Free-Energy Simulations Reveal that Both Hydrophobic and Polar Interactions Are Important for Influenza Hemagglutinin Antibody Binding

Zhen Xia,[†§] Tien Huynh,[†] Seung-gu Kang,[†] and Ruhong Zhou[†‡*]
[†]Computational Biology Center, IBM Thomas J. Watson Research Center, Yorktown Heights, New York; [‡]Department of Chemistry, Columbia University, New York, New York; and [§]Department of Biomedical Engineering, The University of Texas at Austin, Austin, Texas

ABSTRACT Antibodies binding to conserved epitopes can provide a broad range of neutralization to existing influenza subtypes and may also prevent the propagation of potential pandemic viruses by fighting against emerging strands. Here we propose a computational framework to study structural binding patterns and detailed molecular mechanisms of viral surface glycoprotein hemagglutinin (HA) binding with a broad spectrum of neutralizing monoclonal antibody fragments (Fab). We used rigorous free-energy perturbation (FEP) methods to calculate the antigen-antibody binding affinities, with an aggregate underlying molecular-dynamics simulation time of several microseconds (~2 μs) using all-atom, explicit-solvent models. We achieved a high accuracy in the validation of our FEP protocol against a series of known binding affinities for this complex system, with $<$0.5 kcal/mol errors on average. We then introduced what to our knowledge are novel mutations into the interfacial region to further study the binding mechanism. We found that the stacking interaction between Trp-21 in HA2 and Phe-55 in the CDR-H2 of Fab is crucial to the antibody-antigen association. A single mutation of either W21A or F55A can cause a binding affinity decrease of $\Delta\Delta G > 4.0$ kcal/mol (equivalent to an ~1000-fold increase in the dissociation constant K_d). Moreover, for group 1 HA subtypes (which include both the H1N1 swine flu and the H5N1 bird flu), the relative binding affinities change only slightly ($< \pm 1$ kcal/mol) when nonpolar residues at the αA helix of HA mutate to conservative amino acids of similar size, which explains the broad neutralization capability of antibodies such as F10 and CR6261. Finally, we found that the hydrogen-bonding network between His-38 (in HA1) and Ser-30/Gln-64 (in Fab) is important for preserving the strong binding of Fab against group 1 HAs, whereas the lack of such hydrogen bonds with Asn-38 in most group 2 HAs may be responsible for the escape of antibody neutralization. These large-scale simulations may provide new insight into the antigen-antibody binding mechanism at the atomic level, which could be essential for designing more-effective vaccines for influenza.

INTRODUCTION

Influenza A virus is one of the most fatal infectious diseases in humans and poultry (1,2). Each year, this virus causes hundreds of thousands of deaths in humans and tens of millions of deaths in birds worldwide (3–9). The recent H1N1 swine flu pandemic and the spread of highly pathogenic H5N1 avian flu caused a great public-health concern (10–18). Current vaccines usually respond to a limited number of strains and often fail to neutralize emerging new strains because of rapid genetic evolution (19–24). Although some neuraminidase (NA) inhibitors, such as oseltamivir (Tamiflu) and zanamivir (Relenza), have shown effective suppression against influenza viruses, their efficacy is often challenged by increasing drug resistance (25–32).

Therefore, broad and potent cross-protective host immunity is in great demand for influenza prevention and treatment. High-affinity neutralizing antibodies (nAbs), such as F10 and CR6261, have been selected by phage display on recombinant H5 HA (13–15). Both F10 and CR6261 have shown broad neutralization to all group 1 subtypes, including the H1N1 Spanish flu and the H5N1 bird flu. In general, 16 hemagglutinin (HA) subtypes of influenza A viruses are categorized into two major phylogenetic groups: group 1 (H1, H2, H5, H6, H8, H9, H11, H12, H13, and H16) and group 2 (H3, H4, H7, H10, H14, and H15). The cocrystal structures of H5 HA with antibodies reveal that both antibodies can block the viral infection by inserting their heavy chain into a conserved helical stem region in HA1 and HA2, thereby possibly preventing membrane fusion (13–15). We believe that the similar neutralization effect of these nAbs could result from some important common features that are critical for developing new broad-spectrum vaccines. However, several underlying key mechanisms remain unclear, and it remains to be determined at which molecular level antibodies have a broad range of neutralization, and why most group 2 influenza viruses (e.g., H3 and H7) cannot be neutralized by F10 or CR6261. To address these issues, we performed large-scale, free-energy perturbation (FEP) simulations to characterize key residues and their mutation effects on HA-Fab binding at the atomic level. The FEP method has been widely used to calculate binding affinities for a variety of chemical and biological systems, such as solvation free-energy calculation, ligand-receptor binding, protein–protein interaction, and protein-DNA (RNA) binding (17,18,33–44), and is often regarded as the most rigorous and reliable method for free-energy calculations. Many previous FEP calculations achieved high accuracy for various protein-protein and protein-ligand binding affinities when compared with experiments (37,45–47). Our

Submitted September 29, 2011, and accepted for publication January 27, 2012.

*Correspondence: ruhongz@us.ibm.com

Editor: Alexandre Bonvin.

