | A/Suita/ Japan    |
| 4  | H1          | 554562           | USSR              |
| 5  | H2          | 305115           | Japan             |
| 6  | H2          | 408525           | Krasnodar         |
| 7  | H3          | 493639098        | Hong Kong         |
| 8  | H3          | 77861715         | Hong Kong         |
| 9  | H3          | 18615859         | Udorn             |
| 10 | H3          | 71558927         | Moscow            |
| 11 | H5          | 71013498         | Hong Kong         |
| 12 | H5          | 2865380          | Hong Kong         |
| 13 | H5          | 2833657          | Hong Kong         |
| 14 | H5          | 3421252          | Hong Kong         |
| 15 | H9          | 8894694          | Hong Kong         |

Table 4: Neuraminidase sequences used for BLAST and multiple sequence alignment.

| S# | Influenza A | Accession Number | Country of Origin |
|----|-------------|------------------|-------------------|
| 1  | N1          | 31096405         | China             |
| 2  | N1          | 31096413         | Hong Kong         |
| 3  | N1          | 31096415         | Hong Kong         |
| 4  | N1          | 2833659          | Hong Kong         |
| 5  | N1          | 3335423          | Hong Kong         |
| 6  | N1          | 47834892         | Hong Kong         |
| 7  | N1          | 47834894         | Hong Kong         |
| 8  | N1          | 324578           | Russia            |
| 9  | N2          | 38524538         | Japan             |
| 10 | N2          | 4973497          | Japan             |
| 11 | N2          | 418078           | China             |
| 12 | N2          | 8894698          | Hong Kong         |
| 13 | N2          | 21632606         | China             |
| 14 | N2          | 21632610         | China             |
| 15 | N2          | 81174818         | Hong Kong         |
| 16 | N2          | 8250669          | Russia            |
| 17 | N2          | 71533740         | Russia            |