Free Energy Simulations Reveal a Double Mutant Avian H5N1 Virus Hemagglutinin with Altered Receptor Binding Specificity

PAYEL DAS,[1] JINGYUAN LI,[2] AJAY K. ROYYURU,[1] RUHONG ZHOU[1,2]
[1]*Computational Biology Center, IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598*
[2]*Department of Chemistry, Columbia University, New York, New York 10027*

Received 19 December 2008; Revised 11 February 2009; Accepted 19 February 2009
DOI 10.1002/jcc.21274
Published online in Wiley InterScience (www.interscience.wiley.com).

Abstract: Historically, influenza pandemics have been triggered when an avian influenza virus or a human/avian reassorted virus acquires the ability to replicate efficiently and become transmissible in the human population. Most critically, the major surface glycoprotein hemagglutinin (HA) must adapt to the usage of human-like (α-2,6-linked) sialylated glycan receptors. Therefore, identification of mutations that can switch the currently circulating H5N1 HA receptor binding specificity from avian to human might provide leads to the emergence of pandemic H5N1 viruses. To define such mutations in the H5 subtype, here we provide a computational framework that combines molecular modeling with extensive free energy simulations. Our results show that the simulated binding affinities are in good agreement with currently available experimental data. Moreover, we predict that one double mutation (V135S and A138S) in HA significantly enhances α-2,6-linked receptor recognition by the H5 subtype. Our simulations indicate that this double mutation in H5N1 HA increases the binding affinity to α-2,6-linked sialic acid receptors by 2.6 ± 0.7 kcal/mol per HA monomer that primarily arises from the electrostatic interactions. Further analyses reveal that introduction of this double mutation results in a conformational change in the receptor binding pocket of H5N1 HA. As a result, a major rearrangement occurs in the hydrogen-bonding network of HA with the human receptor, making the human receptor binding pattern of double mutant H5N1 HA surprisingly similar to that observed in human H1N1 HA. These large scale molecular simulations on single and double mutants thus provide new insights into our understanding toward human adaptation of the avian H5N1 virus.

 J Comput Chem 00: 000–000, 2009

Key words: H5N1 Avian Flu; receptor specificity; single mutation; free energy perturbation; molecular dynamics simulations; binding free energy

Introduction

The influenza pandemics of 1918, 1957, and 1968 were caused by either human adaptation or reassortment of the H1N1, H2N2, and H3N2 subtypes of avian influenza A virus.[1,2] The currently circulating avian influenza H5N1 viruses are highly pathogenic and produce high-mortality rates in humans. Therefore, the global spread of the avian H5N1 virus has raised serious public health concern that an H5N1 virus may seed the next pandemic. A typical influenza infection is initiated with the binding of the viral surface glycoprotein hemagglutinin (HA) to sialylated glycans on the host cell surface.[3–7] The linkage between sialic acids (SA) and the penultimate sugar (usually galactose) in the host cell receptor is believed to determine the host range of influenza viruses.[4–6,8,9] HA molecules of avian influenza viruses bind to α-2,3-linked receptors, whereas those of human influenza viruses prefer α-2,6-linked receptors.[4–6,8,9] A switch in the receptor specificity from α-2,3- to α-2,6-linked sialylated glycans is believed to facilitate bird-to-human as well as human-to-human transmission of influenza viruses.[4–6,8–11]

The crystal structures of H1, H3, and H5 influenza virus HAs have provided valuable molecular insights into the binding modes of human and avian influenza HAs with host cell receptors.[12–18] HAs are homotrimers with each monomer comprised of two subunits. The receptor binding domain (RBD), situated at the membrane distal end of the molecule, is formed by the 220-loop (residues 221–228), 130-loop (residues 134–138) and 190-helix (residues 188–190).* In the H1, H2, and H3 subtypes, as few as two mutations in the RBD of HA were required to switch

*The numbering of HA amino acids is based on H3 HA.

Additional Supporting Information may be found in the online version of this article.

Correspondence to: R. Zhou; e-mail: ruhongz@us.ibm.com

receptor specificity from avian to human. However, mutations that can convert receptor specificity from avian to human are different between H1 and H3 subtypes of HA, indicating the importance of characteristic structural requirements of different HA subtypes for their individual receptor specificity. For instance, mutations at residues 190 (D190E) and 225 (D225G) in 1918 H1 HA completely changed its receptor specificity from human to avian.[19] In contrast, mutations at positions 226 and 228 (Q226L and G228S) are the primary determinants of receptor specificity for human H3 HA.[20,21] However, introduction of any of these critical mutations into an avian H5 HA does not alter its receptor specificity dramatically,[17,18,22] although the double mutation (Q226L and G228S) on H5 HA does increase the binding affinity for some α-2,6-sialylated glycans (more later).[18] Thus, the mutations in H5 HA causing a shift in receptor specificity from avian to human are currently not well established. Identification of such future mutations in H5 HA will allow development of vaccines before pandemic strains emerge.

Glycan arrays comprising a diverse range of oligosaccharide motifs have been widely used to characterize the receptor binding of wild-type and mutant H1, H3, and H5 HAs.[17,18,22–24] Recently, Wilson and coworkers have shown that introduction of the Q226L, G228S, or (Q226L and G228S) mutations into an H5 HA does not effect a dramatic switch in receptor specificity.[18] However, their results indicate a significant reduction in α-2,3-glycan binding of the double mutant (Q226L and G228S). The double mutant, as well as the single mutant G228S, shows substantial binding to a natural, branched α-2,6 biantennary glycan.[18] Another glycan array study by Skehel and coworkers reports that G143R, N186K, or Q196R mutations in H5 HA may cause an increase in binding affinity to α-2,6-glycans.[17] In contrast, a different study by Nabel and coworkers that uses a resialylated hemagglutination assay in addition to a slightly modified glycan array finds no α-2,6-glycan recognition for any of the single H5 HA mutants, Q226L, G228S, N186K, or Q196R.[22] However, the same study reports S137A, T192I, and (S137A and T192I) mutants with increased human-like glycan binding.[22] These inconsistent results obtained using slightly different glycan arrays indicate the inherent difficulties and limitations of the current techniques in estimating the binding affinities quantitatively. The difficulties in determining HA-receptor binding affinity determination arise from the fact that the binding between HA and a sialylated glycan molecule is weak (typically in the millimolar range) and nonspecific.[23,25–27] Consequently, a quantitative measurement of the HA-glycan binding affinity highly depends on the experimental conditions such as the concentration of the protein.[25] Notably, currently available glycan array data provide only qualitative estimation of the receptor binding affinities of H5 HA mutants; no K_d or K_i values for the above-mentioned H5 HA mutants are reported in literature. This illustrates the complexity associated with the HA-receptor binding affinities calculation, both in experiments and in simulation. On the other hand, recent studies demonstrate that the binding specificity is regulated by the structural topology of the sialylated glycans, which, in turn, is determined by the SA-Gal linkage: Human-adapted HA favors binding with long α-2,6 glycans with a flexible, umbrella-like topology, whereas avian HA prefers α-2,3-glycans as well as short α-2,6-glycans with a rigid, cone-like topology.[25,26] Therefore, an atomic level investigation of HA-glycan binding can provide valuable insights into the effect of previously reported mutations on the relative receptor binding affinities. Such knowledge may also help to identify probable future mutations that can be critical for human receptor recognition of currently circulating H5N1 influenza viruses.