0006-3495/12/03/1453/9 $2.00

doi: 10.1016/j.bpj.2012.01.043

previous work on H3N2 HA-antibody and H5N1 HA-receptor binding systems also showed excellent agreement between the calculated binding free energies and experimental values (17,18). Among several available computational methods developed in past years, the FEP method based on the all-atom explicit-solvent model is probably the most accurate approach for estimating the relative antigen-antibody binding affinity (45,46). In this study, we first validated our FEP protocol by comparing the simulated binding affinity changes with available experimental data. We then extended our FEP calculation to novel (to our knowledge) mutations at the interfacial region on either group 1 HA or monoclonal antibody fragment (Fab). We found that the stacking interaction is critical for HA-antibody recognition, and the nonspecific hydrophobic interaction is responsible for the broad neutralization of antibody. To understand why F10-like antibodies successfully neutralize most group 1 influenza subtypes but fail in group 2 subtypes, we further investigated the role of a highly conserved His residue observed in almost all important group 1 HA subtypes (e.g., H1 and H5) by mutating it to Asn in four of the six group 2 HA subtypes (i.e., H3, H7, H10, and H15). Our results show that such a computational approach can serve as a complementary tool for interpreting and predicting critical mutations for HA-antibody binding.

METHOD AND SYSTEM

Molecular systems

The H5 HA (Viet04/H5) with Fab F10 complex (PDB entry 3FKU) was used for antigen-antibody binding, where the HA (HA1 and HA2) monomer was bound to one Fab with both heavy and light chains (13). The HA-Fab complex was solvated in a 71.5 Å × 81.5 Å × 160.0 Å water box, and then seven sodium ions were added to neutralize the system, with a total number of ~88,000 atoms. The solvated system was first energy-minimized by 20,000 steps, followed by a 500,000-step equilibration. The snapshot during the equilibration was randomly picked as the starting point for the FEP calculations. The unbound (free) state was modeled with the HA or F10 only, solvated in water, and equilibrated with a similar process. The particle-mesh Ewald method was used for the long-range electrostatic interactions with a cutoff distance of 12 Å (48). All underlying molecular-dynamics simulations, which are widely used to simulate biological systems (49–58), were performed using the NAMD2 (59) molecular modeling package with a 1.5 fs time step in the NPT ensemble at 1 atm and 300 K. The CHARMM22 force field (60) and TIP3P water model (61) were used.

FEP protocol

When an antigenic variation occurs, the change in the HA binding affinity to a neutralizing antibody can be calculated by the FEP method (17,18,33–35,37–40). The Helmholtz free energy of a system can be expressed as

$$G = -kT \ln Z = kT \ln\left\{\iint dpdq \exp[-\beta H(p,q)]\right\}, \quad (1)$$

where Z is the partition function; $H(p,q)$ is the Hamiltonian of the system; p and q represent the momentum and the coordinate, respectively; k is the Boltzmann constant; T is the temperature; and β is equal to $1/kT$. The binding free-energy change ΔG due to a mutation in HA or antibody can then be calculated as

$$\Delta G_\lambda = -kT \ln\langle \exp(-\beta[V(\lambda+\Delta\lambda) - V(\lambda)])\rangle_\lambda \quad (2)$$

$$\Delta G = \sum_\lambda \Delta G_\lambda, \quad (3)$$

where $V(\lambda) = (1-\lambda)\,V_1 + \lambda\,V_2$, with V_1 representing the potential energy of the wild-type (WT) and V_2 representing the potential energy of the mutant. The FEP parameter λ changes from 0 (V_1) to 1 (V_2) when the system changes from the WT to the mutant, and $<\ldots>_\lambda$ represents the ensemble average at potential $V(\lambda)$. In typical FEP calculations, to ensure a smooth transition from state A to state B, many perturbation windows have to be used. To avoid singularity at small interaction distances, when λ approaches 0 or 1 (a situation often referred to as endpoint-catastrophe), we used soft-core potentials for the Lennard-Jones interactions, with the 12-6 LJ function modified as follows (62,63):

$$V_{vdW} = \varepsilon_{ij}\left[\left(\frac{R_{ij}^2}{r_{ij}^2 + \delta(1-\lambda)}\right)^6 - \left(\frac{R_{ij}^2}{r_{ij}^2 + \delta(1-\lambda)}\right)^3\right], \quad (4)$$

where ε_{ij} is the depth of the potential well, R_{ij} is the radius, r_{ij} is the distance between a pair of atoms, and δ is the shift parameter that allows a smooth transition from the original Lennard-Jones potential to zero or vice versa. The electrostatic interactions are handled with the normal Coulomb law but are switched on for the appearing atoms only after $\lambda > 0.1$, thus allowing the soft-core Lennard-Jones potentials to repel the possible overlapping before introducing the electrostatic interactions. Similarly, for the disappearing atoms, the electrostatic interactions are switched off after $\lambda > (1-0.1) = 0.9$.

In general, it is difficult to directly calculate the absolute binding affinity change ΔG_A for the binding process between a viral surface protein and an antibody due to the long timescale and complicated binding process. However, we can avoid this problem by designing a thermodynamic cycle to calculate the relative binding free-energy change, i.e., $\Delta\Delta G_{AB}$. Instead of calculating the difficult direct binding energies ΔG_A and ΔG_B, we calculate the free-energy changes for the same mutation in both the bound state (HA bound to antibody, ΔG_1) and the free state (HA or antibody not bounded, ΔG_2). Within a complete thermodynamic cycle, the total free-energy change should be zero, which gives the relative binding affinity due to the mutation from A → B as:

$$\Delta\Delta G_{bind} = \Delta G_B - \Delta G_A = \Delta G_1 - \Delta G_2. \quad (5)$$

In the current setup, we adopted a 20-window scheme with a soft-core potential (λ = 0.00001, 0.0001, 0.001, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 0.99, 0.999, 0.9999, 0.99999, 1). For each mutation, we performed at least five independent runs starting from different initial configurations (taken from the equilibration) for better convergence. The simulation time for each window was 0.3 ns. In each FEP λ-window, the first 4000 steps were for the further equilibration. Larger window sizes and longer simulation durations were also tested in our previous studies, and we found that the current setup gives us a reasonable convergence in the final binding affinities. Therefore, at least 60 ns (20 windows × 0.3 ns × 5 runs × 2 states) of simulation time were generated for each mutation, and the total aggregate simulation time for this study was ~2 μs, which is much longer than most FEP calculations currently reported in the literature.

