Influenza A nucleoprotein binding sites for antivirals: current research and future potential

Andreas Kukol* & Hershna Patel

School of Life and Medical Sciences, University of Hertfordshire, Hatfield AL10 9AB, UK

*) Author for correspondence: a.kukol@herts.ac.uk

The influenza A virus is a significant human and agricultural pathogen that causes annually recurring epidemics of a respiratory disease which leads to 250,000 to 500,000 human deaths worldwide [1] as well as to significant losses among domesticated birds [2]. Additionally there have been several pandemics including the latest 'Swine Flu' with an estimated death toll of 284,000 people over the first year [3]. Due to its ability to genetic reassortment, i.e. exchange of RNA segments between different virus strains, when an individual is infected by two different strains at the same time, new pandemic strains may emerge in the future [4]. This is facilitated by close contact of humans with domesticated birds that occurs in some parts of East Asia. Current avian influenza A virus subtypes that are cause for concern are the H5N1 virus that has caused 386 deaths until January 2014 and the H7N9 virus that has caused 45 deaths until October 2013 [5]. The concern is that these highly pathogenic avian viruses may acquire the capability of efficient human to human transmission. Given the history of the virus and rate of transmission, the occurrence of another pandemic should not be underestimated.

The seasonal flu vaccine formulated against the expected circulating strains cannot provide long term protection and must be administered at least two weeks before the infection occurs. Antivirals may be used to treat patients after an infection has occurred. Current antivirals available in most countries include the M2 proton channel inhibitors amantadine/rimantadine and neuraminidase inhibitors. Currently many virus strains are resistant against the older M2 channel inhibitors, while resistance against the newer neuraminidase inhibitors is emerging [6]. In the authors' opinion the problem of emerging resistance affirms the need for investigating other viral proteins as drug targets as well as including the evolutionary record of these proteins into the early stages of drug target

characterisation and drug discovery as exemplified by studies of the non-structural protein [7] or the nucleoprotein [8].

The influenza A virus belongs to the family Orthomyxoviridae and has a lipid/protein envelope that encloses eight segments of negative-strand RNA, which are associated with the nucleoprotein (NP) and the three polymerase subunits PA, PB1 and PB2. The viral genome encodes for up to 17 proteins, whose expression depend on the particular virus strain (reviewed in [9]). The classification into subtypes $HxNy$ is determined by the antigenic properties of the surface proteins haemagglutinin and neuraminidase; until now 18 haemagglutinin and 11 neuraminidase subtypes have been identified. The RNA segment five encodes for the nucleoprotein (NP), which consists of 496 amino acid residues. The structure of NP from an H1N1 virus [10] and an H5N1 virus [11] was determined by x-ray crystallography. NP can be divided into three domains, namely a head domain, a body domain and a flexible tail loop. The main function of NP is to participate in formation of ribonucleoprotein (RNP) complexes that include NP, the viral RNA and the three polymerase subunits PA, PB1 and PB2. In the virus particle, the RNA is wrapped around NP molecules by binding to an arginine-rich RNA binding groove, while NP forms oligomers by insertion of the tail loop of one NP monomer into the tail-loop binding pocket of another NP monomer involving an ionic interaction between Glu339 and Arg416. NP has multiple functions during virus replication that include RNP trafficking, facilitating viral RNA synthesis, polymerase regulation and interaction with cellular proteins [12]. The amino acid sequence of NP shows various nuclear localisation signals (NLS), which are important for the nuclear import of the polymerase complex. An unconventional NLS can be found in the N-terminus, this becomes deactivated upon phosphorylation of Ser-3 . Additional NLS were found in the RNA binding groove between residues 198 and 216 and between residues 320 and 400 based on deletion studies. The N-terminal unconventional NLS appears to be the major determinant of nuclear import [13], while a nuclear accumulation signal between residues 327 and 345 may cause retention of the NP in the nucleus [14]. Taken together the important functions of NP which successful infection is dependent upon establishes this protein as a promising drug target for the development of influenza A virus replication inhibitors.

Drug discovery and development is a lengthy and costly procedure, emphasising the necessity of a thorough characterisation of the drug target at the beginning. There are many examples of the failure of clinical trials due to the choice of an inappropriate drug target. In case of the influenza virus and other RNA viruses a

further problem is the high mutation rate due to the lack of proof-reading capability of the viral RNA polymerase [15]. This could render years of drug development efforts fruitless within a few years after the drug comes to the market due to the development of resistance. In addition to mutations genetic variability arises as a consequence of genetic reassortment between different virus strains, a process which frequently occurs in pigs that can be infected by both avian and human viruses. The recent H1N1 swine flu virus resulted from a complex reassortment of human H3N2, swine H1N1 and avian H1N1 genetic elements [16]. In the early stages of anti-influenza drug target characterisation, it is therefore of utmost importance to consider sequence conservation among all possible hosts and subtypes together with potential binding sites for antiviral drugs. To our knowledge there are two studies, which have taken this approach for the non-structural protein 1 and nuclear export protein [7] and for the nucleoprotein [8].

For the nucleoprotein 4430 amino acid sequences were analysed using a combination of multiple sequence alignment and evolutionary trees as implemented in the ConSurf method [17]. It is essential to incorporate the evolutionary relationships between sequences, since the content of sequence databases does not contain a random sample of sequences. The incorporation of the evolutionary tree ensures that conservation obtained from sequences closely related in evolution is weighted less than conservation obtained from sequences distant in evolution. The sequence conservation was combined with prediction of binding sites for small organic molecules. The results identified predicted binding sites that overlapped with high evolutionary conservation. One binding site identified was the tail loop binding pocket, which showed high conservation of residues Ser165, Arg267, His272, Glu339, Arg342 and Phe489. Another potential binding site with high conservation was detected in the RNA binding region; it involved the highly conserved sites Gln58, Thr62, Tyr78, Tyr97, Ser141, Thr134, His135, Thr171 and Lys273. Other conserved binding sites were found near the conserved residues Glu339, Asp340 and Arg74, Arg175 and Asp145, Arg150. Some of these sites have already been targeted by experimental inhibitors, such as the tail loop binding pocket for which an organic molecule was discovered that showed an inhibitory activity (IC) of IC_{50} = 2.7 μM [18]. The anti-inflammatory drug naproxen has been show to bind near Asp145, Arg150 and it reduced viral titers with an IC_{50} = 11 μM [19]. The main reason for investigating conserved binding sites for drug discovery is the hypothesis that it is difficult for the virus to become resistant against inhibitors targeted to those sites. Interestingly, this has been confirmed by recent experimental studies. Naproxen was used during six passages of selection and no resistant mutants were found [19],

while in another study the inhibitor nucleozin and analogues were investigated and after five passages of selection a resistant mutant was found with a mutation Tyr289His [20]. Tyr289 together with Arg305 and Arg309 are part of a binding site predicted in the previous study [8] that was not highly conserved. At position 289 the amino acid residues Tyr, Phe and His were found, while Arg305 and Arg309 were even less conserved.

In conclusion, due to its multiple functions the influenza A virus nucleoprotein is a valid drug target for antivirals and various evolutionary conserved binding sites were identified that could provide starting points for future development of antiviral drugs [8]. Initial experimental investigation of replication inhibitors indicate that inhibitors targeted at highly conserved binding sites have less potential for inducing resistance than inhibitors targeted at less conserved sites [19,20]. We believe that the strategy of combining sequence conservation analysis with binding site prediction should be applied not only to other influenza A virus proteins but in general to antiviral drug discovery leading to better ways in achieving antiviral protection in the future.

Financial & competing interests disclosure

1. WHO. Influenza (seasonal). *Fact sheet*, No. 211 (2009).
2. Leibler JH, Otte J, Roland-Holst D *et al.* Industrial Food Animal Production and Global Health Risks: Exploring the Ecosystems and Economics of Avian Influenza. *Ecohealth*, 6(1), 58-70 (2009).
3. Dawood FS, Iuliano AD, Reed C *et al.* Estimated global mortality associated with the first 12 months of 2009 pandemic influenza A H1N1 virus circulation: a modelling study. *Lancet Infect Dis*, 12(9), 687-695 (2012).
4. Lin YP, Shaw M, Gregory V *et al.* Avian-to-human transmission of H9N2 subtype influenza A viruses: Relationship between H9N2 and H5N1 human isolates. *Proceedings of the National Academy of Sciences of the United States of America*, 97(17), 9654-9658 (2000).
5. WHO. China-WHO Joint Mission on Human Infection with Avian Influenza A(H7N9) Virus. (Ed.^(Eds) (Bejing, 2013)
6. Samson M, Pizzorno A, Abed Y, Boivin G. Influenza virus resistance to neuraminidase inhibitors. *Antiviral Research*, 98(2), 174-185 (2013).

7. Darapaneni V, Prabhaker VK, Kukol A. Large-scale analysis of influenza A virus sequences reveals potential drug target sites of non-structural proteins. *Journal of General Virology*, 90, 2124-2133 (2009).

8. Kukol A, Hughes DJ. Large-scale analysis of influenza A virus nucleoprotein sequence conservation reveals potential drug-target sites. *Virology*, 454-455, 40-47 (2014).

9. Vasin AV, Temkina OA, Egorov VV, Klotchenko SA, Plotnikova MA, Kiselev OI. Molecular mechanisms enhancing the proteome of influenza A viruses: An overview of recently discovered proteins. *Virus Res*, 185C, 53-63 (2014).

10. Ye QZ, Krug RM, Tao YZJ. The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature*, 444(7122), 1078-1082 (2006).

11. Ng AK, Zhang H, Tan K *et al.* Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. *FASEB J*, 22(10), 3638-3647 (2008).

12. Li Z, Watanabe T, Hatta M *et al.* Mutational Analysis of Conserved Amino Acids in the Influenza A Virus Nucleoprotein. *Journal of Virology*, 83(9), 4153-4162 (2009).

13. Cros JF, Garcia-Sastre A, Palese P. An unconventional NLS is critical for the nuclear import of the influenza A virus nucleoprotein and ribonucleoprotein. *Traffic*, 6(3), 205-213 (2005).

14. Davey J, Dimmock NJ, Colman A. Identification of the sequence responsible for the nuclear accumulation of the influenza virus nucleoprotein in Xenopus oocytes. *Cell*, 40(3), 667-675 (1985).

15. Holland J, Spindler K, Horodyski F, Grabau E, Nichol S, Vandepol S. Rapid Evolution of Rna Genomes. *Science*, 215(4540), 1577-1585 (1982).

16. Smith GJD, Vijaykrishna D, Bahl J *et al.* Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature*, 459(7250), 1122-U1107 (2009).

17. 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 research*, 38(Web Server issue), W529-533 (2010).

18. Shen YF, Chen YH, Chu SY *et al.* E339 ... R416 salt bridge of nucleoprotein as a feasible target for influenza virus inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 108(40), 16515-16520 (2011).

19. Lejal N, Tarus B, Bouguyon E *et al.* Structure-Based Discovery of the Novel Antiviral Properties of Naproxen against the Nucleoprotein of Influenza A Virus. *Antimicrob Agents Ch*, 57(5), 2231-2242 (2013).

20. Kao RY, Yang D, Lau LS *et al.* Identification of influenza A nucleoprotein as an antiviral target. *Nature Biotechnology*, 28(6), 600-U688 (2010).