Toward this goal, we here present an unprecedented level of molecular modeling, combined with rigorous free energy perturbation (FEP) simulations,[28–34] to characterize the effect of mutations on HA-glycan binding specificity at an atomic level. The FEP method has been widely used to calculate binding affinities for a variety of biophysical phenomena, such as solvation free energies calculation, ligand-receptor binding, protein–protein interaction, and protein-DNA (RNA) binding.[28–33,35–41] Among several available computational methods developed in past years, FEP using an all-atom explicit solvent model probably provides the most accurate method for our current needs in estimating the relative protein-glycan binding affinity.[36,37] However, such simulations for realistic biological systems often require extensive computational resources. In this study, we have used IBM Bluegene/L supercomputer for these computationally expensive FEP simulations of HA/glycan complexes.[34] We find the simulated binding affinities of H5 HA mutants matching well with experimental data. Further analyses reveal critical molecular factors governing the receptor binding specificity of H5 HA. Moreover, such analyses help us to identify one double mutation that enables H5 HA to bind a human receptor analog in a similar manner to what was observed in the human H1 HA-human receptor binding. Thus, this double mutant may represent a step toward human adaptation of the avian H5 influenza virus. Taken together, our results show that such a computational approach can serve as a complementary tool to interpret and predict critical mutations for HA-receptor binding.

Results and Discussions

We model the wild-type H5 HA (structure taken from the highly pathogenic H5N1 influenza virus A/Vietnam/1203/2004, Viet04) in complex with the avian (or human) receptor analog by aligning the backbone of its RBD to that of the 1934 human H1 HA (see Fig. 1 and Model and Methods section). To elucidate the interactions between HA and the glycan receptor in this modeled wild-type complex structures (see Fig. 1), we perform short (1 ns long) molecular dynamics (MD) simulations of these wild-type HA complexes in explicit water. The HA-glycan hydrogen-bonding interactions that appear to be important in the wild-type H5 HA complexes are shown in Figure 2 and the average probabilities of those contact formation are reported in Table 1[†]: An

[†]In addition to the hydrogen-bonding interactions, the modeled H5 HA forms several nonhydrogen-bonding contacts with both avian and human receptor. Those contacts involve residues L134, S145, W153, I155, H183, P185, N186, L194, and S227 of H5 HA. A nonhydrogen-bonding interaction between an amino acid and the receptor is considered if a heavy atom of that amino acid is found within 5.0 Å of any heavy atom of the receptor.

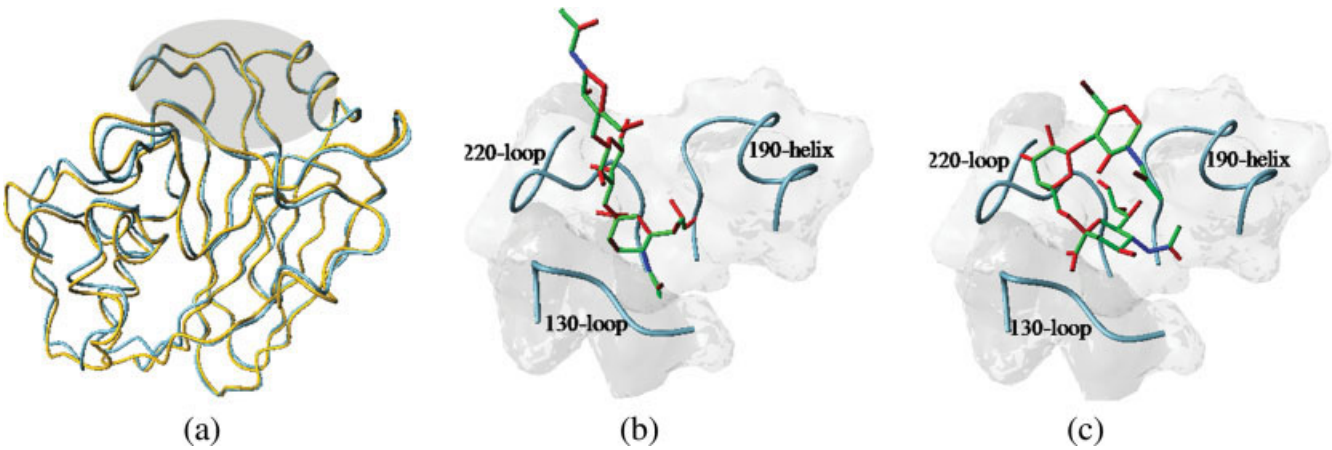

Figure 1. (a): Ribbon representation of the avian H5 HA structure (PDB ID: 2IBX, shown in cyan) with the backbone of the receptor-binding domain (RBD) aligned to that of 1934 H1 HA (PDB ID: 1RVX/1RVZ, shown in yellow). RBD, located on the top of the molecule, is highlighted with gray. RBD of H5 HA in complex with α-2,3-glycan (b) and α-2,6-glycan (c). The ribbon structure of RBD is shown as well as the surface representation. The receptor molecule is shown in green, with oxygen atoms in red and nitrogen atoms in blue. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

average of 3.5 hydrogen bonds is observed between the 130-loop (residues 135–138) and the human/avian receptor analog (see Fig. 2 and Table 1). We also find residues Y95, E190, G225, and Q226 to form hydrogen bonds with both avian and human receptor analogs, although the G225-receptor and the E190-receptor hydrogen bonds are found with lower probability with the α-2,3-glycan compared with that with α-2,6-glycan. All of these hydrogen-bonding interactions were found to be present in the cocrystal structures of avian H5 HA from A/duck/Singapore/3/97 virus (Sing97) complexed with α-2,3- and α-2,6-sialylated glycans.[42] However, the hydrogen bond between K193 and α-2,6-glycan, as well as the one between G225 and α-2,3- (or α-2,6-) glycan, was not observed in the cocrystal structures of avian H5 HA.[42] Table 1 further indicates presence of a weak hydrogen-bonding interaction between K222 of the modeled H5 HA and α-2,6-glycan, which was observed in a human H1 HA structure in complex with an α-2,6-glycan.[12] Overall, these observations clearly indicate that the HA-receptor glycan hydrogen-bonding pattern of the modeled HA is substantially similar to that of avian H5 HA with some significant differences. Those differences in the receptor binding pattern presumably result from the different conformation of the receptor binding site of the modeled H5 HA compared with that in the H5 HA cocrystal structures from the Sing97 isolate. Therefore, any mutations that are predicted to reconfigure H5 HA such that its receptor binding site more closely resembles that of a human H1 HA will be considered as a candidate mutation for the virus to gain foothold in the human population. Such mutations are expected to show a decrease in the free energy difference for HA-receptor binding between wild-type H5 and mutant H5 ($\Delta\Delta G < 0$). To verify the "effectiveness" of the predicted mutation in "humanizing" the H5 virus, we will test the same mutation in H5 HA structures isolated from both Viet04 and Sing97 virus. However, the avian H5 HA/human receptor complex structure shows only a poorly defined SA moiety, indicating the low affinity of avian H5 HA for α-2,6-glycan. Therefore, we have reconstructed the coordinates of the missing atoms in the human receptor moiety by using the receptor structure from the 1934 human H1 HA/α-2,6-glycan complex as the structural template (see the Model and Methods section).