Despite the controversy in the literature about the meaningfulness of breaking down the total free energy into components, and the ambiguity associated with a path-dependent decomposition (64–67), we believe a breakup of the total binding free energy into its van der Waals (vdW)

and electrostatic components can provide useful information about the underlying physical interactions involved in HA-Fab binding. In this study, we collected the contributions of free-energy change by vdW and electrostatic interactions separately. Because of the nonlinearity of the FEP formulation, there may be a small coupling term in this approach.

RESULTS AND DISCUSSION

Validation of the FEP protocol with known experimental data

Mutagenesis studies of the H5-F10 complex have shown the importance of the binding region between the αA helix of H5 and Fab, which was proposed to prevent the conformational change of αA helix and the subsequent membrane fusion (13–15) (see Fig. 1). We noticed that some experimental data are available for mutagenesis studies in the αA helix region. Therefore, we first used these experimental data to validate our FEP protocol by calculating the relative binding free energies for mutations $N50_2A$, $V52_2A$, $V52_2L$, and $V52_2E$ in the αA helix (subscripts 1 and 2 refer to HA1 and HA2, respectively, and the residue numbering system follows that of H3 (PDB entry: 2HMG)) (13).

The first target site, $Asn\text{-}50_2$, is exposed to the solvent in the αA helix, and hence is not in direct contact with Fab. Presumably, the mutation $N50_2A$ will not affect the binding affinity significantly. Our simulations showed that the $N50_2A$ mutation enhanced the binding affinity slightly, with $\Delta\Delta G = -0.28 \pm 0.28$ kcal/mol (equivalent to an ~1.6-fold enhancement (decrease) in binding dissociation constant K_d; see Table 1), which agreed well with an experimental result of an ~2-fold enhancement in K_d for the same mutation (13). For the second target site, $Val\text{-}52_2$, which is conserved in group 1 HA as directly interacting with the complementarity-determining region H1 (CDR-H1) of

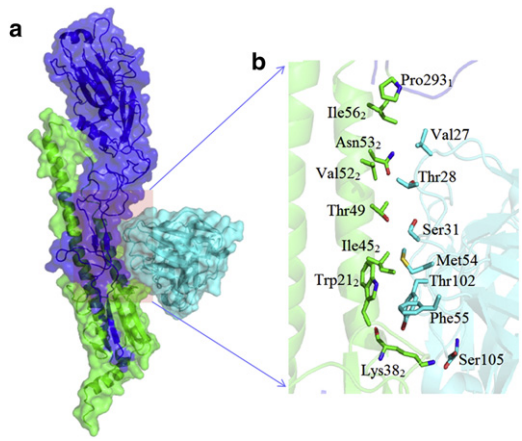

FIGURE 1 Molecular modeling system for HA protein binding with the antibody F10. (*a*) Overview of the HA-antibody complex structure. The HA and nAb are represented by surface and cartoon; HA1 and HA2 are colored blue and green, respectively, and both the heavy chain and light chain of the antibody are colored cyan. (*b*) Detailed view of the antigen-antibody binding interface, with the contact residues rendered with sticks. The total size of simulation system is 88,345 atoms for WT.

TABLE 1 Comparison of FEP simulation results with experimental data for the HA-nAb binding free-energy change due to the mutation in HA

Mutation	Calculated $\Delta\Delta$G (kcal/mol)	Reverse $\Delta\Delta$G (kcal/mol)	Experimental $\Delta\Delta G$ (kcal/mol)
$N50_2A$	−0.28 ± 0.28	0.11 ± 0.48 ($A50_2N$)	−0.4
$V52_2L$	0.57 ± 0.60	−0.20 ± 1.15 ($L52_2V$)	0~0.5
$V52_2A$	0.91 ± 0.29	−1.08 ± 0.59 ($A52_2V$)	1.3~1.6
$V52_2E$	9.62 ± 0.93	−8.22 ± 1.61 ($E52_2V$)	no bound states observed
$I56_2A$	2.42 ± 0.45	−1.81 ± 0.46 ($A56_2I$)	N/A

Total of five independent runs were performed for both the bound and free states for the standard error calculations, with each running 6 ns.

F10, we examined three different mutations: $V52_2A$, $V52_2L$, and $V52_2E$. An experimental site-directed mutagenesis study of $V52_2A$ revealed an ~10- to 20-fold increase in K_d (13) (equivalent to a 1.3~1.6 kcal/mol decrease in binding free energy), which was comparable to our FEP result of $\Delta\Delta G = 0.91 \pm 0.29$ kcal/mol (see Table 1). However, a conservative mutation to Leu of similar size as $Val\text{-}52_2$ had little influence on the binding affinity. We observed only a slight reduction in the binding affinity ($\Delta\Delta G = 0.57 \pm 0.60$ kcal/mol), which was also consistent with experiments in which the concentration of the bound state was decreased by <10% (13). As one would expect for the $V52_2E$ mutation, the inclusion of charged amino acid in the hydrophobic core at the binding interface severely interfered with the antibody association with the αA helix of HA, resulting in a significant decrease in binding affinity, $\Delta\Delta G = 9.62 \pm 0.93$ kcal/mol. The $V52_2E$ mutation on Fab F10 consistently failed to neutralize the HA in experiment (Table 1) (13).

