

Discovery notes

Open Access

Mapping the sequence mutations of the 2009 H1N1 influenza A virus neuraminidase relative to drug and antibody binding sites

Sebastian Maurer-Stroh*¹, Jianmin Ma¹, Raphael Tze Chuen Lee¹,
Fernanda L Sirota¹ and Frank Eisenhaber^{1,2}

Address: ¹Biomolecular Function Discovery Division, Bioinformatics Institute (BII), Agency for Science Technology and Research (A*STAR), 30 Biopolis Street, #07-01, Matrix, 138671, Singapore and ²Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, 117543 Singapore

Email: Sebastian Maurer-Stroh* - sebastianms@bii.a-star.edu.sg; Jianmin Ma - majm@bii.a-star.edu.sg; Raphael Tze Chuen Lee - leetc@bii.a-star.edu.sg; Fernanda L Sirota - fernanda@bii.a-star.edu.sg; Frank Eisenhaber - franke@bii.a-star.edu.sg

* Corresponding author

Published: 20 May 2009

Received: 13 May 2009

Biology Direct 2009, 4:18 doi:10.1186/1745-6150-4-18

Accepted: 20 May 2009

This article is available from: <http://www.biology-direct.com/content/4/1/18>

© 2009 Maurer-Stroh et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

: In this work, we study the consequences of sequence variations of the "2009 H1N1" (swine or Mexican flu) influenza A virus strain neuraminidase for drug treatment and vaccination. We find that it is phylogenetically more closely related to European H1N1 swine flu and H5N1 avian flu rather than to the H1N1 counterparts in the Americas. Homology-based 3D structure modeling reveals that the novel mutations are preferentially located at the protein surface and do not interfere with the active site. The latter is the binding cavity for 3 currently used neuraminidase inhibitors: oseltamivir (Tamiflu®), zanamivir (Relenza®) and peramivir; thus, the drugs should remain effective for treatment. However, the antigenic regions of the neuraminidase relevant for vaccine development, serological typing and passive antibody treatment can differ from those of previous strains and already vary among patients.

Reviewers: This article was reviewed by Sandor Pongor and L. Aravind.

Findings

The recent epidemic of the "2009 H1N1" influenza A virus (also called swine or Mexican flu) has put the world on alert since a new swine flu strain (naturally hosted by pigs) has crossed the species barrier to human and, apparently, acquired the capability for human to human transmission [1,2]. Given earlier experiences with risks of viral pandemics such as SARS and the avian flu [3], global control and public health surveillance mechanisms provided sequences of the new flu strain in public sequence databases within weeks of the outbreak. Here, we analyze the protein sequence of its neuraminidase with respect to sim-

ilarities and differences to known strains and implications on drug treatment and vaccination.

Domain architecture and posttranslational modifications

Sequence and residue numbering in this analysis correspond to the neuraminidase [Genbank: ACP41107.1 <http://www.ncbi.nlm.nih.gov/protein/227809834>] representative for the new strain. Sequence analysis was carried out following an established protocol using the ANNIE resource [4,5]. The 469 amino acid long neuraminidase (NA) protein (Figure 1) is essential for release of the viral particle from the outer membrane of infected cells by

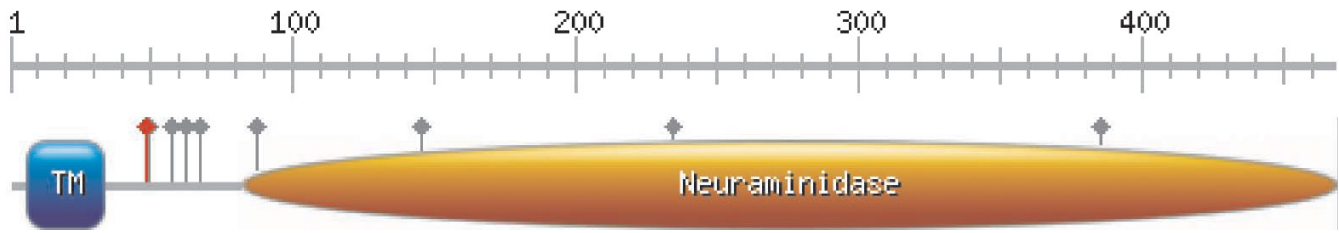


Figure 1

Domain architecture (drawn with <http://au.expasy.org/tools/mydomains/>). Besides the labelled domains (TM ... trans-membrane), grey lollipops indicate known and putative glycosylation sites and the red lollipop marks the conserved cysteine shown in Figure 2.

cleaving sialic acid from host glycoproteins that are recognized by the viral hemagglutinin [6]. As a type II trans-membrane protein, it is N-terminally attached to the membrane [7]. It consists of a tiny cytoplasmic tail at the N-terminus (residues 1 to 6) [8] followed by the trans-membrane region (residues 7 to 34) that is also responsible for translocation of the protein [9].

Next, a presumably unstructured linker region (residues 35 to 82) connects the membrane anchor to the catalytic neuraminidase domain (residues 83 to 469; Figure 1). Such unstructured linker regions are rich in small and polar residues and often harbour sites for posttranslational modifications [10,11]. Probable posttranslational modification sites in the neuraminidase of the new strain are glycosylation motifs involving N88, N146 and N235, which correspond to residues that are also glycosylated in other subtype neuraminidases [12]. However, the minimal and non-specific consensus motif of glycosylation sites (Nx [ST]) is found in total 8 times in the new strain sequence with an apparent clustering (50%) in the unstructured linker region (Figure 1). Interestingly, another putative novel glycosylation site N386, which is

unique to the new strain, would be accessible on the surface, as seen in the structural models.

Comparing among all strains, the sequence variation is largest in the linker region, including large deleted segments. Nevertheless, this region harbours a cysteine (Figure 2) that can be aligned over multiple NA subtypes and is conserved in N1-N5 and N8, but not in N6, N7 and N9. Earlier reports assume that, at least in related viruses, cysteines in the non-globular region could be involved in intermolecular disulfide bridges [13-15]. Alternatively, by analogy to other influenza proteins such as hemagglutinin [16] and M2 protein [17], it cannot yet be excluded that cysteine C49 is palmitoylated and that the anchor localizes the protein to lipid rafts [18].

Phylogenetic relation of new NA to known subtypes

Influenza A virus protein sequences were downloaded from NCBI (as of April 29th). Neuraminidases were identified by BLAST (E-value < 0.001) [19] using the representative NA of the new strain as query [Genbank: ACP41107.1 <http://www.ncbi.nlm.nih.gov/protein/227809834>]. Redundancy was removed with cd-hit at a level of maximal 90% sequence identity [20], the remain-

<i>H1N1_California_04_2009</i>	48	I	E	T	C	N	Q	S	52
<i>H1N1_swine_lowa_1945</i>	48	P	E	T	C	N	Q	S	52
<i>H5N1_HongKong_516_97</i>	48	P	E	P	C	N	Q	S	52
<i>H5N2_chicken_Pennsylvania_1370_1983</i>	50	T	T	P	C	K	P	I	56
<i>H2N3_swine_Missouri_4296424_2006</i>	55	V	P	N	C	S	D	T	61
<i>H4N4_grayteal_Australia_2_79</i>	45	G	Q	P	C	S	N	D	51
<i>H6N5_mallard_Alberta_203_92</i>	43	P	S	P	C	N	T	T	49
<i>H3N8_duck_Ukraine_1_1963</i>	42	N	G	I	C	N	E	T	48

Figure 2

Representative alignment of the sequence environment of the conserved cysteine C49 that could either serve for intermolecular disulfide bridges or as palmitoylation site.

ing sequences were aligned with MAFFT (using L-INS-I settings [21]) and the resulting multiple alignment was visualized and annotated in Jalview [22]. A neighbour joining tree with pairwise gap deletion, Poisson correction as distance measure and 500 bootstrap replicates (generated with MEGA [23]) produces robust groupings consistent with previous studies [24] for the known NA subtypes (clustering of N1, N4, N5+N8 on one side and N2, N3, N6+N7+N9 on the other) and reliably places the new NA with other N1s. Interestingly, inside the N1 cluster, the new NA appeared close to the N1 of H5N1 avian flu viruses. The alignment and corresponding phylogenetic tree are available at <http://mendel.bii.a-star.edu.sg/SEQUENCES/H1N1/>.

Hence, we repeated the analysis (same protocol as outlined above) for a detailed mapping of only the N1 subtype family with the difference of allowing 95% sequence identity for sequences before 2009 but keeping all new NA sequences (as of April 29th). A characteristic clustering emerges (Figure 3) that roughly corresponds to host and geographic distributions, consistent with previous reports [25,26]. The observed clustering is robust in respect to the method used for tree generation (same for maximum parsimony or neighbour joining trees with JTT distance and gamma-distributed variable rates). The 2009 NA is part of a cluster of avian-like swine flu H1N1 strains predominantly found in European pigs. However, previous examples of human infections from swine flu are also part of the same cluster, for example from 2005 in Thailand [27]. This indicates that, similar to the current outbreak, closely related H1N1 strains have crossed species boundaries on previous occasions as also evidenced by further reports in the literature [28-30].

Moreover, neuroaminidases of these new H1N1 swine flu examples are more similar to H5N1 avian flu strains than other H1N1 variants found in the Americas or than that of the historic strains such as the 1918 Spanish flu [31]. This is surprising since avian flu strains typically have different hemagglutinin (HA) subtypes (e.g. H5N1). Combinations of HA and NA subtypes need to be fine-tuned to recognize the same type of sialic acid modifications to allow smooth interplay of the two proteins, which is important for the viral cycle [6]. These results support the notion that, also inside the family of N1 subtypes, a clear distinction can be made between avian-like H1N1, such as the one from the current outbreak, and other existing H1N1 strains.

Structural modelling and mapping of new mutations

The crystal structures of both the historic 1918 NA as well as the avian flu NA are available in complex with currently used drugs. We created a homology model of the new 2009 swine flu NA to map the sequence differences to the three-dimensional structure templates. Using Modeller [32], the sequence of the new neuraminidase [Genbank

Accession: ACP41107.1 <http://www.ncbi.nlm.nih.gov/protein/227809834>] was modelled 50 times onto multiple templates (PDB: [2hu4](#)[33], [3ckz](#)[34], [3b7e](#)[35], [3beq](#)[35]) and the resulting best model (as judged by DOPE score) further refined with short simulated annealing MD simulations in the presence of a bound inhibitor (zanamivir, oseltamivir or peramivir) as implemented in the Yasara Structure package [36]. The final atom-resolution models are available in PDB format at <http://mendel.bii.a-star.edu.sg/SEQUENCES/H1N1/>.

We mapped the level of residue conservation (calculated with the evolutionary trace algorithm [37]) from the multiple alignment of all NA subtypes to its corresponding position in the structure. The results show the strict conservation close to the neuraminidase catalytic site, which also serves as the drug binding pocket (Figure 4A). The remaining conserved patches (for example the sites around N104 or below N146) fit into each other and form the dimerization/tetramerization interfaces [35]. A model of the dimeric version is available in PDB format at <http://mendel.bii.a-star.edu.sg/SEQUENCES/H1N1/>.

Next, we compared the sequences of the new strain with the related H5N1 from avian flu and H1N1 from the Spanish flu (Figure 5). Among 387 residues that were structurally modelled, the "2009 H1N1" neuraminidase differs from the other two in 21 positions. The mapping to the structure (Figure 4B) shows that the novel sequence mutations are distributed all around the surface of the molecule leaving the hydrophobic core, but also the catalytic site, essentially untouched. Importantly, none of the new mutations appears sufficiently close to affect the drug binding pocket. For example, all 17 residues within 3 Å of the zanamivir molecule bound to the active site are fully conserved among all three strains. The closest mutation is the conservative V149I substitution at a distance of ~10 Å to zanamivir and ~7 Å to oseltamivir.

It has to be noted that indirect effects of the mutations that may alter the binding pocket also from a greater distance are difficult to assess and cannot be excluded. To this extent, we have analysed coevolution patterns in an extensive alignment of more than 6000 non-identical Influenza A neuraminidases to eventually identify connected networks of residues using the SCA algorithm [38] as implemented in [39], but no network that would connect the surface directly to the core and catalytic site was found (see supplementary material). In fact, all positions of the observed mutations are at the surface and naturally variable, as judged by the conservation and SCA analysis, which would rather indicate that they do not have an effect on the structure of the more distant binding pocket.

Thus, we conclude that the drug binding pocket remains unchanged in the new strain and, hence, the binding

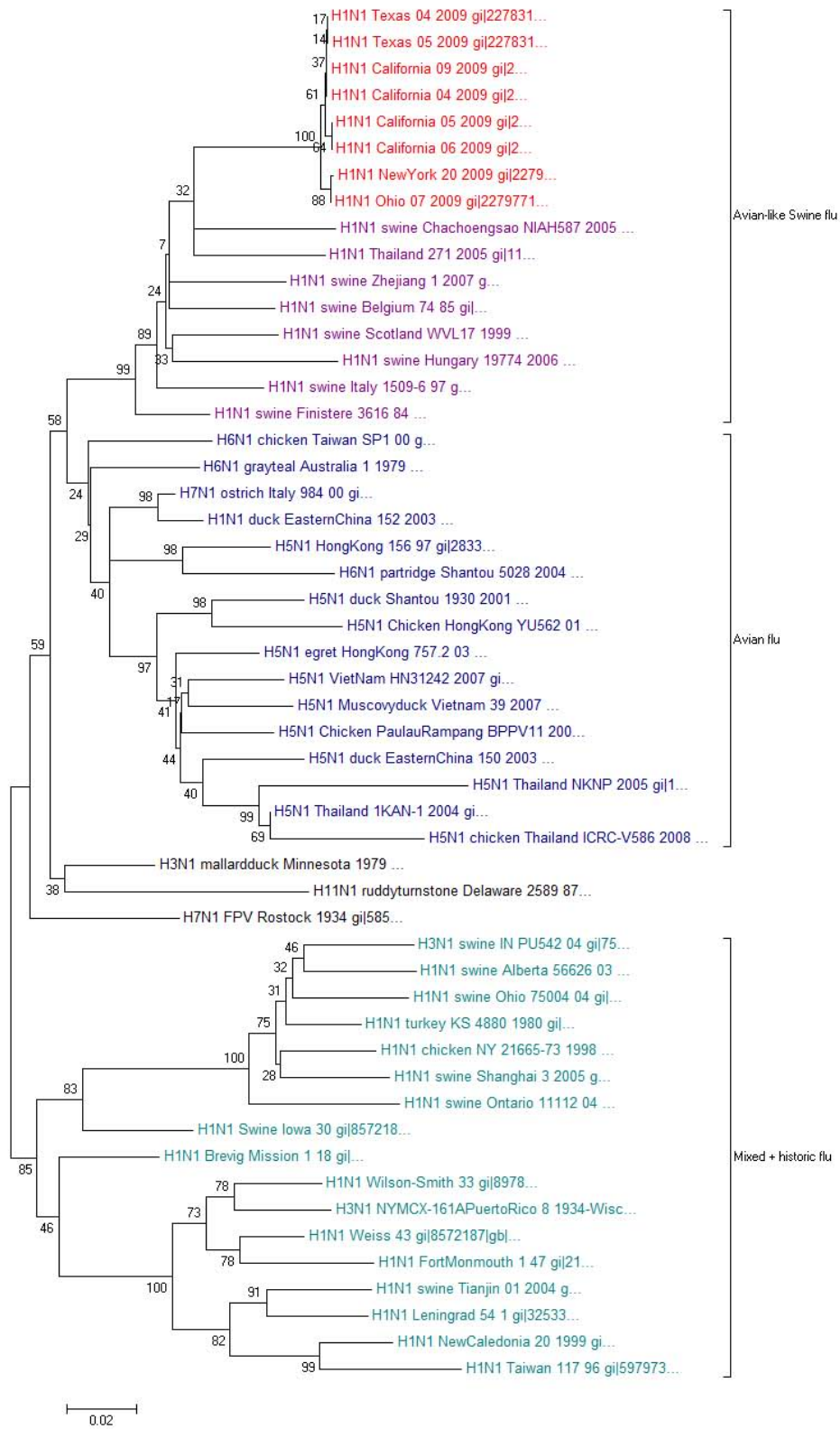


Figure 3
Phylogenetic tree of neuraminidase protein sequences of the N1 subtype family.

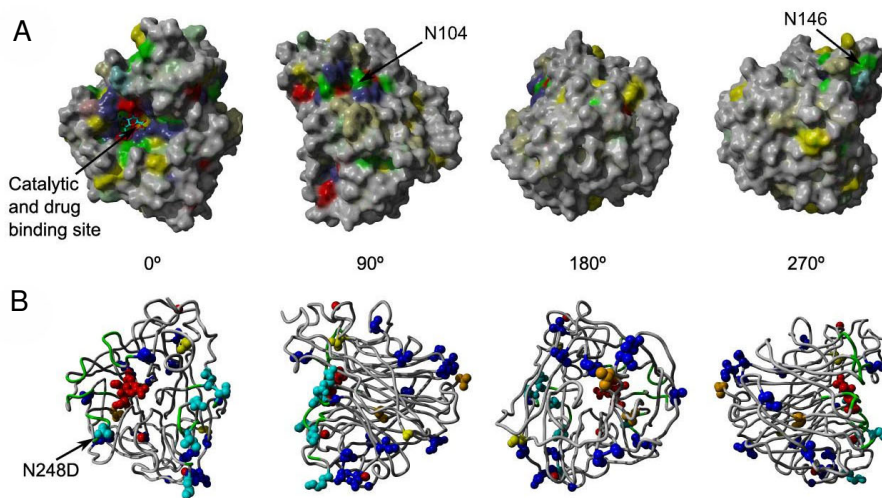


Figure 4

A) Surface representation of the structural model of the neuraminidase domain of the new strain in complex with zanamivir. Coloring is based on sequence conservation over all NA subtypes. Grey means no conservation. Other colors are according to physical properties: yellow ... hydrophobic, green ... polar, blue ... positive charge, red ... negative charge. Color intensities are proportional to strength of conservation. **B) Mapping of new mutations to structure.** Cyan colored residues are mutations at typical antibody recognition sites. Blue residues indicate differences to both the H5N1 avian flu as well as H1N1 from the 1918 Spanish flu. Yellow residues are intra-strain variations occurring in multiple patients of the 2009 H1N1 outbreak and orange if they have only been found in isolated patients, so far. Note that the intra-strain variation N248D is colored cyan since it is part of the antibody recognition site. The backbone of the antibody recognition sites is colored green and the bound drug and 3 calcium ions are shown in red.

behaviour of neuraminidase inhibitors such as oseltamivir (Tamiflu®) and zanamivir (Relenza®) should be unaffected. Indeed, initial clinical reports suggest that the new virus is susceptible to the two drugs [40]. Our findings support this notion and provide a molecular mechanism. Furthermore, the third currently tested neuraminidase inhibitor, peramivir, should also be effective since it also shares the same binding pocket.

Next, we review how the new mutations affect vaccine development through altering antibody interactions as well as antigenic regions. There are 3 crystal structures of related neuraminidases in complex with antibodies [41-43]. In Figure 5, we annotate residues that are within 3 Å distance to the respective bound antibody and, hence, crucial for the interaction. Interestingly, residues in sites recognized by both NC41 and NC10 antibodies appear mutated in the new strain. This would suggest that these old antibodies (that were originally directed against N9 neuraminidase) would probably not bind to NA of the new strain. Nevertheless, using the same regions as epitope may be a viable option for novel vaccine development.

Additionally, several other known antigenic regions, partially derived from surviving patients of previous flu outbreaks (e.g. H5N1), are reported in the literature [44-46] and their location is indicated in the alignment (Figure 5). Another extensive source of epitopes and antigens includ-

ing neuraminidase of influenza A viruses is the immune epitope database (IEDB) [47] and the complete mapping of epitopes for the new H1N1 NA sequence is available at <http://mendel.bii.a-star.edu.sg/SEQUENCES/H1N1/>. While several of the new mutations are found in antigenic regions, it is also apparent that they often occur on positions that are hardly conserved among different NA subtypes. Consequently, these regions are evolutionarily more flexible and may mutate fast. This increases the risk of evading antibody responses of human hosts acquired during previous flu infection or from vaccination.

Sequence variation among patients with the same 2009 H1N1 strain

After the first wave of new patient sequences arrived, it becomes clear that there are at least two major lineages that are distinguishable by only few mutations. Most notably, N248 has mutated to Aspartate (D248) in the New York infection cluster. All intra-strain mutations available before May 8th 2009 are indicated in Figures 4B and 5. As expected, they are predominantly found on the surface and in regions that are known to be variable from the conservation analysis. While the drug binding pocket remains unaffected by these most recent mutations, the N248D substitution changes a central part of an antibody recognition site. This has important consequences for vaccine development, forcing to either avoid this epitope or produce combined vaccines to account for the epitope

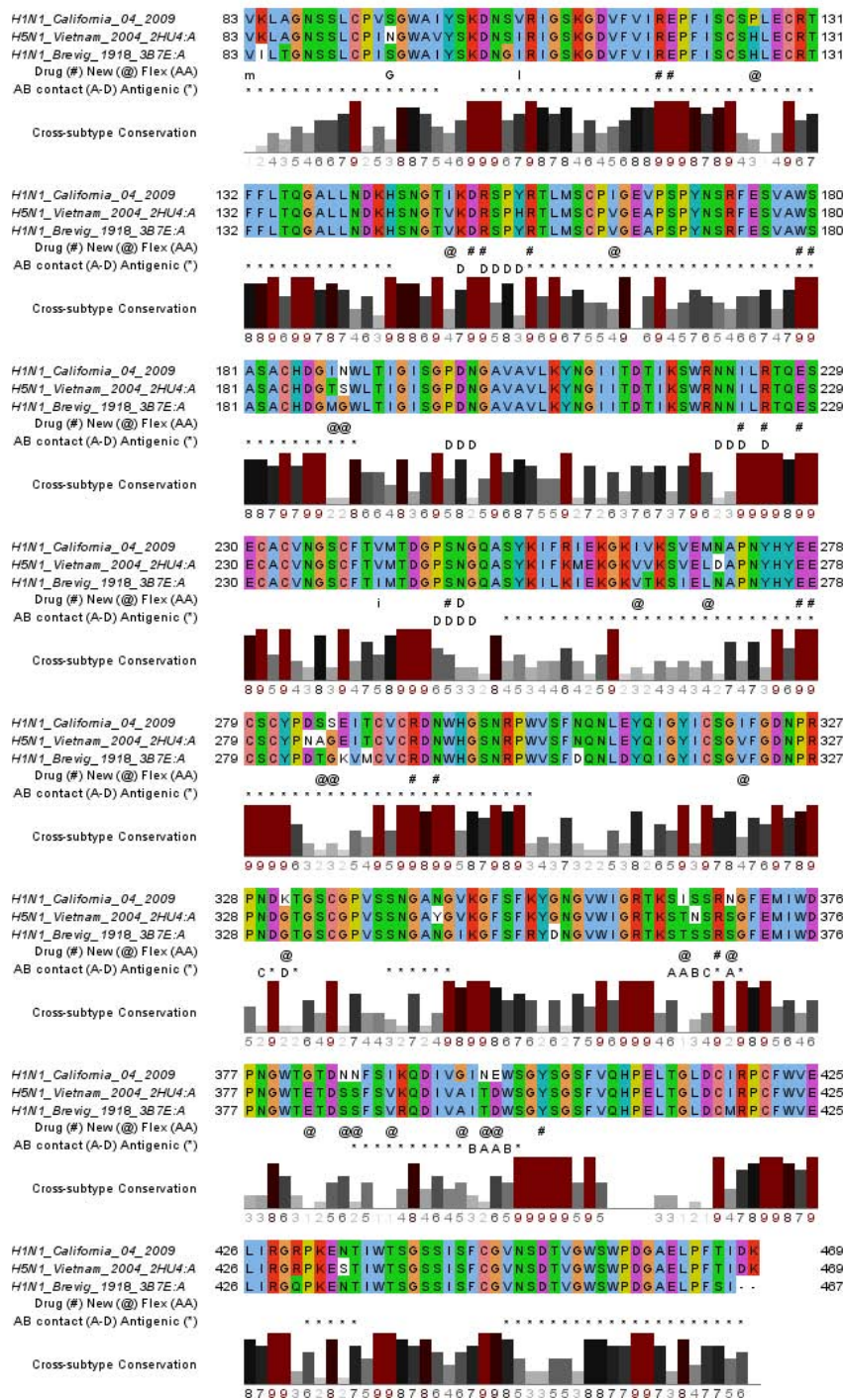


Figure 5
Alignment of the NA domain of the 2009 H1N1 strain with the sequences in crystal structures of H5N1 avian flu as well as H1N1 from the 1918 Spanish flu. Residues within 3 Å of the bound drug are indicated with "#", while residues that are different in the new strain compared to both other structures are marked with "@". Intra-strain variation (Flex [AA] in the first annotation line) is displayed as the respective mutated residue in capital letters if found in multiple patients (e.g. D for the N248D substitution) or lower-case (e.g. "i" for V241I) for single occurrences. In the second annotation row, antigenic regions are labelled as "*". Residues with < 3 Å contact to antibodies are labelled "A" for interactions derived from PDB:Incb, "B" from both PDB:Incb and PDB:Inmb, "C" from PDB:Inmb and "D" from PDB:2aep.

variation observed in different patient groups. Although the mutation pattern will become less transparent over time it may still serve to delineate chains of transmission, retrospectively.

Conclusion

In summary, we provide a sequence analysis and structural modelling of the neuraminidase from the 2009 H1N1 swine flu outbreak. Besides mapping of phylogenetic relationships to other strains, we find that the sequence variation in the new strain does not seem to affect the drug binding site but may very well alter common epitopes. To allow quick analysis of future mutations that could produce drug or vaccine-resistant strains, we provide a tool for 3D visualization of the neuraminidase structure models with mapping of drug and antibody recognition sites on the supplementary webpage <http://mendel.bii.a-star.edu.sg/SEQUENCES/H1N1/>.

Abbreviations

NA: neuraminidase; HA: hemagglutinin.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

SMS did the alignments and phylogenetic trees. FLS contributed the domain architecture analysis and helped with the phylogenetic analysis. MJ did the structural models and conservation mapping. RLTC contributed the study on antigenic regions and the Jmol visualization on the webpage. SMS and FE wrote the manuscript; all authors approved the final version.

Reviewers' comments

Reviewer 1

Sandor Pongor, International Centre for Genetic Engineering and Biotechnology, Trieste, Italy

In this work, the authors carry out a predictive analysis of the "2009 H1N1" (swine or Mexican flu) influenza A virus strain, based on phylogenetic analysis and 3D homology modelling. The results show that this strain is phylogenetically more closely related to European H1N1 swine flu and H5N1 avian flu rather than to the H1N1 counterparts in the Americas.

Homology-modeling of the neuraminidase reveals that the novel mutations are not likely to interfere with the active site so the currently used neuraminidase inhibitors (oseltamivir, zanamivir and peramivir) will be effective against the new virus strain.

The subject is very timely and the approach is adequate. The authors may want to include analysis of more patient data that were published since the analysis was completed. More and more sequences are being published from all over the world that might

be worthwhile to include into this analysis. The authors may consider establishing a periodically updated homepage, if appropriate. In summary, the analysis is careful and carried out in a commendable fashion, and the findings are highly significant.

Response

Indeed, there have been additional "2009 H1N1" neuraminidase sequences since the submission of this manuscript for review. Between April 29th and May 8th, 45 new sequences became available. Overall, 3 mutations occur in multiple (S95G, V106I and N248D) while 2 mutations are restricted to single patient virus isolates (V83M, V241I), so far. We have included a section about these mutations identified in new patient sequences and mapped them to the structure. As discussed in the main text, the intra-strain variation is typically found on the surface and at positions expected to be variable as judged by the conservation analysis among all NA subtypes. Interestingly, one of the mutations among patients (N248D) is critically affecting one of the antibody binding sites.

As the virus will continue to evolve, new mutations will become available and the best way we have found to allow quick mapping and update of new sequence variation in respect to drug and antibody binding sites is to give full access to users/readers via a 3D structure visualization tool at the supplementary webpage <http://mendel.bii.a-star.edu.sg/SEQUENCES/H1N1/> (instructions to map new mutations are given, including an example).

Reviewer 2

L. Aravind, National Center for Biotechnology Information, Bethesda, MD, USA

Maurer-Stroh et al discuss sequence features of the neuraminidase gene of the virus behind the latest influenza outbreak, 2009 H1N1. Of the features of interest they note that this strain has acquired a novel glycosylation consensus site that could be surface accessible. Of course, it remains unknown if this site is indeed used for modification. The key finding in the paper is that the binding cavity for neuraminidase inhibitors is unaffected as suggested by molecular modeling. This finding is of significance in the current situation of an outbreak with pandemic potential. However, one issue needs to be highlighted in this regard – mutations far away from the binding site can potentially affect the shape and or binding affinity of the binding pocket. These are not always captured by homology models, especially the issue of affinity. In principle, authors could use a conservation patterns or co-evolution measures (e.g. as in PMID: 10514373) to determine if there are interaction chains that might connect distant residues to the active site. In the least it would be useful to provide the caveat of distant changes affecting affinity in the current paper.

Response

We totally agree with the referee that also mutations at a greater distance may affect the binding pocket under spe-

cial circumstances and we have added a new paragraph to the manuscript. However, we have to admit that it remains essentially impossible to quantify the influence of non-direct interactions by theoretical means unambiguously. In our experience, co-evolution measures such as the one proposed and several others (PMID: 18056067) produce high rates of false positives and, therefore, the interpretation is difficult. We did the requested analysis with Ranganathan's SCA algorithm to eventually identify connected networks of residues using the webserver implementation at the Gerstein lab, but no network that would connect the surface directly to the core and catalytic site was found (see supplementary webpage for full details). In fact, all positions of the observed mutations are at the surface and naturally variable, as judged by the conservation and SCA analysis, which would rather indicate that they do not have an effect on the structure of the more distant binding pocket. A possible problem why the SCA could not work in this case is the high level of sequence similarity among the neuraminidases which only gives limited numbers of informative correlated mutations. This is a totally different scenario from alignments of highly divergent sequences that still have the same fold, such as the small PDZ domains analyzed by Ranganathan. In the case of a diverse family with shared fold, correlated mutations are indicative of allowed fluctuations also among structurally important residues. However, with the neuraminidases, most variation can be attributed to surface residues, which makes sense given the pressure to avoid immune responses. Additionally, sequence sampling of neuraminidases is not independent but biased by transmission chains and clusters of outbreaks. A possible but also not necessarily more precise alternative to judge indirect effects on drug binding would be to run free energy simulations with the bound drug to judge changes of affinity caused by the mutations but this is a tricky and time-consuming endeavour that would burst the scope of this current manuscript.

unspecific consensus motif: change to "non-specific"

Response
changed.

Phylogenetic analysis: While for sequences at this range of similarity Poisson correction may not have negative consequences it is definitely better to repeat the analysis with JTT and variable rates to see if the clustering remains the same or changes drastically.

Response
We confirmed that the tree clustering inside the N1 subtype family is robust regarding different tree generation methods and added this also to the text. We also provide the suggested JTT distance tree with variable rates as supplementary at <http://mendel.bii.a-star.edu.sg/SEQUENCES/H1N1/>.

"This indicates that, similar to the current outbreak, scenarios of breaches in the species barrier between human and pigs have already arisen out of closely related H1N1 strains as also evidenced by further reports in the literature [26,27]." The wording of this sentence is somewhat unclear. It appears that the authors wish to state that H1N1 like strains have crossed species boundaries on other occasions, but this is not necessarily clear in the sentence.

Response
changed to "This indicates that, similar to the current outbreak, closely related H1N1 strains have crossed species boundaries on previous occasions as also evidenced by further reports in the literature".

"These results support the notion that, also inside the family of N1 subtypes, a clear distinction can be made to distinguish avian-like H1N1, such as the one from the current outbreak, from other existing H1N1 strains." What would be the explanation for this? A recent recombination between an avian-like H1N1 or has it diverged from other avian like H1N1 with the recombination occurring much earlier. Could this information be superimposed in phylogenetic context on current figure 3?

Response
This is, of course, a very interesting question and difficult to deduce from the phylogenetic tree of a single protein without molecular clock, as in our case. However, this and similar questions have already been analyzed to quite some detail for the existing strains. The current knowledge of the scenario for H1N1 is that around the late 1970s to early 1980s, a human-avian reassortant virus started to be detected in European pigs as host (PMID: 8091678). Then, this avian-like swine flu started to move from the European continent to the UK in the early 90s (PMID: 9049404). More reassortments and emergence of the N2 subtype that quickly spread is also well documented. This phylogenetic analysis plugs the new H1N1 strain into the already known clusterings among the NA subtypes and within the N1 family.

References

1. Butler D: **Swine flu goes global.** *Nature* 2009, **458**:1082-1083.
2. Cohen J, Enserink M: **Infectious diseases. As swine flu circles globe, scientists grapple with basic questions.** *Science* 2009, **324**:572-573.
3. Vogel G: **Public health. SARS and bird flu veteran to take WHO helm.** *Science* 2006, **314**:1067.
4. Ooi HS, Kwo CY, Wildpaner M, Sirota FL, Eisenhaber B, Maurer-Stroh S, Wong WC, Schleiffer A, Eisenhaber F, Schneider G: **ANNIE: integrated de novo protein sequence annotation.** *Nucleic Acids Res* 2009 in press.
5. Eisenhaber F: **Prediction of Protein Function: Two Basic Concepts and One Practical Recipe.** In *Discovering biomolecular mechanisms with computational biology* Edited by: Eisenhaber F. Georgetown: Landes Biosciences and Eurekah.com; 2006:39-54.
6. Wagner R, Matrosovich M, Klenk H: **Functional balance between haemagglutinin and neuraminidase in influenza virus infections.** *Reviews in Medical Virology* 2002, **12**:159-166.
7. Blok J, Air GM, Laver WG, Ward CW, Lilley GG, Woods EF, Roxburgh CM, Inglis AS: **Studies on the size, chemical composition, and partial sequence of the neuraminidase (NA) from type A influenza viruses show that the N-terminal region of**

- the NA is not processed and serves to anchor the NA in the viral membrane. *Virology* 1982, **119**:109-121.
8. Mitnaul L, Castrucci M, Murti K, Kawaoka Y: **The cytoplasmic tail of influenza A virus neuraminidase (NA) affects NA incorporation into virions, virion morphology, and virulence in mice but is not essential for virus replication.** *J Virol* 1996, **70**:873-879.
 9. Kundu A, Avalos R, Sanderson C, Nayak D: **Transmembrane domain of influenza virus neuraminidase, a type II protein, possesses an apical sorting signal in polarized MDCK cells.** *J Virol* 1996, **70**:6508-6515.
 10. Eisenhaber B, Eisenhaber F: **Posttranslational modifications and subcellular localization signals: indicators of sequence regions without inherent 3D structure?** *Curr Protein Pept Sci* 2007, **8**:197-203.
 11. Eisenhaber F, Eisenhaber B, Maurer-Stroh S: **Prediction of Post-translational modifications from amino acid sequence: Problems, pitfalls, methodological hints.** In *Bioinformatics and Genomes: Current Perspectives* Edited by: Andrade MM. Wyomondham: Horizon Scientific Press; 2003:81-105.
 12. Ward CW, Elleman TC, Azad AA: **Amino acid sequence of the Pronase-released heads of neuraminidase subtype N2 from the Asian strain A/Tokyo/3/67 of influenza virus.** *Biochem J* 1982, **207**:91-95.
 13. Wang N, Glidden EJ, Murphy SR, Pearce BR, Hebert DN: **The cotranslational maturation program for the type II membrane glycoprotein influenza neuraminidase.** *J Biol Chem* 2008, **283**:33826-33837.
 14. Hausmann J, Kretzschmar E, Garten W, Klenk HD: **Biosynthesis, intracellular transport and enzymatic activity of an avian influenza A virus neuraminidase: role of unpaired cysteines and individual oligosaccharides.** *J Gen Virol* 1997, **78**(Pt 12):3233-3245.
 15. Ward CW, Colman PM, Laver WG: **The disulphide bonds of an Asian influenza virus neuraminidase.** *FEBS Lett* 1983, **153**:29-33.
 16. Chen BJ, Takeda M, Lamb RA: **Influenza virus hemagglutinin (H3 subtype) requires palmitoylation of its cytoplasmic tail for assembly: M1 proteins of two subtypes differ in their ability to support assembly.** *J Virol* 2005, **79**:13673-13684.
 17. Sugrue RJ, Belshe RB, Hay AJ: **Palmitoylation of the influenza A virus M2 protein.** *Virology* 1990, **179**:51-56.
 18. Barman S, Nayak DP: **Analysis of the Transmembrane Domain of Influenza Virus Neuraminidase, a Type II Transmembrane Glycoprotein, for Apical Sorting and Raft Association.** *J Virol* 2000, **74**:6538-6545.
 19. Altschul SF, Koonin EV: **Iterated profile searches with PSI-BLAST – a tool for discovery in protein databases.** *Trends Biochem Sci* 1998, **23**:444-447.
 20. Li W, Godzik A: **Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences.** *Bioinformatics* 2006, **22**:1658-1659.
 21. Katoh K, Kuma K, Toh H, Miyata T: **MAFFT version 5: improvement in accuracy of multiple sequence alignment.** *Nucleic Acids Res* 2005, **33**:511-518.
 22. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ: **Jalview Version 2 – a multiple sequence alignment editor and analysis workbench.** *Bioinformatics* 2009, **25**:1189-1191.
 23. Kumar S, Nei M, Dudley J, Tamura K: **MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences.** *Brief Bioinformatics* 2008, **9**:299-306.
 24. Liu S, Ji K, Chen J, Tai D, Jiang W, Hou G, Chen J, Li J, Huang B: **Panorama phylogenetic diversity and distribution of Type A influenza virus.** *PLoS ONE* 2009, **4**:e5022.
 25. Brown IH, Ludwig S, Olsen CW, Hannoun C, Scholtissek C, Hinshaw VS, Harris PA, McCauley JW, Strong I, Alexander DJ: **Antigenic and genetic analyses of H1N1 influenza A viruses from European pigs.** *J Gen Virol* 1997, **78**(Pt 3):553-562.
 26. Chen JM, Ma HC, Chen JW, Sun YX, Li JM, Wang ZL: **A preliminary panorama of the diversity of N1 subtype influenza viruses.** *Virus Genes* 2007, **35**:33-40.
 27. Komadina N, Roque V, Thawatsupha P, Rimando-Magalong J, Waicharoen S, Bomasang E, Sawanpanyalert P, Rivera M, Iannello P, Hurt AC, Barr IG: **Genetic analysis of two influenza A (H1) swine viruses isolated from humans in Thailand and the Philippines.** *Virus Genes* 2007, **35**:161-165.
 28. Rimmelzwaan GF, de Jong JC, Bestebroer TM, van Loon AM, Claas EC, Fouchier RA, Osterhaus AD: **Antigenic and genetic characterization of swine influenza A (H1N1) viruses isolated from pneumonia patients in The Netherlands.** *Virology* 2001, **282**:301-306.
 29. Newman AP, Reisdorf E, Beinemann J, Uyeki TM, Balish A, Shu B, Lindstrom S, Achenbach J, Smith C, Davis JP: **Human case of swine influenza A (H1N1) triple reassortant virus infection, Wisconsin.** *Emerging Infect Dis* 2008, **14**:1470-1472.
 30. Gregory V, Bennett M, Thomas Y, Kaiser L, Wunderli W, Matter H, Hay A, Lin YP: **Human infection by a swine influenza A (H1N1) virus in Switzerland.** *Arch Virol* 2003, **148**:793-802.
 31. Tumpey TM, Basler CF, Aguilar PV, Zeng H, Solórzano A, Swaine DE, Cox NJ, Katz JM, Taubenberger JK, Palese P, Garcia-Sastre A: **Characterization of the reconstructed 1918 Spanish influenza pandemic virus.** *Science* 2005, **310**:77-80.
 32. Eswar N, Eramian D, Webb B, Shen M, Sali A: **Protein structure modeling with MODELLER.** *Methods Mol Biol* 2008, **426**:145-159.
 33. Russell RJ, Haire LF, Stevens DJ, Collins PJ, Lin YP, Blackburn GM, Hay AJ, Gamblin SJ, Skehel JJ: **The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design.** *Nature* 2006, **443**:45-49.
 34. Collins PJ, Haire LF, Lin YP, Liu J, Russell RJ, Walker PA, Skehel JJ, Martin SR, Hay AJ, Gamblin SJ: **Crystal structures of oseltamivir-resistant influenza virus neuraminidase mutants.** *Nature* 2008, **453**:1258-1261.
 35. Xu X, Zhu X, Dwek RA, Stevens J, Wilson IA: **Structural characterization of the 1918 influenza virus H1N1 neuraminidase.** *J Virol* 2008, **82**:10493-10501.
 36. Krieger E, Koraimann G, Vriend G: **Increasing the precision of comparative models with YASARA NOVA – a self-parameterizing force field.** *Proteins* 2002, **47**:393-402.
 37. Lichtarge O, Yao H, Kristensen DM, Madabushi S, Mihalek I: **Accurate and scalable identification of functional sites by evolutionary tracing.** *J Struct Funct Genomics* 2003, **4**:159-166.
 38. Lockless SW, Ranganathan R: **Evolutionarily conserved pathways of energetic connectivity in protein families.** *Science* 1999, **286**:295-299.
 39. Yip KY, Patel P, Kim PM, Engelman DM, McDermott D, Gerstein M: **An integrated system for studying residue coevolution in proteins.** *Bioinformatics* 2008, **24**:290-292.
 40. Centers for Disease Control and Prevention (CDC): **Update: drug susceptibility of swine-origin influenza A (H1N1) viruses, April 2009.** *MMWR Morb Mortal Wkly Rep* 2009, **58**:433-435.
 41. Tulip WR, Varghese JN, Webster RG, Laver WG, Colman PM: **Crystal structures of two mutant neuraminidase-antibody complexes with amino acid substitutions in the interface.** *J Mol Biol* 1992, **227**:149-159.
 42. Malby RL, Tulip WR, Harley VR, McKimm-Breschkin JL, Laver WG, Webster RG, Colman PM: **The structure of a complex between the NC10 antibody and influenza virus neuraminidase and comparison with the overlapping binding site of the NC41 antibody.** *Structure* 1994, **2**:733-746.
 43. Venkatramani L, Bochkareva E, Lee JT, Gulati U, Graeme Laver W, Bochkarev A, Air GM: **An epidemiologically significant epitope of a 1998 human influenza virus neuraminidase forms a highly hydrated interface in the NA-antibody complex.** *J Mol Biol* 2006, **356**:651-663.
 44. Fanning TG, Reid AH, Taubenberger JK: **Influenza A virus neuraminidase: regions of the protein potentially involved in virus-host interactions.** *Virology* 2000, **276**:417-423.
 45. Castrucci MR, Bilsel P, Kawaoka Y: **Attenuation of influenza A virus by insertion of a foreign epitope into the neuraminidase.** *J Virol* 1992, **66**:4647-4653.
 46. Ho HT, Qian HL, He F, Meng T, Szyporta M, Prabhu N, Prabakaran M, Chan KP, Kwang J: **Rapid Detection of H5N1 Subtype Influenza Viruses by Antigen-Capture Enzyme-Linked Immunosorbent Assay Using H5 and N1 specific Monoclonal Antibodies.** *Clin Vaccine Immunol* 2009, **16**:726-32.
 47. Bui H, Peters B, Assarsson E, Mbowuikie I, Sette A: **Ab and T cell epitopes of influenza A virus, knowledge and opportunities.** *Proc Natl Acad Sci USA* 2007, **104**:246-251.