As discussed earlier, the majority of HA-glycan hydrogen bonds involves the 130-loop that functions as an anchor of the SA moiety of the avian/human receptor, indicating importance of the residues 135–138 in receptor binding. A notable change in the 130-loop conformation between the human H1 and H3 HAs have been correlated with more intimate interactions between the human receptor and H1 HA compared with H3 HA.[12] Moreover, previous sequence and structural analyses have proposed a direct correlation between the A138 mutation to S138 in H1 HA and an enhanced specificity for α-2,6-linked receptors that is observed after 1977 to date.[43] Experimental data has attributed this enhanced human receptor specificity to the possible interaction between S138 and Q226, thereby favoring the binding of Q226 to the human receptor. Taken together, these observations underscore the critical role of the 130-loop of HA in determining receptor binding specificity and suggest that mutations in this region may allow for a switch in H5 HA receptor specificity. Here, we consider mutations in residues V135 and A138, leaving S136, a highly conserved site, as well as S137 unperturbed, such that important interactions between sialylated glycans and S136 and S137 remain intact after mutation (see Fig. 2).

Before analyzing the effect of the novel mutations at residues 135 and 138 on receptor recognition of H5 HA, we validate our FEP simulation protocol (see Model and Methods section for details) by comparing the simulated binding affinities against

experimentally available data. We employ a thermodynamic cycle (Supporting Information Figure S1) method to estimate the relative free energy change ($\Delta\Delta G$) of HA-receptor binding caused by a mutation directly from the FEP simulations. However, such a calculation of $\Delta\Delta G$s from FEP simulations is extremely nontrivial because of the low affinity between the protein HA and sialylated glycans, with K_d typically in the order of low millimolar.[23,25,27,36] Thus, the accuracy of simulated $\Delta\Delta G$s highly depends on the sampling convergence (in addition to the usual force field accuracy). To obtain an accurate estimate of the relative binding affinities from FEP calculations, we have performed extensively long‡ MD simulations of a single HA monomer with (and without) a single glycan receptor. The typical values of simulated binding free energy changes ($\Delta\Delta G$) obtained in this study are low (in the order of 0–3 kcal/mol), consistent with the millimolar dissociation constant measurements. However, any meaningful change in the binding affinity due to a mutation will be amplified because of multivalency from the presence of multiple copies of the HA trimer and the glycan receptor molecules, as found in physiological conditions.

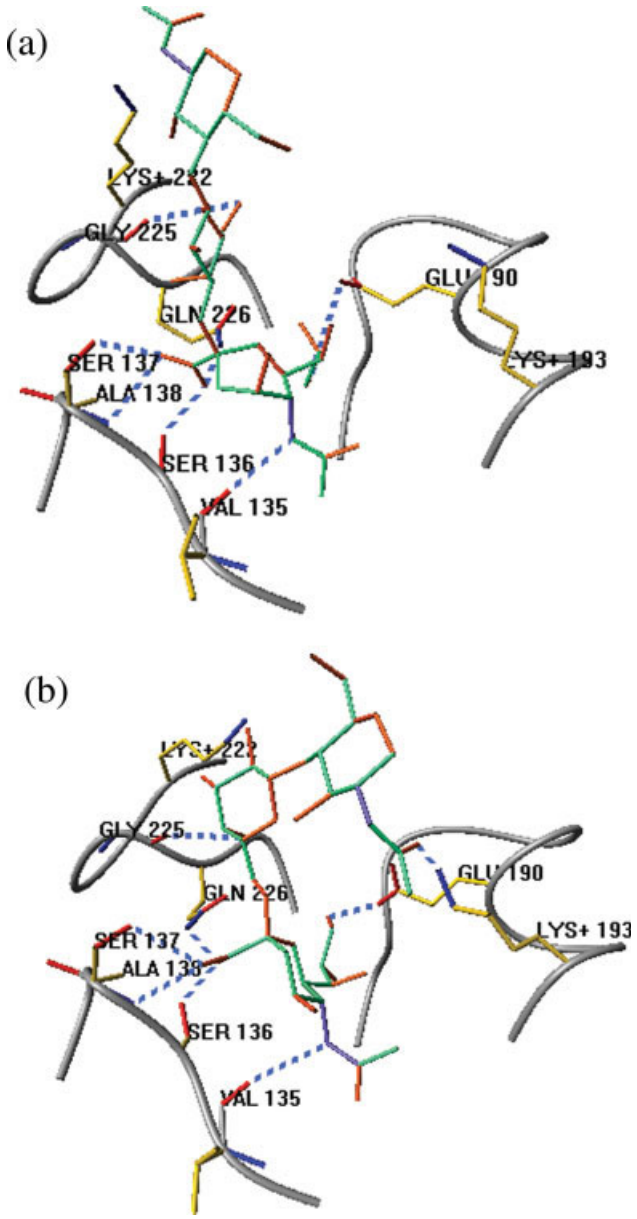

Figure 2. Hydrogen-bonding network of modeled WT avian H5 HA (backbone in gray with sidechain in yellow) with α-2,3-glycan (a) and α-2,6-glycan (b), as typically observed during MD. For clarity, residue Y95 is not shown. The sugar structure, shown in green, is adopted from the cocrystal structure of 1934 H1N1 HA. The hydrogen bonds are shown in cyan. Nitrogen atoms are shown in shades of blue with oxygen atoms in shades of red. Sidechains of few critical residues are also shown.

Table 1. Average Probability of Hydrogen-Bond Formation Between Critical Amino Acids of the Wild-Type H5 HA and Avian/Human Receptor Analogs.

Linkage	Y95	V135	S136	S137	E190	K193	K222	G225	Q226
α-2,3-	0.36	0.65	0.95	1.87	0.61	0.11	<0.1	0.25	0.55
α-2,6-	0.21	0.75	0.90	1.83	0.98	0.69	0.22	0.62	0.65

Table 2 summarizes the average and the standard deviation of the simulated binding free energy changes, $\Delta\Delta G$s, for a number of experimentally reported mutations. The first four mutants in Table 2 that show no meaningful enhancement in the simulated α-2,6-glycan binding affinity are the ones reported to switch the receptor specificity of either H1 (mutations E190D and G225D, slightly negative $\Delta\Delta G$s)[19] or H3 (mutations Q226L and G228S, both positive $\Delta\Delta G$s)[18] HA from avian to human. On the other hand, individual single mutations, G143R, Q196R, or S227N, in modeled H5 HA reduces its binding affinity to α-2,3-glycan, whereas enhancing its α-2,6-glycan binding, in good agreement with glycan array results.[17] The enhanced α-2,6-glycan recognition of the S227N mutant of H5 HA found in this study is also consistent with the presence of this mutation in human H5N1 isolates.[45] The FEP simulations for N186K,[17] however, show a reduced affinity to the human receptor analog, in line with the resialylated HA assay result.[22] The resialylated HA assay study[22] finds the two single mutants, S137A and T192I, with increased α-2,6-glycan binding, for which our simulations indicate slightly positive $\Delta\Delta G$s for α-2,6-glycan binding. Taken together, the binding affinity changes for different mutations, as obtained from FEP simulations, agree fairly well with the available glycan array data.[17,18]

For the three single mutants with enhanced simulated α-2,6-glycan binding affinity, we also analyze the individual contributions of specific interactions toward the negative $\Delta\Delta G$s for human-like glycan binding: the free energy component analysis (as discussed in the Model and Methods section) reveals that the

‡At least 66 ns of MD is performed to calculate $\Delta\Delta G$ for every mutation including five independent 6.6 ns MD runs for both the bound and the free state of the protein, which is much longer than most FEP simulations reported in refs. 38–40 and 43–46.

Table 2. Simulated Binding Affinity Changes of Avian H5 HA on Few Experimentally Reported Single Mutations.

H5 HA mutation	α-2,3-			α-2,6-		
	ΔG_b	ΔG_f	$\Delta\Delta G$	ΔG_b	ΔG_f	$\Delta\Delta G$
E190D	−17.36 (0.7)	−17.04 (0.48)	−0.32 (0.85)	−18.58 (1.04)	−18.02 (0.82)	−0.56 (1.32)
G225D	−162.27 (0.48)	−162.28 (1.27)	0.01 (1.36)	−163.48 (0.9)	−163.03 (0.73)	−0.45 (1.16)
Q226L	47.13 (0.73)	47.15 (0.87)	−0.02 (1.13)	48.45 (0.86)	45.94 (0.88)	2.51 (1.24)
G228S	11.6 (0.24)	12.11 (0.44)	−0.51 (0.5)	12.63 (0.73)	11.98 (0.88)	0.65 (1.14)
G143R	−378.15 (1.75)	−381.02 (1.35)	2.87 (2.21)	−381.5 (1.21)	−380.35 (0.97)	−1.15 (1.55)
Q196R	−327.87 (1.66)	−330.97 (1.5)	3.1 (2.24)	−330.32 (1.61)	−328.0 (1.18)	−2.32 (2.00)
S227N	−81.86 (0.46)	−83.46 (0.13)	1.60 (0.48)	−84.18 (0.24)	−82.45 (0.72)	−1.73 (0.76)
N186K	−88.87 (0.61)	−89.32 (1.43)	0.45 (1.55)	−85.24 (0.86)	−88.2 (1.83)	2.96 (2.02)
S137A	−3.07 (0.43)	−3.45 (0.25)	0.38 (0.5)	−3.45 (0.25)	−3.85 (0.15)	0.40 (0.29)
T192I	24.38 (0.57)	23.55 (0.56)	0.83 (0.8)	23.53 (0.18)	23.04 (0.69)	0.49 (0.71)

All free energy units are in kcal/mol. Standard deviations for the free energy values are reported in parentheses.

electrostatic interactions mainly contribute to the total free energy change of HA-human receptor binding due to both G143R and Q196R single mutations. In contrast, van der Waals component plays the major role in determining the free energy change of human receptor binding of S227N mutant, because N227 with its longer side chain can form a van der Waals contact with the human-like receptor more favorably than S227.

Having validated our FEP simulation protocol (in addition to the previous successful prediction of the H3N2 HA-antibody binding affinity with a similar protocol[34]), we perform a number of single/double amino acid substitutions at residues V135 and/or A138 to find mutations causing a switch in the H5 HA receptor specificity from avian to human. However, such a mutation may not involve a rigorous one which might disrupt the key interactions present between HA and sialoglycans. Therefore, we calculate the binding free energy changes for a number of mutations (both single and double) at positions 135 and/or 138 of the modeled H5 HA-receptor complexes (Table 3). Simulated binding affinities for those mutations suggest that majority of them results in either no change or a decrease in binding affinity to α-2,6-glycan over α-2,3-glycan (Table 3). However, introduction of V135S and A138S single mutations result in a small preference for α-2,6-glycan, with a $\Delta\Delta G$ of -0.6 ± 0.19 kcal/mol for V135S and -0.41 ± 0.32 kcal/mol for A138S (Table 4). Finally, the double mutation (V135S and A138S) in H5 HA significantly enhances the human receptor binding ($\Delta\Delta G = -2.56 \pm 0.73$ kcal/mol) over the avian one ($\Delta\Delta G = 0.84 \pm 1.02$ kcal/mol). We have confirmed the effect of the same double mutation in the structures of H5 HA/receptor complexes of the Sing97 virus. The double mutation (V135S and A138S) in H5 HA from the Sing97 isolate also reveals a substantial increase in the human receptor binding ($\Delta\Delta G = -1.18 \pm 0.57$ kcal/mol) over the avian one ($\Delta\Delta G = -0.15 \pm 0.99$ kcal/mol).

Although $\Delta\Delta G$ of -2.56 ± 0.73 kcal/mol (-1.18 ± 0.57 kcal/mol for Sing97 H5 HA) per monomer of H5 HA may seem small at a glance, we must consider the fact the HA-glycan binding is weak in general (there is reason for virus not to bind too strongly in order for efficient transmission), which can be

Table 3. Receptor Binding Free Energy Changes of Avian H5 HA on a Number of Mutations at V135 and A138 that does not Affect or Reduce α-2,6-Linkage Recognition.

H5 HA mutations	α-2,3-			α-2,6-		
	ΔG_b	ΔG_f	$\Delta\Delta G$	ΔG_b	ΔG_f	$\Delta\Delta G$
V135T	−18.83 (0.85)	−18.00 (0.51)	−0.83 (0.99)	−18.29 (0.21)	−18.57 (0.9)	0.28 (0.92)
V135K	−166.52 (0.21)	−168.61 (0.87)	2.09 (0.89)	−168.7 (0.3)	−170.4 (0.42)	1.57 (0.52)
V135N	−79.05 (0.61)	−78.76 (0.63)	−0.29 (0.88)	−78.8 (0.04)	−79.51 (0.96)	0.71 (0.94)
V138I	5.84 (0.21)	7.15 (0.47)	−1.31 (0.51)	5.77 (0.57)	6.37 (0.72)	−0.60 (0.92)
A138I	8.64 (0.57)	6.61 (0.56)	2.03 (0.8)	11.15 (0.86)	7.42 (0.47)	3.63 (0.98)
A138V	2.31 (0.2)	1.18 (0.36)	1.13 (0.41)	2.83 (0.49)	2.66 (0.49)	0.17 (0.69)
A138N	−77.13 (0.4)	−77.52 (0.45)	0.39 (0.6)	−76.42 (0.7)	−78.82 (0.59)	3.05 (1.35)
A138K	−166.36 (0.21)	−164.88 (1.15)	−1.68 (1.17)	−161.59 (0.31)	−161.16 (1.61)	−0.43 (1.64)
A138T	−18.75 (0.15)	−19.27 (0.37)	0.52 (0.4)	−16.65 (0.32)	−17.1 (0.5)	0.45 (0.59)
V135T + A138T	−34.44 (0.5)	−34.86 (0.75)	0.48 (0.9)	−34.9 (0.77)	−34.6 (0.64)	−0.3 (1.00)

All free energy units are in kcal/mol. Standard deviations for the free energy values are reported in parentheses.

Table 4. Receptor Binding Free Energy Changes of Avian H5 HA on V135S, A138S and (V135S and A138S) Mutations.

H5 HA mutation	α-2,3-			α-2,6-		
	ΔG_b	ΔG_f	$\Delta\Delta G$	ΔG_b	ΔG_f	$\Delta\Delta G$
V135S	5.68 (0.58)	4.56 (0.4)	1.12 (0.7)	3.87 (0.12)	4.47 (0.15)	−0.60 (0.19)
A138S	4.39 (0.1)	4.52 (0.58)	−0.13 (0.59)	4.25 (0.22)	4.66 (0.24)	−0.41 (0.32)
V135S + A138S	9.48 (0.8)	8.64 (0.63)	0.84 (1.02)	7.85 (0.42)	10.41 (0.6)	−2.56 (0.73)

The double mutation that markedly enhances α-2,6-linkage recognition. All free energy units are in kcal/mol. Standard deviations for the free energy values are reported in parentheses.

amplified if a multivalent system is taken into account. In addition, a -2.56 ± 0.73 kcal/mol of simulated binding affinity increase to the human-like glycan for a mutation in HA is rarely obtained for the many mutations performed throughout this study. Therefore, after the consideration of the multivalency in HA trimers binding, a 2–3 kcal/mol decrease in $\Delta\Delta G$ due to a mutation in HA, as obtained from FEP calculations, should be regarded as a strong indication for a substantial increase in HA-receptor binding affinity.

A free energy component analysis[34] indicates that the electrostatic interactions primarily contribute to the free energy change associated with human receptor binding of the double mutant (V135S and A138S). Typical numbers for the free energy components are as follows: of the 8.01 kcal/mol total bound state free energy change ΔG_b calculated from a particular run for the α-2,6 human-like glycan, 8.43 kcal/mol is from the electrostatic interactions, −0.37 kcal/mol from van der Waals, and the coupling term is about 0.05 kcal/mol. Similarly, of the total 10.32 kcal/mol total free state free energy change ΔG_f obtained for a particular run, 10.15 kcal/mol is from the electrostatic interactions, 0.07 kcal/mol from van der Waals, and 0.1 kcal/mol from the coupling. As a result, in the final binding affinity change $\Delta\Delta G$ for the double mutation (V135S and A138S), the electrostatic interactions also dominate the contribution, with about 80% from electrostatic and 20% from van der Waals interactions. Consistent with this result, we find a major rearrangement in the HA-glycan hydrogen-bonding network in the double mutant. First of all, α-2,6-glycan, but not α-2,3-glycan, forms favorable hydrogen bonds with Y95 and S135 (see Table 5 and Fig. 3). This results in the repositioning of the SA moiety of the human receptor, compared with that of the avian receptor, lower in the RBD of the double mutant. Table 5 further indicates that K222 in the double mutant is explicitly positioned to make a hydrogen bond with the human receptor analog with an average probability of 46% (see Fig. 3), which was much weaker in the wild-type protein complex. In addition, G225 in the (V135S and A138S) mutant interacts preferentially with the human receptor analog, but not the avian one, as seen in Table 5 and also in Figure 3. Another consequence of the double mutation is loss of the hydrogen-bonding interaction of α-2,6-glycan with K193 (see Fig. 3). Taken together, the favorable HA-glycan hydrogen-bonding interactions at residues 95, 135–137, 190, 222, 225, and 226 enable the double mutant H5 HA to bind the human receptor analog exclusively. We also observe similar HA-α-2,6-glycan hydrogen-bonding pattern in two of the experimentally reported avian H5 single mutants, G143R and S227N, with α-2,6-glycan binding specificity. However, the hydrogen bond between K222 and the human receptor analog is not found in S227N mutant. Instead, in this case N227, in contrast to S227, clearly favors interacting with the human receptor analog resulting in a higher binding affinity than the wild-type protein.

To ascertain the essential structural changes that trigger the altered hydrogen-bonding network formation between the double mutant H5 HA and α-2,6-glycan, we compare the RBD of the double mutant protein with the wild-type one. Figure 4 shows the receptor-binding domain of one representative conformation for the wild-type HA (in yellow) and for its double mutant (in cyan). Clearly, the 130-loop experiences a notable conformational change due to the double mutation. This altered conformation of the 130-loop in the double mutant enables S138 to make stronger interactions with N224 and Q226, which, in turn, affects the 220-loop conformation and makes the side-chain orientation of K222 in the double mutant different from that in the wild-type (see Fig. 4). We also notice a discrete conformational change in the 190-helix of the double mutant HA, most likely caused by the interactions with the modified 220-loop (see Fig. 4).[§] Taken together, the altered RBD conformation in double mutant H5 HA, with ∼0.8 Å of backbone RMSD from wild-type RBD, facilitates binding with α-2,6-glycan by providing optimal contacts with Y95, S135, S136, S137, E190, K222, G225, and Q226. Many of these positions, such as residues 135–

Table 5. Average Probability of Hydrogen-Bond Formation Between Critical Amino Acids of the (V135S and A138S) mutant H5 HA and Avian/Human Receptor Analogs.

Linkage	Y95	S135	S136	S137	E190	K193	K222	G225	Q226
α-2,3-	0.18	0.52	0.48	1.87	0.19	0.23	<0.1	<0.1	0.69
α-2,6-	0.56	0.85	0.91	1.86	0.92	<0.1	0.46	0.81	0.59

[§]Although the major rotamer states of K193 and K222 in the double mutant protein are the ones shown in Figure 4, the wild-type rotamer states are also populated in the double mutant protein to a much lesser extent.

137, 190, 222, and 225, were found to be in contact with the human receptor in human H1 HAs.[12] This observation indicates that this double mutant of avian H5 HA binds to the human receptor analog in a similar fashion to that of a human H1 HA. In addition, these findings are consistent with the human receptor specificity of H1 HAs with a serine at position 138 that was observed in human H1 HA after 1977 to date. Thus, this double mutation may represent a path toward human adaptation of avian H5 HA. Moreover, the results presented here underline the importance of such a large-scale simulation study, in addition to experiments, for early detection of viruses able to cause a pandemic flu.

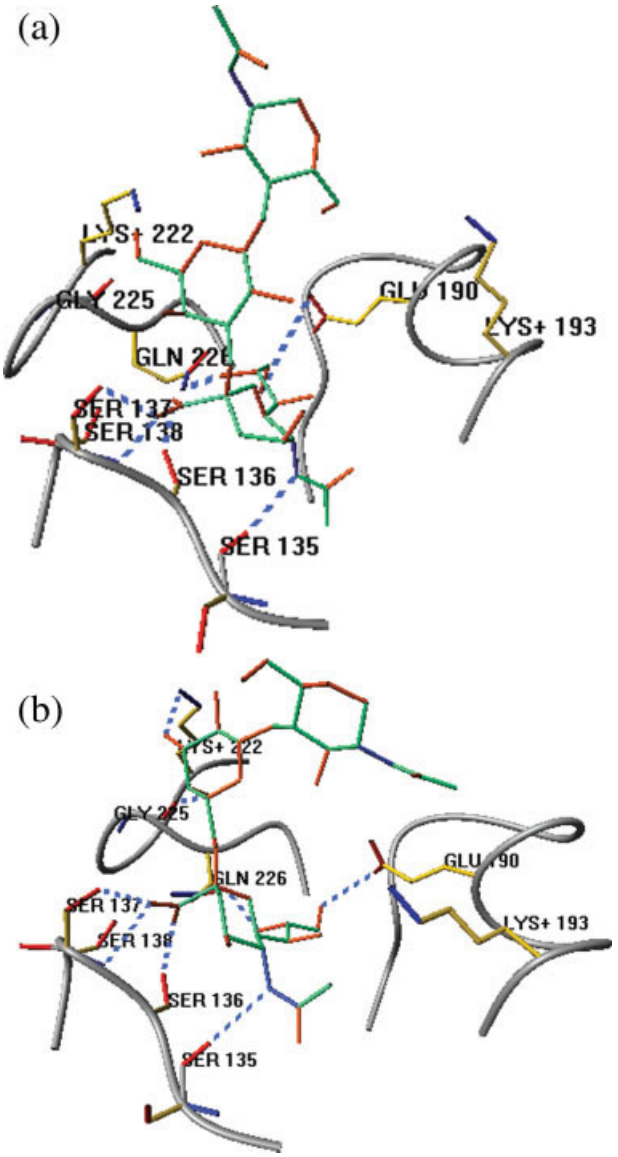

Figure 3. Hydrogen-bonding network of the modeled (V135S and A138S) mutant avian H5 HA with α-2,3-glycan (a) and α-2,6-glycan (b), as typically observed during 0.3 ns MD. For clarity, residue Y95 is not shown. The colors used are same as in Figure 2. On double mutation, the hydrogen bond between G225 and galactose sugar of the avian receptor is completely lost. On the other hand, K222 in the double mutant HA forms a hydrogen bond with much higher probability with the galactose sugar of the human receptor analog, whereas the hydrogen bond between K193 and SA of the human receptor analog disappears.

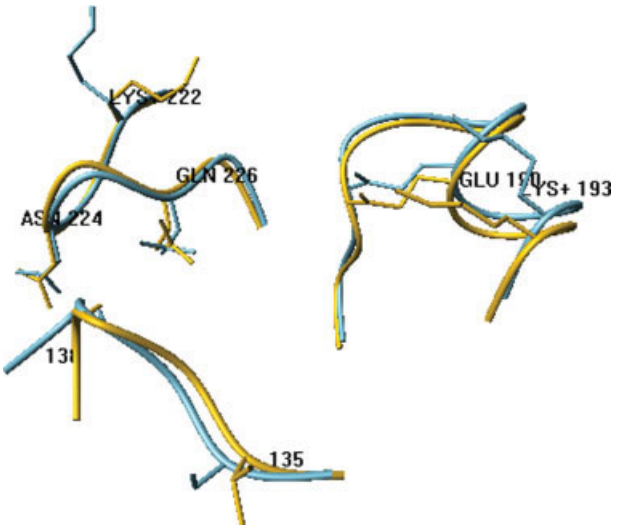

Figure 4. The conformational change observed in the free state of modeled avian H5 HA on the (V135S and A138S) mutation. The backbone of the WT protein is shown in yellow, whereas the backbone of the double mutant protein is shown in cyan. Sidechains are shown for the two mutation sites, 135 and 138, in addition to the residues E190, K193, K222, N224, and Q226. Clearly, introduction of the double mutation alters the conformation of the receptor binding pocket of H5 HA, which significantly facilitates human receptor binding.

Conclusions

In conclusion, this study presents a computational approach that uses rigorous FEP simulations to quantitatively estimate binding affinities of H5 HA to sialylated glycan receptors. The simulated binding affinities agree fairly well with the currently available glycan array data for a few H5 HA mutants. In addition, we identify a double mutation (V135S and A138S) in H5 HA which dramatically enhances its human receptor specificity. The simulated binding affinity change of this double mutant H5 HA for the human receptor is substantially high ($\Delta\Delta G = -2.6 \pm 0.7$ kcal/mol per HA monomer). Such an enhancement is primarily due to the electrostatic interactions between double mutant HA and the human receptor analog, as revealed from a free energy component analysis. We further notice a major rearrangement in the HA-glycan hydrogen-bonding network as a result of the double mutation, which enables the double mutant H5 HA to preferentially bind with the human receptor analog. A detailed analysis of the simulated configurational ensemble indicates distinct conformational change in the binding pocket of (V135S and A138S) mutant H5 HA from that of the wild-type one. Residues

Y95, S135, S136, S137, E190, K222, G225, and Q226 of this altered RBD form favorable hydrogen-bonding interactions with the human-like receptor—similar to what was seen for previous human-adapted HAs. Thus, introduction of the (V135S and A138S) double mutation in avian H5 HA may provide a path to gain foothold in human population. The crucial presence of a serine at position 138 in human H1 HA further emphasizes the importance of this predicted double mutation in avian H5 HA in gaining the pandemic potential. Moreover, the findings from this study provide valuable insight into the atomic-level mechanism of HA-glycan binding by detecting few amino acids that are primary determinants for α-2,3- or α-2,6-glycan specificity. Such knowledge can aid in prompt identification of emerging viruses with a potential of human adaptation.

Method and System

The complete cocrystal structure of H5 HA with the human SA receptor is not available to date. However, the crystal structure of H5 HA from a human isolate reveals high-structural similarity to 1934 human H1 HA.[17,18] Therefore, to obtain the initial complex structures of avian H5 HA of the Viet04 isolate with avian/human SA receptor we exploit the availability of cocrystal structures of H1N1 HA with both α-2,3-glycan (PDB ID: 1RVX) and α-2,6-glycan (PDB ID: 1RVZ). The backbone of the residues corresponding to RBD (residue 135–138, 190–193, 222–228) of Viet04 HA (PDB ID: 2IBX) are aligned to that of the H1 HA configuration complexed with either α-2,3- or α-2,6-glycan (see Fig. 1). In this study, we only consider residues 56–263 of a single monomer of avian H5 HA in complex with a trisaccharide as the receptor moiety, as shown in Figure 1. The structural coordinates for the avian H5 HA complexes of the Sing97 virus are taken from residues 52–258 of the PDB structures 1JSN and 1JSO.[42] However, Sing97 HA/α-2,6-glycan complex structure (PDB ID: 1JSO) shows only a poorly defined SA moiety. The remaining two groups of the trisaccharide sugar moiety are not visible in X-ray experiments, indicating low affinity of H5 HA for the human receptor. Therefore, we have reconstructed the remaining two groups of the trisaccharide human receptor moiety by using the receptor structure from the human H1 HA/α-2,6-glycan complex (PDB ID: 1RVZ) as the structural template. For this purpose, the heavy atoms of the SA moiety of the H5 HA/human receptor structure are aligned to those of the H1 HA/human receptor structure and the penultimate sugars are then positioned accordingly.

The resultant H5 HA complex structures are solvated in a 67 Å × 67Å × 77Å water box with a total of ~9260 water molecules. The system of a total of ~32,000 atoms is then subjected to a 10,000 steps of energy minimization followed by a 1 ns equilibration, during which positions of the backbone of the RBD residues, in addition to those heavy atoms of the receptor that are within 8 Å of RBD residues, remain constrained. The configuration at the end of this long MD simulation is used as the starting point for the FEP calculations. The particle-mesh Ewald method is used for the long-range electrostatic interactions with a cutoff distance of 12 Å. All MD simulations are performed using NAMD2[47,48] molecular modeling package with 1.5 fs time step in NPT ensemble at 1 atm and 300 K. The CHARMM22 force field[49] and TIP3P water model[50] are used. All the parameters for the sialylated glycan receptors are obtained from the CHARMM22 force-field, except for the charges which are not available to our best knowledge. Instead, these charges are calculated from quantum mechanics ESP fitting in solvent at the HF-631G* basis level. These quantum charges might not be optimal, but we believe the difference in receptor binding affinity is largely due to the different orientation of hydroxyl groups and/or the different topology in different linkages (α-2,3 vs. α-2,6-linkage).[¶] We have also verified that the structures of the glycan receptors with the above-mentioned set of parameters remain stable during MD simulation, both in gas phase and in solution.

We have used a FEP approach[34,38–40] to estimate the binding affinity changes between HA and glycans on mutation. Although FEP simulations require massive computational resources to perform, they provide the most accurate estimation of free energy changes by considering the effect of the conformational flexibility of the complex system. The long timescale involved makes it nearly impossible to simulate the complete protein-ligand binding processes and calculate the direct binding free energies with the currently available computational resources. However, the *relative* binding free energy change ($\Delta\Delta G_{\text{binding}} = \Delta G_{\text{wt}} - \Delta G_{\text{mut}}$) due to a mutation can be calculated by using a thermodynamic cycle (see Supporting Information Figure S1). In this thermodynamic cycle, the binding free energy change due to a mutation, $\Delta\Delta G_{\text{binding}}$, is estimated from the difference in the free energy changes caused by the same mutation for the bound state (ΔG_b) and for the free state (ΔG_f) of the protein, as shown in Supporting Information Figure S1. The free energy change involved in such a transformation in either the bound state or free state (ΔG), in which one or more amino acids of the wild-type protein gradually mutates over the course of a simulation, is calculated using the FEP formula[51]:

$$\Delta G_\lambda = -RT \ln\langle \exp(-[V_{\lambda+\Delta\lambda} - V_\lambda]/RT)\rangle\lambda,$$

$$\Delta G = \sum_{\lambda=0}^{\lambda=1} \Delta G\lambda,$$

where $V(\lambda) = (1 - \lambda)V_1 + \lambda V_2$, V_1 and V_2 representing the potential energies of the wild-type and the mutant, respectively. The FEP parameter λ changes from 0 (V_1) to 1 (V_2) when the system mutates from the wild-type to the mutant, and $<\ldots>_\lambda$ represents the ensemble average at potential $V(\lambda)$.

In this study, a mutation from residue A (state A) to residue B (state B) is performed over 22 FEP windows with more windows near the two endpoints to obtain better convergence of ΔGs. The simulation time for each window is 0.3 ns, resulting in a 6.6 ns long simulation per run, and at least five independent runs starting from different initial configurations are run for each state, resulting in 33 ns run for both the bound and free

[¶]All force-field parameters are available on request.

states (total 66 ns per mutation), much longer than the length of similar simulations currently reported in literature.[37–40] A highly parallel and specially optimized version (for BlueGene/L) of NAMD2[48] is used to perform these FEP simulations. For the mutation of a neutral residue to a charged one, the system charge is counter-balanced by mutating another neutral residue far from the binding site to an oppositely charged one.[34] The positions of the residues that lie at the monomer–monomer interface of HA trimer remain constrained during the FEP simulations to account for the presence of other monomers. For every single mutation, at least five independent runs are performed to obtain the averages and the standard deviations of the binding free energy changes.

Despite the controversy in the literature about the meaningfulness of breaking the total free energy into components[44,52–54] and the ambiguity associated with a path-dependent decomposition, a break-up of the total binding free energy change into its van der Waals and electrostatic components can provide useful information about the energetic interactions involved in HA-glycan binding. In this study, we decompose the free energy change associated, ΔG_λ, into:

$$\Delta G_\lambda = \Delta G_\lambda^{\mathrm{vdw}} + \Delta G_\lambda^{\mathrm{elec}} + \delta G_{\mathrm{coup}},$$

where δG_{coup} is the coupling free energy that arises from the nonadditivity of the free energy components to ΔG. Next, ΔG_{vdw} and ΔG_{elec} (as well as ΔG) are estimated simultaneously using the same conformational ensemble $<\ldots>_\lambda$ at potential $V(\lambda)$, by collecting the van der Waals and electrostatic interaction contributions separately, i.e. $V(\lambda) = V(\lambda)_{\mathrm{vdw}} + V(\lambda)_{\mathrm{elec}}$.

The conformational ensemble generated at the end (with $\lambda = 1$) of a FEP simulation is considered for further structural characterization. For hydrogen-bond analysis, a distance cutoff of 3.5 Å and an angle cutoff of 150° are used. The representative conformation for a particular ensemble is selected using a cluster analysis algorithm[55] as following: we pick the neighbors for each conformation such that the pairwise $\mathrm{RMSD}_{C\alpha}$ of the RBD is <0.5 Å. Next, the structure with the largest number of neighbors is considered as the representative conformation of the cluster that contains all its neighbors. This cluster is then eliminated from the pool of structures and the process is repeated until no structures are left in the pool.

Acknowledgments

We thank Ian Wilson and Damian Ekiert for critical reading of the manuscript. We also thank Peter Palese, Jim Paulson, Dennis Burton, Bruce Berne, David Silverman, Alice McHardy, and Isidore Rigoutsos for their helpful discussions. We thank Sameer Kumar for numerous help with porting NAMD2, particularly the free energy perturbation (FEP) module, onto IBM BlueGene/L. We would also like to acknowledge the contributions of the BlueGene/L hardware, system software, and science application teams whose efforts and assistance made it possible for us to use the BlueGene/L supercomputer at the IBM Watson Center.

References

1. Johnson, N.; Mueller, J. Bull Hist Med 2002, 76, 105.
2. Scholtissek, C.; Rohde, W.; Vonhoyningen, V.; Rott, R. Virology 1978, 87, 13.
3. Kuiken, T.; Holmes, E. C.; McCauley, J.; Rimmelzwaan, G. F.; Williams, C. S.; Grenfell, B. T. Science 2006, 312, 394.
4. Russell, R. J.; Stevens, D. J.; Haire, L. F.; Gamblin, S. J.; Skehel, J. J. Glycoconjugate J 2006, 23, 85.
5. Shinya, K.; Ebina, M.; Yamada, S.; Ono, M.; Kasai, N.; Kawaoka, Y. Nature 2006, 440, 435.
6. Skehel, J. J.; Wiley, D. C. Annu Rev Biochem 2000, 69, 531.
7. van Riel, D.; Munster, V. J.; de Wit, E.; Rimmelzwaan, G. F.; Fouchier, R. A. M.; Osterhaus, A.; Kuiken, T. Science 2006, 312, 399.
8. van Riel, D.; Munster, V. J.; de Wit, E.; Rimmelzwaan, G. F.; Fouchier, R. A. M.; Osterhaus, A.; Kuiken, T. Am J Pathol 2007, 171, 1215.
9. Nicholls, J. M.; Bourne, A. J.; Chen, H.; Guan, Y.; Peiris, J. S. M. Respir Res 2007, 8, 73.
10. Parrish, C. R.; Kawaoka, Y. Annu Rev Microbiol 2005, 59, 553.
11. Suzuki, Y.; Ito, T.; Suzuki, T.; Holland, R. E.; Chambers, T. M.; Kiso, M.; Ishida, H.; Kawaoka, Y. J Virol 2000, 74, 11825.
12. Gamblin, S. J.; Haire, L. F.; Russell, R. J.; Stevens, D. J.; Xiao, B.; Ha, Y.; Vasisht, N.; Steinhauer, D. A.; Daniels, R. S.; Elliot, A.; Wiley, D. C.; Skehel, J. J. Science 2004, 303, 1838.
13. Stevens, J.; Corper, A. L.; Basler, C. F.; Taubenberger, J. K.; Palese, P.; Wilson, I. A. Science 2004, 303, 1866.
14. Nobusawa, E.; Ishihara, H.; Morishita, T.; Sato, K.; Nakajima, K. Virology 2000, 278, 587.
15. Matrosovich, M.; Tuzikov, A.; Bovin, N.; Gambaryan, A.; Klimov, A.; Castrucci, M. R.; Donatelli, I.; Kawaoka, Y. J Virol 2000, 74, 8502.
16. Glaser, L.; Stevens, J.; Zamarin, D.; Wilson, I. A.; Garcia-Sastre, A.; Tumpey, T. M.; Basler, C. F.; Taubenberger, J. K.; Palese, P. J Virol 2005, 79, 11533.
17. Yamada, S.; Suzuki, Y.; Suzuki, T.; Le, M. Q.; Nidom, C. A.; Sakai-Tagawa, Y.; Muramoto, Y.; Ito, M.; Kiso, M.; Horimoto, T.; Shinya, K.; Sawada, T.; Kiso, M.; Usui, T.; Murata, T.; Lin, Y.; Hay, A.; Haire, L. F.; Stevens, D. J.; Russell, R. J.; Gamblin, S. J.; Skehel, J. J.; Kawaoka, Y. Nature 2006, 444, 378.
18. Stevens, J.; Blixt, O.; Tumpey, T. M.; Taubenberger, J. K.; Paulson, J. C.; Wilson, I. A. Science 2006, 312, 404.
19. Tumpey, T. M.; Maines, T. R.; Van Hoeven, N.; Glaser, L.; Solorzano, A.; Pappas, C.; Cox, N. J.; Swayne, D. E.; Palese, P.; Katz, J. M.; Garcia-Sastre, A. Science 2007, 315, 655.
20. Connor, R. J.; Kawaoka, Y.; Webster, R. G.; Paulson, J. C. Virology 1994, 205, 17.
21. Rogers, G. N.; Paulson, J. C.; Daniels, R. S.; Skehel, J. J.; Wilson, I. A.; Wiley, D. C. Nature 1983, 304, 76.
22. Yang, Z. Y.; Wei, C. J.; Kong, W. P.; Wu, L.; Xu, L.; Smith, D. F.; Nabel, G. J. Science 2007, 317, 825.
23. Blixt, O.; Head, S.; Mondala, T.; Scanlan, C.; Huflejt, M. E.; Alvarez, R.; Bryan, M. C.; Fazio, F.; Calarese, D.; Stevens, J.; Razi, N.; Stevens, D. J.; Skehel, J. J.; van Die, I.; Burton, D. R.; Wilson, I. A.; Cummings, R.; Bovin, N.; Wong, C. H.; Paulson, J. C. Proc Natl Acad Sci USA 2004, 101, 17033.
24. Stevens, J.; Blixt, O.; Glaser, L.; Taubenberger, J. K.; Palese, P.; Paulson, J. C.; Wilson, I. A. J Mol Biol 2006, 355, 1143.
25. Srinivasan, A.; Viswanathan, K.; Raman, R.; Chandrasekaran, A.; Raguram, S.; Tumpey, T. M.; Sasisekharan, V.; Sasisekharan, R. Proc Natl Acad Sci USA 2008, 105, 2800.

26. Chandrasekaran, A.; Srinivasan, A.; Raman, R.; Viswanathan, K.; Raguram, S.; Tumpey, T. M.; Sasisekharan, V.; Sasisekharan, R. Nat Biotechnol 2008, 26, 107.
27. Blixt, O.; Collins, B. E.; van den Nieuwenhof, I. M.; Crocker, P. R.; Paulson, J. C. J Biol Chem 2003, 278, 31007.
28. Simonson, T.; Archontis, G.; Karplus, M. Acc Chem Res 2002, 35, 430.
29. Tembe, B. L.; McCammon, J. A. Comput Chem 1984, 8, 281.
30. Jorgensen, W. L. Acc Chem Res 1989, 22, 184.
31. Warshel, A.; Sussman, F.; King, G. Biochemistry 1986, 25, 8368.
32. Kollman, P. Chem Rev 1993, 93, 2395.
33. Rami Reddy, M.; Erion, M. D.; Agarwal, A. Rev Comput Chem 2000, 16, 217.
34. Zhou, R.; Das, P.; Royyuru, A. K. J Phys Chem B 2008, 112, 15813.
35. Wang, W.; Donini, O.; Reyes, C. M.; Kollman, P. A. Annu Rev Biophys Biomol Struct 2001, 30, 211.
36. Pathiaseril, A.; Woods, R. J. J Am Chem Soc 2000, 122, 331.
37. Laederach, A.; Reilly, P. J. Proteins: Struct Funct Bioinfo 2005, 60, 591.
38. Tirado-Rives, J.; Jorgensen, W. L. J Med Chem 2006, 49, 5880.
39. Deng, Y.; Roux, B. J Chem Theo Comp 2006, 2, 1255.
40. Jayachandran, G.; Shirts, M. R.; Park, S.; Pande, V. S. J Chem Phys 2006, 125, 084901.
41. Almlof, M.; Aqvist, J.; Smalas, A. O.; Brandsdal, B. O. Biophys J 2006, 90, 433.
42. Ha, Y.; Stevens, D. J.; Skehel, J. J.; Wiley, D. C. Proc Natl Acad Sci USA 2001, 98, 11181.
43. Rogers, G. N.; Dsouza, B. L. Virology 1989, 173, 317.
44. Mark, A. E.; Vangunsteren, W. F. J Mol Biol 1994, 240, 167.
45. Gambaryan, A.; Tuzikov, A.; Pazynina, G.; Bovin, N.; Balish, A.; Klimov, A. Virology 2006, 344, 432.
46. Thorpe, I. F.; Brooks, C. L., III. Proc Natl Acad Sci USA 2007, 104, 8821.
47. Zhou, R.; Eleftheriou, M.; Royyuru, A. K.; Berne, B. J. Proc Natl Acad Sci USA 2007, 104, 5824.
48. Kumar, S.; Huang, C.; Zheng, G.; Bohm, E.; Bhatele, A.; Phillips, J. C.; Yu, H.; Kale, L. V. IBM J Res Dev 2008, 52, 177.
49. MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph, D.; Kuchnir, L.; Kuczera, K.; Lau, F.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. J Phys Chem B 1998, 102, 3586.
50. William, L. J.; Jayaraman, C.; Jeffry, D. M.; Roger, W. I.; Michael, L. K. J Chem Phys 1983, 79, 926.
51. Zwanzig, R. W. J Chem Phys 1954, 22, 1420.
52. Boresch, S.; Karplus, M. J Mol Biol 1995, 254, 801.
53. Brady, G. P.; Sharp, K. A. J Mol Biol 1995, 254, 77.
54. Bren, M.; Florian, J.; Mavri, J.; Bren, U. Theor Chem Acc 2007, 117, 535.
55. Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. Angew Chem Int Ed 1999, 38, 236.