We further validated our FEP approach by performing extra reverse FEP calculations in which five mutants ($Ala\text{-}50_2$, $Leu\text{-}52_2$, $Ala\text{-}52_2$, $Glu\text{-}52_2$, and $Ala\text{-}56_2$) were mutated back to their WTs ($Asn\text{-}50_2$, $Val\text{-}52_2$, $Val\text{-}52_2$, $Val\text{-}52_2$, and $Ile\text{-}56_2$), respectively (Table 1). We found that all five backward FEP simulations (mutants to WT) and forward FEP calculations (WT to mutants) showed similar binding affinities but with an opposite sign. The results of both methods are in good agreement with experimental results. Considering the large size of the HA-Fab system, our FEP calculation provided a relatively small deviation of <0.5 kcal/mol on average compared with available experimental data. Therefore, the FEP simulation protocol may be employed as a reliable tool for antigen-antibody complex binding affinity studies.

The stacking interaction and hydrophobic environment are crucial for HA binding to CDR-H2

Inspired by encouraging results from the above validation with known mutations, we extended our FEP method to novel (to our knowledge) mutations found in either HA or the Fab side at the binding interface. The cocrystal

structures of both H5-F10 and H5-CR6261 complexes suggest that the interactions between the αA helix of HA and CDRs of the Fab may be important for neutralization (13,14). As a key interaction site, we first focused on Trp-21_2 and Ile-56_2, two conserved residues that are observed in all subtypes of HAs, which make hydrophobic interactions with Met-54 and Phe-55 in the CDR-H2 of Fab F10 (equivalent to Ile-53 and Phe-54, respectively, in antibody CR6261).

In addition to the general hydrophobic interaction, a stacking interaction between the indole ring of Trp-21_2 and phenyl group of Phe-55 (or Phe-54) of Fab F10 (or CR6261) is believed to play a critical role in binding recognition between the αA helix of HA and the CDR-H2 of Fab F10 (i.e., by providing the binding specificity). Here, we performed computational mutagenesis for these sites to evaluate how interactions associated with the aromatic side chains ($\pi-\pi$ stacking) aid in the recognition of their binding partners. We started by mutating Trp-21_2 to Ala using our FEP simulation. Knocking out Trp-21_2 with a smaller amino-acid Ala resulted in a binding affinity decrease of $\Delta\Delta G = 4.02 \pm 0.34$ kcal/mol, which corresponds to an ~1000-fold increase in the dissociation constant K_d. Similarly, a counter mutation on the antibody's Phe-55 with Ala reduced the binding affinity by $\Delta\Delta G = 4.24 \pm 0.76$ kcal/mol, implying that the stacking interaction between two aromatic rings is indeed important, contributing ~4 kcal/mol to the binding affinity. The mutation effect on the binding site structure is clearly depicted in Fig. 2 for the WT and W21_2A at the end of simulation. Due to the lack of an aromatic ring and a shorter side chain in the mutant Ala-21_2, the intermolecular hydrophobic cluster made of Trp-21_2, Phe-55, and Met-54 is largely disrupted. On the other hand, we noticed enhancement of an intramolecular interaction between Met-54 and Phe-55 in CDR-H2 of Fab. A free-energy decomposition confirmed the role of vdW interactions (see Method and System for details). Almost all (98%) of the total contribution to the binding affinity originated from the vdW interactions (4.11 kcal/mol), whereas the electrostatic interactions played only a minor role (−0.05 kcal/mol). The coupling term was 0.04 kcal/mol. Note that our estimated stacking interaction (~4 kcal/mol) and the vdW component of the total free energy (4.11 kcal/mol) are very similar to the previously reported value of stacking interactions between two idealized benzene rings (4.08 kcal/mol) (68). Overall, our energetic analysis confirmed that the large free-energy loss in the W21_2A mutation is attributed to the destruction of stacking interactions between two aromatic residues from both HA and antibody. We further evaluated the effect of hydrophobic interactions with two nonconservative mutations: F55E (mutating to an acidic residue) and F55K (mutating to a basic residue). Much larger binding affinity decreases of 13.58 ± 0.55 kcal/mol and 7.66 ± 0.97 kcal/mol were found for F55E and F55K, respectively (Table 2), which supports the notion that the recognition of HA by nAb F10 is highly dependent on the hydrophobic interactions at the binding interface.

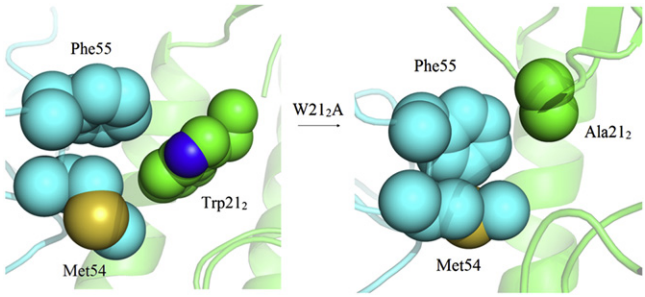

FIGURE 2 Structural comparison of F10 nAb bound to H5 HA around CDR-H2 due to the W21_2A mutation (HA, *green*; nAb, *cyan*). The overall complex is represented as a cartoon and the residues at interaction face are rendered with spheres. The interactions of Trp-21_2 (HA) with Met-54/Phe-55 (nAbs) are largely diminished by W21_2A substitution.

Given the crucial role of the hydrophobic interactions between these two aromatic residues (HA2's Trp-21 and F10's Phe-55), we further investigated the nearby nonpolar residue Met-54 in F10 (Ile-53 in CR6261), which displays strong contacts with Trp-21_2 and Ile-45_2 in HA2, forming a hydrophobic cluster. As expected, the nonconservative M54E or M54K mutations broke the hydrophobic core and caused ~10 kcal/mol free-energy decreases in our FEP simulations (Table 2). However, substituting Met with Ile, as found in Fab CR6261, or with Ala, was shown to destabilize the protein association by $\Delta\Delta G = 1.10 \pm 0.31$ kcal/mol and 2.13 ± 0.36 kcal/mol, respectively (Table 2). Note that the free-energy decreases in conservative mutations of M54I or M54A are smaller than those of aromatic residue deletions in W21_2A or F55A. This indicates that the stacking interaction between aromatic residues is more important for the Fab F10 neutralization, even though a hydrophobic environment around these key aromatic residues is still required.

Overall, our FEP simulations revealed in detail the determining factors in the interaction between CDR-H2 of Fab and HA. The stacking interactions between the aromatic rings from both HA and Fab were essential for the antigen-antibody binding. Preserving a hydrophobic environment

TABLE 2 FEP simulation results for the HA-nAb binding free-energy change due to the mutation in HA2/CDR-H2

Mutation	Calculated $\Delta\Delta G$ (kcal/mol)
W21_2A	4.02 ± 0.34
M54A	2.13 ± 0.36
M54L	1.10 ± 0.31
M54E	10.02 ± 1.11
M54K	9.29 ± 1.20
F55A	4.24 ± 0.76
F55E	13.58 ± 0.55
F55K	7.66 ± 0.97

around these key aromatic residues was found to be important as well, i.e., the existence of nonpolar residues, such as Met-54 in F10 and Ile-53 in CR6261, also helped.

Nonspecific hydrophobic interactions are responsible for the broad antibody neutralization

In addition to the hydrophobic interactions between HA and CDR-H2, another hydrophobic core is formed in the αA helix and the CDR-H1 (Fig. 3), where Pro-293_1, Val-52_2, and Ile-56_2 in HA are in contact with Val-27 and the methyl group of Thr-28 in F10. We first investigated the role of Val-52_2 using two types of FEP mutations: one with other nonpolar residues to keep the hydrophobic environment intact, and one with charged ones to break the hydrophobic core. The FEP calculation results are listed in Table 3. In general, for mutations to similarly sized hydrophobic residues, minor binding free-energy changes could be expected, but for mutations to charged residues, significant decreases in binding affinity or even disassociations of the HA-Fabs complex could be expected. Indeed, our simulations showed that the polarity of the substituted residue at site 52 of HA2 is critical for preserving the WT binding affinity. Of more interest, the conservative hydrophobic mutations displayed more or less similar binding affinities, with an ~ ± 1 kcal/mol variation (Table 3), indicating that the exact hydrophobic residue type is less critical in this binding environment. The mutations to charged amino acids, on the other hand, clearly broke up the antigen-antibody association with a high free-energy penalty of 7.28~15.47 kcal/mol (see Table 3). It should be noted that $V52_2I$ substitution could have even stronger binding with Fab than the WT, where the binding affinity enhanced by $\Delta\Delta G = -1.08 \pm 0.59$ kcal/mol, indicating that Fab is even more effective for HAs that possess an $I52_2$ residue. It is interesting to observe that Ile, Leu, and Val are the most common residue types at site 52 of HA2 for all of the subtypes of group 1 HA: Ile and Leu are two dominant amino acids in subtype H13/H16 and H12, respectively, whereas Val is the majority amino acid in other subtypes of group 1 HAs (Table 4). Thus, the strong binding affinity and the limited sequence diversity (within aliphatic amino acids Ile, Leu, or Val) at site 52 explains why antibody F10 could have broad neutralization toward various HAs.

TABLE 3 FEP simulation results for the HA-nAb binding free-energy change due to the mutation in HA/CDR-H1

Mutation	Calculated $\Delta\Delta G$ (kcal/mol)
H381N	1.32 ± 0.78
V522A	0.91 ± 0.29
V522I	−1.08 ± 0.59
V522L	0.57 ± 0.60
V522F	1.44 ± 0.87
V522E	10.43 ± 0.83
V522D	15.47 ± 1.09
V522R	13.45 ± 0.98
V522K	7.28 ± 1.51
I562A	2.42 ± 0.45
I562V	−0.48 ± 0.46
I562L	0.84 ± 0.50
I562E	2.24 ± 1.59
I562K	−0.38 ± 0.65

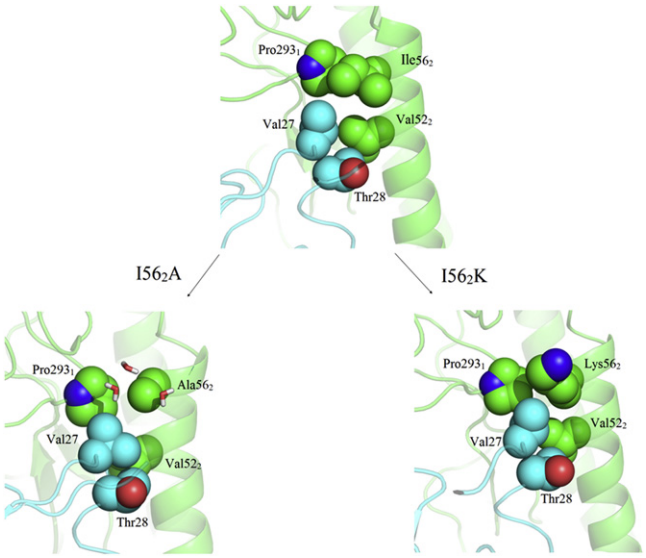

FIGURE 3 Structural comparison of F10 nAb bound to H5 HA around CDR-H2 due to $I56_2A$ and $I56_2K$ mutations (HA, *green*; nAb, *cyan*). The overall complex is represented as cartoon and the residues at interaction face are rendered with spheres. The hydrophobic core (shown in *spheres*) is broken by the $I56_2A$ mutation but is preserved in the $I56_2K$ mutation.

In contrast to the mutational effect at Val-52_2, which shows a clear binary dependence on residue polarity, the nearby hydrophobic residue Ile-56_2 displayed a more sophisticated pattern for the individual amino acid substitutions. For example, similarly sized $I56_2V$ and $I56_2L$ mutations showed similar binding activities as the WT, within ± 1 kcal/mol, whereas mutation to a smaller but still hydrophobic Ala residue reduced the binding affinity by ~2.42 ± 0.45 kcal/mol (Table 3). The free-energy decomposition analysis revealed that most of the free-energy loss (2.50 kcal/mol) was due to the weaker vdW interactions. This is clearly shown in Fig. 3, where the hydrophobic interaction in the WT Ile-56_2 with Pro-293_1 in HA1 and Val-27 of Fab has been weakened by the short side chain of Ala, and the extra space is partially filled by two water molecules making contact with Ala.

A more dramatic difference was observed in the $I56_2K$ mutation, where a nonconservative charged residue was introduced; however, the binding affinity was not affected or was slightly enhanced by $\Delta\Delta G = -0.38 \pm 0.65$ kcal/mol. Fig. 3 shows how the long side chain (also hydrophobic) and positively charged $NH3^+$ group of Lys could be packed in the binding interface. The terminal amine group is pointing out from the hydrophobic binding interface and interacts with hydrophilic residues such as Asn-53_2 and Asn-60_2, whereas

TABLE 4 Comparison of the sequence conservation among 16 HA subtypes

			HA1				HA2																
Group	Cluster	Subtype	17	18	38	40	18	19	20	21	38	41	42	44	45	46	48	49	52	53	55	56	111
Group 1	H1a	H2	Y	H	H	K	V	D	G	W	K	T	Q	A	I	D	I	T	V	N	V	I	H
		H5	Y	H	H	Q	V	D	G	W	Q	T	Q	A	I	D	V	T	V	N	I	V	H
		H1	Y	H	H	V	V	D	G	W	K	T	Q	A	I	D	I	T	V	N	V	I	H
		H6	Y	H	H	V	I	D	G	W	K	T	Q	A	I	D	I	T	V	N	I	I	H
	H1b	H13	Y	L	S	I	I	N	G	W	K	T	Q	A	I	D	I	T	I	N	I	I	H
		H16	Y	L	S	V	I	N	G	W	K	T	Q	A	I	D	I	T	I	N	I	I	H
		H11	Y	L	S	V	I	N	G	W	K	T	Q	A	I	D	I	T	V	N	I	V	H
	H9	H8	Y	Q	Q	M	I	D	G	W	Q	T	Q	A	I	D	I	T	V	N	I	I	H
		H12	Y	Q	Q	E	V	A	G	W	R	T	Q	A	I	D	M	Q	L	N	V	I	H
		H9	Y	Q	H	K	V	A	G	W	R	T	Q	A	I	D	I	T	V	N	I	V	H
Group 2	H3	H4	H	H	T	Q	I	D	G	W	L	T	Q	A	I	D	I	T	L	N	L	I	T
		H14	H	H	S	K	I	D	G	W	L	T	Q	A	I	D	I	N	L	N	L	I	T
		H3	H	H	N	T	V	D	G	W	L	T	Q	A	I	D	I	N	L	N	V	I	T
	H15	H15	H	H	N	T	I	D	G	W	Y	T	Q	A	I	D	I	T	L	N	L	I	A
		H7	H	H	N	T	I	D	G	W	Y	T	Q	A	I	D	I	T	L	N	L	I	A
		H10	H	H	N	T	V	D	G	W	Y	T	Q	A	I	D	I	T	L	N	L	I	A

the long aliphatic chain of Lys still makes favorable contacts with hydrophobic residues Val-27, Val-52_2, and Pro-293_1, thereby maintaining a strong interaction between the αA helix and CDR-H1.

Compared with the highly conservative aromatic residues in the binding region of HA and Fab-CDR-H2 discussed above, we found that the binding interface between HA and Fab-CDR-H1 exhibited relatively more variability in amino-acid selection in HA. Hydrophobic residues were usually required for site 52_2 to maintain the Fab binding. However, the binding interface near site 56_2 required a tightly packed hydrophobic environment. We found that the $V52_2I$ or $I56_2V$ mutation could actually increase the binding affinity by ~1 kcal/mol and ~0.5 kcal/mol, respectively. A strong but nonspecific binding between HA and antibody could be an important principle to consider in designing vaccines with a broad neutralization.

Asn-38 in group 2 HA1 may contribute to the antibody neutralization escape

The molecular mechanism that prevents most group 2 HAs (e.g., H3 and H7) from being neutralized by F10 or CR6261 is not fully understood. In previous studies, glycosylation at position 38_1 on group 2 HA was proposed as the main reason for neutralization escape (13,14). However, we found that two of the six subtypes in group 2 are not glycosylated at position 38_1, yet are able to escape the neutralization. That is, the glycosylation may not be the only mechanism to explain the immunity escape of group 2 subtypes. We investigated possible molecular mechanisms beyond the glycosylation by searching the sequence diversity around the HA-Fab binding interface. The sequences of all HA subtypes were collected from the National Center for Biotechnology Information Flu Database (Table 4). We hypothesized that a residue site in HA would be important for the antibody's escape from neutralization if the amino-acid types of that specific site were diversified over different groups but conserved within each group. That is, antibodies could neutralize HA subtypes in one group with a specific amino-acid type conserved at that site, whereas HA subtypes in the other group might escape from the same antibodies due to a different amino-acid type being conserved at the same site. For site 38_1, His is a well-conserved residue within group 1 (e.g., H1, H5, H2, H6, and H9), whereas Asn is a conserved residue within group 2 (in four out of six subtypes (H3, H7, H10, and H15)). Other sites, such as 40_1 and 38_2, are not fully conserved either across groups or within individual groups, indicating that they may play a less critical role in the neutralization (Table 4).

To confirm our hypothesis from the above sequence analysis and to reveal the atomic detail of the binding specificity, we performed the FEP mutation for His-38_1 in our current group 1 H5N1 HA to Asn-38, mimicking the conservative amino acid in group 2 HA. The simulated FEP results showed that the binding affinity between HA and Fab decreased by 1.32 $\pm$ 0.78 kcal/mol, which is an ~10-fold increase in the dissociation constant (K_d) value. When we compared the final structure with the WT, we found that the two native hydrogen bonds (nitrogen atoms at the side chain of His-38_1 to the hydroxyl group of Ser-30 or Gln-64 in Fab) were broken by the mutation (Fig. 4). The loss of side-chain hydrogen bonding was partially compensated by waters entered in the binding interface, but no direct contact was found between mutated Asn-38_1 and Fab. Therefore, the $H38_1N$ mutation weakened the interaction between HA and Fab by breaking the hydrogen bonds between HA1 and the antibody. However, it should be noted that the binding affinity decrease of 1.32 $\pm$ 0.78 kcal/mol in the $H38_1N$ mutation is not overwhelming, indicating that antibody neutralization is a complicated process, and that although this His-38 residue may be an important

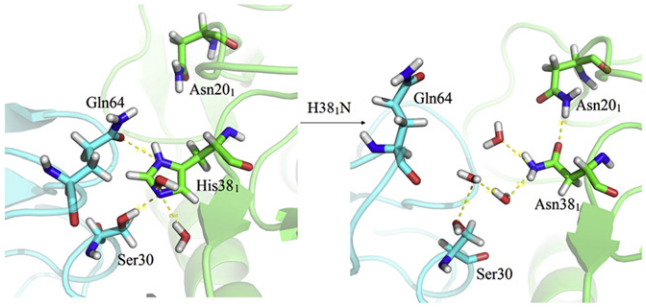

FIGURE 4 Structural comparison of F10 nAb bound to H5 HA around CDR-H2 due to H38$_1$N mutations (HA, *green*; nAb, *cyan*). The overall complex is represented as cartoon and the residues at interaction face are rendered with sticks. The side-chain hydrogen bonds between His-38$_1$/Gln64 and His-38$_1$/Ser-30 (shown with *yellow dashed line*) are broken by H38$_1$N mutation.

contributing factor, it is probably not the only factor (other factors, such as glycosylation, may contribute as well (13,14)). Indeed, several other group 1 subtypes that are bound/neutralized by F10, such as H8 and H11, do not have His at position 38, and instead have Gln or Ser, respectively (13). Nevertheless, our FEP calculations indicate that in addition to glycosylation, the existence of Asn at position 38$_1$ may also be an important contributing factor in neutralization escape in group 2 HA subtypes.

CONCLUSION

In this work, we performed rigorous FEP calculations to estimate influenza antigen-antibody binding affinities (using H5 HA and F10 Fab as a template) and study the characteristics of antibodies with a broad neutralization capability (such as F10 and CR6261). The simulated binding affinities between HA and Fab were in excellent agreement with the currently available experimental data. We identified several key residues in the HA-Fab binding regions and further examined them by means of in silico mutagenesis studies to explore the molecular mechanism of HA-Fab binding.

We found that the stacking interaction of Trp-21$_2$ (HA) and Phe-55 (Fab) is critical for strong binding between the αA helix and the CDR-H2 of antibody. Our FEP simulations suggest that either W21$_2$A in (HA side) or F55A (antibody side) will cause a significant binding affinity decrease of $\Delta\Delta G > 4.0$ kcal/mol (equivalent to an ~1000-fold increase in the binding dissociation constant K_d). In addition, neighboring hydrophobic residues were also required to preserve a stable hydrophobic network around the aromatic side chains. Furthermore, we observed more general hydrophobic interactions between HA and the CDR-H1 of Fab.

The HA residue sites 52$_2$ and 56$_2$ appeared to be more tolerable with various hydrophobic mutations with similar binding ability as the WT, which could explain the wide neutralization of Fabs among all group 1 subtypes. In addition, we found that the V52$_2$I and I56$_2$V substitutions could increase the binding affinity by ~1 kcal/mol and ~0.5 kcal/mol, respectively, which potentially could be used as a way to improve the efficiency of current antibodies.

In addition to the hydrophobic interactions, the hydrogen bonding between His-38$_1$ and Ser-30/Gln-64 were also found to be important for antibody neutralization. When His-38$_1$ was mutated to group 2-like Asn-38$_1$, two hydrogen bonds were lost, substituted by hydration around Asn-38$_1$ in between the HA and the Fabs, with a net decrease of ~1.3 kcal/mol in binding affinity. This could be another important contributing factor for the neutralization escape in group 2 subtypes, in addition to the glycosylation.

We thank Bruce Berne, Payel Das, Ajay Royyuru, Pengyu Ren, Steven Turner, and David Topham for many useful discussions. We also thank Ian Wilson, Jim Paulson, and Peter Palese for helpful comments at the beginning of our influenza modeling effort.

REFERENCES

1. Cox, N. J., and K. Subbarao. 2000. Global epidemiology of influenza: past and present. *Annu. Rev. Med.* 51:407–421.
2. Hilleman, M. R. 2002. Realities and enigmas of human viral influenza: pathogenesis, epidemiology and control. *Vaccine.* 20:3068–3087.
3. Stevens, J., O. Blixt, …, I. A. Wilson. 2006. Glycan microarray analysis of the hemagglutinins from modern and pandemic influenza viruses reveals different receptor specificities. *J. Mol. Biol.* 355: 1143–1155.
4. Fleury, D., S. A. Wharton, …, T. Bizebard. 1998. Antigen distortion allows influenza virus to escape neutralization. *Nat. Struct. Biol.* 5:119–123.
5. Yang, Z. Y., C. J. Wei, …, G. J. Nabel. 2007. Immunization by avian H5 influenza hemagglutinin mutants with altered receptor binding specificity. *Science.* 317:825–828.
6. Tumpey, T. M., T. R. Maines, …, A. García-Sastre. 2007. A two-amino acid change in the hemagglutinin of the 1918 influenza virus abolishes transmission. *Science.* 315:655–659.
7. Stevens, J., A. L. Corper, …, I. A. Wilson. 2004. Structure of the uncleaved human H1 hemagglutinin from the extinct 1918 influenza virus. *Science.* 303:1866–1870.
8. Stevens, J., O. Blixt, …, I. A. Wilson. 2006. Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. *Science.* 312:404–410.
9. Gamblin, S. J., L. F. Haire, …, J. J. Skehel. 2004. The structure and receptor binding properties of the 1918 influenza hemagglutinin. *Science.* 303:1838–1842.
10. Lin, T., G. Wang, …, K. Y. Yuen. 2009. The hemagglutinin structure of an avian H1N1 influenza A virus. *Virology.* 392:73–81.
11. Igarashi, M., K. Ito, …, A. Takada. 2010. Predicting the antigenic structure of the pandemic (H1N1) 2009 influenza virus hemagglutinin. *PLoS ONE.* 5:e8553.
12. Xu, R., D. C. Ekiert, …, I. A. Wilson. 2010. Structural basis of preexisting immunity to the 2009 H1N1 pandemic influenza virus. *Science.* 328:357–360.
13. Sui, J., W. C. Hwang, …, W. A. Marasco. 2009. Structural and functional bases for broad-spectrum neutralization of avian and human influenza A viruses. *Nat. Struct. Mol. Biol.* 16:265–273.
14. Ekiert, D. C., G. Bhabha, …, I. A. Wilson. 2009. Antibody recognition of a highly conserved influenza virus epitope. *Science.* 324:246–251.
15. Throsby, M., E. van den Brink, …, J. Goudsmit. 2008. Heterosubtypic neutralizing monoclonal antibodies cross-protective against H5N1 and

H1N1 recovered from human IgM^+ memory B cells. *PLoS ONE.* 3:e3942.

16. Russell, R. J., P. S. Kerry, …, J. J. Skehel. 2008. Structure of influenza hemagglutinin in complex with an inhibitor of membrane fusion. *Proc. Natl. Acad. Sci. USA.* 105:17736–17741.
17. Das, P., J. Li, …, R. Zhou. 2009. Free-energy simulations reveal a double mutant avian H5N1 virus hemagglutinin with altered receptor binding specificity. *J. Comput. Chem.* 30:1654–1663.
18. Zhou, R., P. Das, and A. K. Royyuru. 2008. Single mutation induced H3N2 hemagglutinin antibody neutralization: a free-energy perturbation study. *J. Phys. Chem. B.* 112:15813–15820.
19. Chen, R., and E. C. Holmes. 2006. Avian influenza virus exhibits rapid evolutionary dynamics. *Mol. Biol. Evol.* 23:2336–2341.
20. Xia, Z., G. Jin, …, R. Zhou. 2009. Using a mutual information-based site transition network to map the genetic evolution of influenza A/H3N2 virus. *Bioinformatics.* 25:2309–2317.
21. Smith, D. J., A. S. Lapedes, …, R. A. Fouchier. 2004. Mapping the antigenic and genetic evolution of influenza virus. *Science.* 305:371–376.
22. Du, X., Z. Wang, …, T. Jiang. 2008. Networks of genomic co-occurrence capture characteristics of human influenza A (H3N2) evolution. *Genome Res.* 18:178–187.
23. Yamada, S., Y. Suzuki, …, Y. Kawaoka. 2006. Haemagglutinin mutations responsible for the binding of H5N1 influenza A viruses to human-type receptors. *Nature.* 444:378–382.
24. Apisarnthanarak, A., S. Erb, …, L. M. Mundy. 2005. Seroprevalence of anti-H5 antibody among Thai health care workers after exposure to avian influenza (H5N1) in a tertiary care center. *Clin. Infect. Dis.* 40:e16–e18.
25. Mehta, T., E. McGrath, …, G. J. Alangaden. 2010. Detection of oseltamivir resistance during treatment of 2009 H1N1 influenza virus infection in immunocompromised patients: utility of cycle threshold values of qualitative real-time reverse transcriptase PCR. *J. Clin. Microbiol.* 48:4326–4328.
26. Stephenson, I., J. Democratis, …, M. Zambon. 2009. Neuraminidase inhibitor resistance after oseltamivir treatment of acute influenza A and B in children. *Clin. Infect. Dis.* 48:389–396.
27. Beigel, J., and M. Bray. 2008. Current and future antiviral therapy of severe seasonal and avian influenza. *Antiviral Res.* 78:91–102.
28. Lowen, A. C., and P. Palese. 2007. Influenza virus transmission: basic science and implications for the use of antiviral drugs during a pandemic. *Infect. Disord. Drug Targets.* 7:318–328.
29. Fick, J., R. H. Lindberg, …, B. Olsen. 2007. