ELSEVIER

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM



Short communication

Homology modeling study toward identifying structural properties in the HA2 B-loop that would influence the HA1 receptor-binding site



Marni E. Cueno^{a,*}, Kenichi Imai^a, Kazufumi Shimizu^b, Kuniyasu Ochiai^{a,*}

- ^a Department of Microbiology, Dental Research Center, Nihon University School of Dentistry, Tokyo 101-8310, Japan
- ^b Department and Division of Obstetrics and Gynecology, Nihon University School of Medicine, Tokyo 173-8610, Japan

ARTICLE INFO

Article history: Accepted 27 May 2013 Available online 13 June 2013

Keywords: B-loop Hemagglutinin Homology modeling Influenza A H1N1 subtype Receptor-binding site

ABSTRACT

Influenza hemagglutinin (HA) consists of a fibrous globular stem (HA2) inserted into the viral membrane supporting a globular head (HA1). HA1 receptor-binding has been hypothesized to be structurally correlated to the HA2 B-loop, however, this was never fully understood. Here, we elucidated the structural relationship between the HA2 B-loop and the HA1 receptor-binding site (RBS). Throughout this study, we analyzed 2486 H1N1 HA homology models obtained from human, swine and avian strains during 1976–2012. Quality of all homology models were verified before further analyses. We established that amino acid residue 882 is putatively strain-conserved and differs in the human (K882), swine (H882) and avian (N882) strains. Moreover, we observed that the amino acid at residue 882 and, similarly, its orientation has the potential to influence the HA1 RBS diameter measurements which we hypothesize may consequentially affect influenza H1N1 viral infectivity, immune escape, transmissibility, and evolution.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

Influenza A virus is an RNA virus that initiates infection to the host cell by binding to sialic acids found in the cell surface [1,2]. Binding of the influenza virus to the host cell is primarily attributed to the hemagglutinin (HA) glycoprotein. HA is a homotrimeric glycoprotein that constitutes most of the virus surface [3,4] and consists of a fibrous globular stem inserted into the viral membrane supporting a globular head containing three sialic acid binding sites [3,5,6]. The HA glycoprotein has multiple roles in the virus life cycle [6,7], most notably receptor-binding.

The early stages of influenza infection and transmission involves the binding of the HA glycoprotein to host cell receptors and the ability of HA to bind to sialic acids found in the host surface determines which host can be infected [8]. HA receptor-binding properties are a critical determinant of influenza evolution [3,9] and are influenced by several factors such as glycosylation [3], neutralizing antibodies [3], and inter-residue atomic interactions [10]. HA receptor-binding avidity has been correlated to HA antigenic drift [11] which is hypothesized to, subsequently, influence neuraminidase antigenic drift [12]. This highlights the importance of

E-mail addresses: marni.cueno@nihon-u.ac.jp (M.E. Cueno), ochiai.kuniyasu@nihon-u.ac.jp (K. Ochiai).

the HA receptor-binding domain (RBD) in viral infectivity, immune escape, transmissibility and evolution [1,8,11,13].

It was previously shown that the RBD is stabilized by salt bridges existing between the HA1 110-helix and the HA2 B-loop [7,14,15]. A salt bridge is a combination of two non-covalent interactions (hydrogen bonding and electrostatic interactions) and, in the influenza HA, salt bridges are commonly affected by pH [14–17]. The most common salt bridges occur from the anionic carboxylate of either aspartic acid or glutamic acid and the cationic ammonium from lysine or the guanidinium of arginine [17]. These amino acid residues can be found within the HA1 110-helix and HA2 B-loop [7,14]. Salt bridges in the influenza HA have been correlated to the stability of the ectodomain and coordination of cooperative motions essential for the fusion process [14] which is suggested to induce a partial relaxation in the HA [18]. This would suggest that any amino acid alteration that may influence the salt bridge between the HA1 110-helix and HA2 B-loop may similarly influence the HA RBD structure. HA1 receptor-binding has been hypothesized to be structurally associated to the HA2 B-loop [14], however, this was never fully understood.

Here, we elucidated possible structural properties found in the HA2 B-loop that could putatively influence the HA1 receptor-binding site. We utilized H1N1 HA amino acid sequences obtained from human, swine and avian strains during 1976–2012 and generated HA homology models for each sequence. In addition, we verified the quality of all HA homology models before further analyses. Moreover, we found that a strain-conserved amino acid residue and its orientation within the HA2 B-loop of each host strain could

^{*} Corresponding authors at: Department of Microbiology, Nihon University School of Dentistry, 1-8-13 Kanda-Surugadai, Chiyoda-ku, Tokyo 101-8310, Japan. Tel.: +81 332198125.

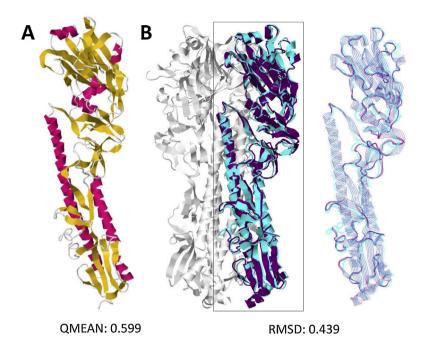


Fig. 1. Quality estimation of influenza A H1N1 hemagglutinin homology models generated. (A) Ribbon structure and model quality estimation of a representative 1918 influenza A H1N1 HA homology model. QMEAN score is indicated below. (B) Structural alignment of representative 1918 influenza A H1N1 HA homology model and crystal structure. RMSD value is indicated below. (*Left panel*) Ribbon structure of the 1918 H1N1 HA crystal structure superimposed to the 1918 H1N1 HA homology model. Chains C and D of the crystal structure (purple) superimposed to the homology model (blue) are indicated. Superimposed area is boxed. (*Right panel*) Strand structure of the superimposed area. Chains C and D of the crystal structure (purple) superimposed to the homology model (blue) are indicated. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

putatively affect the HA1 receptor-binding site (RBS) diameter measurement.

2. Methodology

2.1. Data mining

H1N1 HA amino acid sequences of human, swine and avian strains from 1976 to 2012 were collected from the National Center for Biological Information (NCBI) Web site. A total of 2486 amino acid sequences were obtained and analyzed. Representative strains used were the following: 1918 human strain with Genebank accession number AAD17229.1; 2008 human, swine, and avian strains with Genebank accession numbers ADC45736, ADM18095, and BAJ14545, respectively; and 1988 human, swine, and avian strains with Genebank accession numbers ABU80400, ABR29595, and ACD88703, respectively. Representative crystal structure used for superimposition [19] was the 1918 human strain (PDB ID: 4GXX). Throughout the study, we followed the numbering scheme as previously published [20] and we designated subscripts 1 and 2 to refer to HA₁ and HA₂, respectively.

2.2. Homology modeling and model quality estimation

We generated the HA protein structure using the Phyre server [21]. The most recent version (Phyre2) was utilized in this study. Briefly, the Phyre server utilizes known protein structures taken from the SCOP and PDB database. Each predicted 3D structure generated is scanned against a non-redundant sequence database. Each user-submitted sequence is scanned against the non-redundant sequence database afterwards a profile is constructed. The query secondary structure is predicted by three independent secondary structure prediction programs, namely: Psi-Pred [22], SSPro [23] and JNet [24]. Subsequently, confidence values of each program

are averaged and the consensus is calculated. Profile and secondary structures are scanned against the fold library using a profile–profile algorithm [25] and this process returns a score based on the alignment rankings. These scores are fitted to an extreme value distribution to generate an *E*-value wherein the top ten highest scoring alignments are then used to construct full 3D models. We compared the HA2 B-loops from all influenza strains obtained from 1976 to 2012, identified the strain-conserved amino acid residues, and tallied the consensus strain-conserved amino acid residue occurring from 1976 to 2012. Altered forms of the HA homology models were likewise generated by substituting the identified strain-conserved amino acid residues found in one strain with those found in other strains.

Qualitative Model Energy Analyses (QMEAN) scores were determined to estimate the quality of each homology model generated [26,27]. QMEAN scoring function is based on the linear combination of six structural descriptors and reflects the predicted global model reliability ranging from 0 to 1. Moreover, generated HA homology models were superimposed with known H1N1 HA crystal structures using SuperPose [19] to determine the Root Mean Square Deviation (RMSD) values of the superimposed $C\alpha$ backbone. RMSD values closer to 0 would insinuate that the homology models are reliable.

2.3. Strain-conserved amino acid residue identification and structural analyses

We determined the strain-conserved amino acid residues found in the HA2 B-loop and correlated it to the HA1 RBS. Identified strain-conserved amino acid residues were substituted with amino acid residues observed in other host strains and the corresponding homology model was similarly generated. Moreover, amino acid orientation of the strain-conserved amino acid residues were likewise established and correlated to the HA1 RBS.

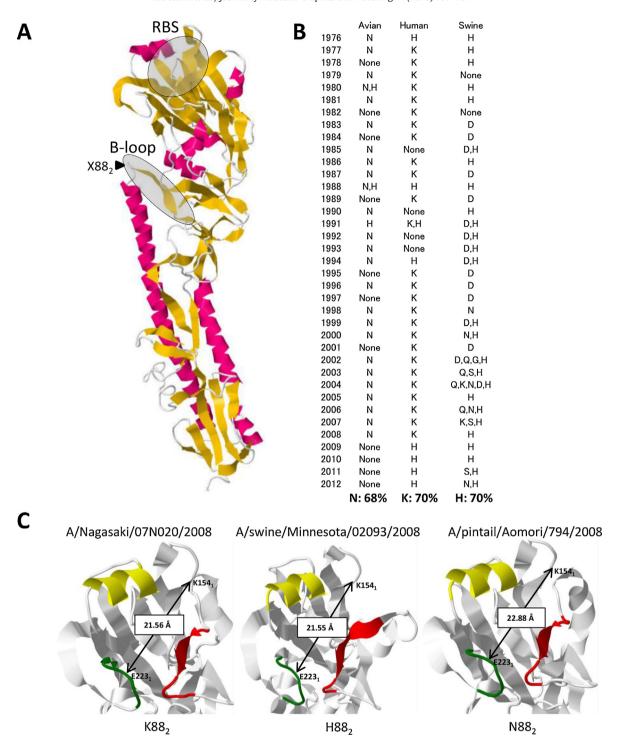


Fig. 2. Amino acid residue 882 differs in the human, swine and avian strains of the influenza A H1N1 hemagglutinin. (A) Representative influenza A H1N1 HA homology model. HA1 receptor-binding site (RBS) and HA2 B-loop are encircled and shaded. Strain-conserved amino acid residue found in the HA2 B-loop is indicated by an arrow and represented by X. (B) Varying amino acids at residue 882 of avian, human, and swine strains from 1976 to 2012. (*Upper panel*) Varying amino acids at residue 882 detected in each year. (*Lower panel*) Predominant amino acid residue 882 found in each host. (C) Representative HA1 RBS diameters of the 2008 avian, human, and swine strains. HA1 RBS diameter was measured between K1541 and E2231. Amino acid residue 882 is indicated below. RBS rim composed of the 220-loop (green), 130-loop (red), and 190-helix (yellow) is shown. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

We analyzed all homology models using the Jmol applet [28]. Briefly, Jmol is a free open source applet developed for the interactive display of three-dimensional chemical structures which also have features for biomolecules. Jmol uses molecular coordinate files for input. The most usual files (pdb, mol, xyz, cif) are accepted and are automatically recognized upon reading the file. Jmol has been designed to be modular since the file parser is independent

of the core functionality. In addition to reading molecular models, Jmol can also read script files and allow instructions or commands to be applied to the predicted model. We used Jmol to determine the RBS rim positions (220-loop, 130-loop and 190-helix) as previously published [13,29]. Throughout the study, the ribbon structure of HA homology models and their altered forms were used. Considering the HA1 RBS positions are found in the membrane distal-tip

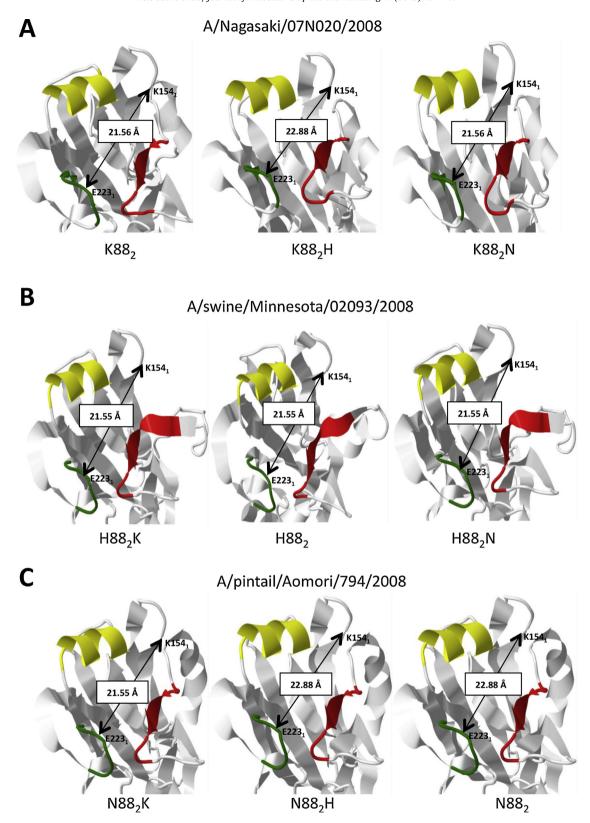


Fig. 3. Amino acid residue 88₂ influences the HA1 RBS diameter in influenza A H1N1 strains. Strain-conserved amino acid residue 88₂ and the substituted forms of the (A) human; (B) swine; and (C) avian influenza strains with their corresponding HA1 RBS diameter are indicated. HA1 RBS diameter was measured between K154₁ and E223₁. Amino acid at residue 88₂ is indicated below. RBS rim composed of the 220-loop (green), 130-loop (red), and 190-helix (yellow) is shown. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

of the RBD which coincidently can be circularly aligned [13,29], we used the RBS diameter to determine any HA1 variations. Differences in RBS diameter were established by measuring the distance between K154₁ and E223₁. The K154₁ residue is found in the

220-loop and in the opposite region found across the RBS is the $E223_1$ residue which, coincidently, is part of the RBD region that is affected whenever amino acid residue 88_2 was substituted. This is consistent with a previously published hypothesis

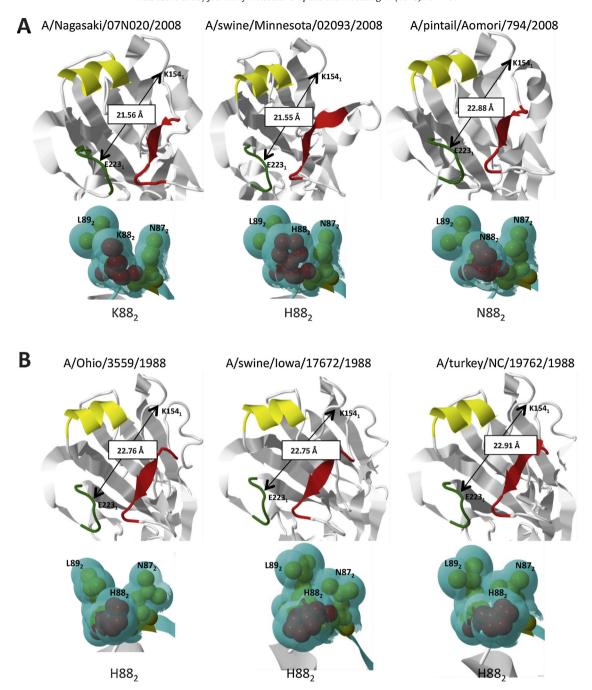


Fig. 4. Amino acid orientation at residue 88₂ affects the HA1 RBS diameter in influenza A H1N1 strains. (A) Representative 2008 strains with similar amino acid orientation but varying amino acids at residue 88₂. (B) Representative 1988 strains with varying amino acid orientation but with the same amino acid at residue 88₂. (*Upper panels*) HA1 RBS diameter measurements. HA1 RBS diameter was measured between K154₁ and E223₁. (*Lower panels*) Amino acid orientation at residue 88₂. Particular amino acid at residue 88₂ is indicated below. RBS rim composed of the 220-loop (green), 130-loop (red), and 190-helix (yellow) is shown.

[14] and would make both amino acid residues ideal for measuring the RBS diameter. All distance measurements are made in Å.

3. Results and discussion

3.1. Generated HA homology model estimation and superimposition

The availability of a protein structural model is a key element to understand the biological processes at the molecular level [26,30]. In generating homology models from amino acid sequences, it is important to establish the accuracy and quality

of each model. In producing experimental or theoretical models of protein structures, it is always important to assess model accuracy and reliability before proceeding to further analyses [31]. To confirm the quality of each HA homology model generated, model quality estimation using QMEAN and SuperPose were performed. Throughout this study, we utilized 2486 amino acid sequences of the H1N1 HA subtype obtained from 1976 to 2012 and generated HA homology models with QMEAN scores >0.50 suggesting that the homology models produced are reliable (Fig. 1A). In addition, superimposing the representative 1918 H1N1 HA homology model with the 1918 H1N1 HA crystal structure showed RMSD values <0.5 further indicating that the homology model is reliable (Fig. 1B).

3.2. Both the amino acid at residue 88_2 and its orientation influences the HA1 RBS diameter

The location of both the HA1 RBS and the HA2 B-loop (shaded in gray) in the representative HA homology model is shown in Fig. 2A (left panel). The RBS is located at the membrane-distal tip of the RBD [10,29] where each site is comprised of three secondary structural elements (190-helix, 130-loop and 220-loop) that together form the RBS rim [13,29]. The HA2 B-loop is hypothesized to affect the HA1 RBS [14] and in order to establish the amino acid residue/s in the HA2 B-loop that can influence the HA1 RBS, 2486 HA homology models obtained from 1976 to 2012 were compared and the conserved amino acid residues were determined. We found that within the HA2 B-loop, amino acid residue 882 is strain-conserved and, likewise, we established that amino acid residue 882 differs in the human (K88₂), swine (H88₂), and avian (N88₂) strains (Fig. 2A, right panel). Furthermore, utilizing the 2008 human, swine, and avian strains as representative influenza strains, we showed that the RBS diameter differs in each species (Fig. 2B). Our results would suggest that the varying HA1 RBS diameter measurements are possibly correlated with the strain-conserved amino acid residue 882 found in the HA2 B-loop.

To further confirm whether amino acid variations at residue 88₂ could putatively affect the HA1 RBS diameter measurements, we substituted the strain-conserved amino acid residue 88₂ of the representative human, swine, and avian strains with those identified from other species (Fig. 2A, bottom right panel). We found that in human and avian strains, RBS diameter measurements changed when the amino acid at residue 88₂ was substituted (Fig. 3A and C). This shows that amino acid residue 88₂ could potentially affect RBS diameter measurements. Surprisingly, RBS diameter in the swine strain was unchanged after amino acid substitution (Fig. 3B). This led us to consider that the amino acid at residue 88₂ alone is not enough to influence the HA1 RBS diameter. We suspect that amino acid orientation at residue 88₂ similarly plays a role in influencing the HA1 RBS diameter.

To determine whether amino acid orientation at residue 882 has the potential to influence the HA1 RBS diameter, we verified the amino acid orientation at residue 882 of the representative 2008 strains and, likewise, selected strains with a common amino acid residue 882. We made use of the representative 1988 strains since H882 is found in all three strains (Fig. 2A, right panel). Subsequently, we visualized the amino acid orientation at residue 88₂ and, likewise, measured the HA1 RBS diameter of each strain. As seen in Fig. 4A, the 2008 strains have varying RBS diameters while having a similar amino acid orientation at residue 882 which would further imply that only the amino acid at residue 88₂ putatively affected the HA1 RBS diameter. However, as seen in Fig. 4B, the 1988 strains have varying RBS diameter measurements and amino acid orientation at residue 882 regardless of having a common H882. This would insinuate that amino acid orientation at residue 88₂ could equally influence the HA1 RBS diameter. Amino acid residues form hydrogen bonds in proteins and protein-protein complexes which, coincidently, are dependent on amino acid orientation [32,33]. We hypothesize that varying amino acid orientations at residue 882 potentially affected the orientationdependent hydrogen bonding of residue 882 with neighboring amino acid residues which, in turn, indirectly affected the HA1 RBS diameter. Thus, we propose that both structural properties (amino acid at residue 882 and its orientation) found in the HA2 B-loop has the potential to influence HA1 RBS diameter measurements which we hypothesize may consequentially affect influenza H1N1 viral infectivity, immune escape, transmissibility, and evolution. Additional work is recommended to further prove this hypothesis.

4. Conclusion

In summary, we found that amino acid residue 88_2 is putatively strain-conserved and varies in the human (K88₂), swine (H88₂), and avian (88₂) strains of the influenza A H1N1 subtype from 1976 to 2012. In addition, we found that the HA1 RBS diameter in each strain potentially varies and, moreover, when the strain-conserved amino acid at residue 88_2 is substituted with amino acid residues found in other strains, the HA1 RBS diameter could possibly also change. Furthermore, we showed that varying amino acid orientations at residue 88_2 has the potential to influence the HA1 RBS diameter regardless of having a similar amino acid at residue 88_2 . Thus, we propose that amino acid residue 88_2 and its orientation in the HA2 B-loop is somehow structurally correlated to the HA1 RBS and, likewise, could putatively influence the HA1 RBS diameter.

Acknowledgements

This work was supported by the following grants: Grants-in-Aid for Scientific Research, Grants-in-Aid for Young Scientist (2479210), Nihon University President's Grant for Multidisciplinary Research, Uemura Fund and Dental Research Center-Nihon University School of Dentistry (Tokyo), "Strategic Research Base Development" Program for Private Universities subsidized by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) 2010 (S1001024).

References

- [1] A. Kongchanagul, O. Suptawiwat, P. Kanrai, M. Uiprasertkul, P. Puthavathana, P. Auewarakul, Positive selection at the receptor-binding site of haemagglutinin H5 in viral sequences derived from human tissues, J. Gen. Virol. 89 (2008) 1805–1810.
- [2] J.J. Skehel, D.C. Wiley, Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin, Annu. Rev. Biochem. 69 (2000) 531–569.
- [3] S.R. Das, P. Puigbo, S.E. Hensley, D.E. Hurt, J.R. Bennink, J.W. Yewdell, Glyco-sylation focuses sequence variation in the influenza A virus H1 hemagglutinin globular domain, PLoS Pathog. 6 (2010) e1001211.
- [4] J.L. Cherry, D.J. Lipman, A. Nikolskaya, Y.I. Wolf, Evolutionary dynamics of N-glycosylation sites of influenza virus hemagglutinin, PLoS Curr. 1 (2009), RRN1001.
- [5] C.C. Wang, J.R. Chen, Y.C. Tseng, C.H. Hsu, Y.F. Hung, S.W. Chen, et al., Glycans on influenza hemagglutinin affect receptor binding and immune response, Proc. Natl. Acad. Sci. U.S.A. 106 (2009) 18137–18142.
- [6] R. Daniels, B. Kurowski, A.E. Johnson, D.N. Hebert, N-linked glycans direct the cotranslational folding pathway of influenza hemagglutinin, Mol. Cell 11 (2003) 79–90
- [7] R.M. DuBois, J.M. Aguilar-Yanez, G.I. Mendoza-Ochoa, Y. Oropeza-Almazan, S. Schultz-Cherry, M.M. Alvarez, et al., The receptor-binding domain of influenza virus hemagglutinin produced in *Escherichia coli* folds into its native, immunogenic structure, J. Virol. 85 (2011) 865–872.
- [8] Y. Cao, X. Koh, L. Dong, X. Du, A. Wu, X. Ding, et al., Rapid estimation of binding activity of influenza virus hemagglutinin to human and avian receptors, PLoS ONE 6 (2011) e18664.
- [9] V. Soundararajan, K. Tharakaraman, R. Raman, S. Raguram, Z. Shriver, V. Sasisekharan, et al., Extrapolating from sequence—the 2009 H1N1 'swine' influenza virus, Nat. Biotechnol. 27 (2009) 510–513.
- [10] V. Soundararajan, S. Zheng, N. Patel, K. Warnock, R. Raman, I.A. Wilson, et al., Networks link antigenic and receptor-binding sites of influenza hemagglutinin: mechanistic insight into fitter strain propagation, Sci. Rep. 1 (2011) 200.
- [11] S.E. Hensley, S.R. Das, A.L. Bailey, L.M. Schmidt, H.D. Hickman, A. Jayaraman, et al., Hemagglutinin receptor binding avidity drives influenza A virus antigenic drift, Science 326 (2009) 734–736.
- [12] S.E. Hensley, S.R. Das, J.S. Gibbs, A.L. Bailey, L.M. Schmidt, J.R. Bennink, et al., Influenza A virus hemagglutinin antibody escape promotes neuraminidase antigenic variation and drug resistance, PLoS ONE 6 (2011) e15190.
- [13] S.J. Gamblin, L.F. Haire, R.J. Russell, D.J. Stevens, B. Xiao, Y. Ha, et al., The structure and receptor binding properties of the 1918 influenza hemagglutinin, Science 303 (2004) 1838–1842.
- [14] P.S. Rachakonda, M. Veit, T. Korte, K. Ludwig, C. Bottcher, Q. Huang, et al., The relevance of salt bridges for the stability of the influenza virus hemagglutinin, FASEB J. 21 (2007) 995–1002.
- [15] R. Xu, I.A. Wilson, Structural characterization of an early fusion intermediate of influenza virus hemagglutinin, J. Virol. 85 (2011) 5172–5182.
- [16] R.M. DuBois, H. Zaraket, M. Reddivari, R.J. Heath, S.W. White, C.J. Russell, Acid stability of the hemagglutinin protein regulates H5N1 influenza virus pathogenicity, PLoS Pathog. 7 (2011) e1002398.

- [17] S. Kumar, R. Nussinov, Close-range electrostatic interactions in proteins, Chembiochem. 3 (2002) 604–617.
- [18] B. Isin, P. Doruker, I. Bahar, Functional motions of influenza virus hemagglutinin: a structure-based analytical approach, Biophys. J. 82 (2002) 569–581.
- [19] R. Maiti, G.H. Van Domselaar, H. Zhang, D.S. Wishart, SuperPose: a simple server for sophisticated structural superposition, Nucleic Acids Res. 32 (2004) W590–W594.
- [20] I.A. Wilson, J.J. Skehel, D.C. Wiley, Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution, Nature 289 (1981) 366–373.
- [21] L.A. Kelley, M.J. Sternberg, Protein structure prediction on the Web: a case study using the Phyre server, Nat. Protoc. 4 (2009) 363–371.
- [22] L.J. McGuffin, K. Bryson, D.T. Jones, The PSIPRED protein structure prediction server, Bioinformatics 16 (2000) 404–405.
- [23] G. Pollastri, D. Przybylski, B. Rost, P. Baldi, Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles, Proteins 47 (2002) 228–235.
- [24] C. Cole, J.D. Barber, G.J. Barton, The Jpred 3 secondary structure prediction server, Nucleic Acids Res. 36 (2008) W197–W201.
- [25] R.M. Bennett-Lovsey, A.D. Herbert, M.J. Sternberg, L.A. Kelley, Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre, Proteins 70 (2008) 611–625.

- [26] P. Benkert, S.C. Tosatto, D. Schomburg, QMEAN: a comprehensive scoring function for model quality assessment, Proteins 71 (2008) 261–277.
- [27] P. Benkert, M. Kunzli, T. Schwede, QMEAN server for protein model quality estimation, Nucleic Acids Res. 37 (2009) W510–W514.
- [28] A. Herraez, Biomolecules in the computer: Jmol to the rescue, Biochem. Mol. Biol. Educ. 34 (2006) 255–261.
- [29] R.J. Russell, D.J. Stevens, L.F. Haire, S.J. Gamblin, J.J. Skehel, Avian and human receptor binding by hemagglutinins of influenza A viruses, Glycoconj. J. 23 (2006) 85–92.
- [30] M. Wiederstein, M.J. Sippl, ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, Nucleic Acids Res. 35 (2007) W407–W410.
- [31] H.M. Berman, S.K. Burley, W. Chiu, A. Sali, A. Adzhubei, P.E. Bourne, et al., Outcome of a workshop on archiving structural models of biological macromolecules, Structure 14 (2006) 1211–1217.
- [32] T. Kortemme, A.V. Morozov, D. Baker, An orientation-dependent hydrogen bonding potential improves prediction of specificity and structure for proteins and protein–protein complexes, J. Mol. Biol. 326 (2003) 1239–1259.
- [33] F. Fabiola, R. Bertram, A. Korostelev, M.S. Chapman, An improved hydrogen bond potential: impact on medium resolution protein structures, Protein Sci. 11 (2002) 1415–1423.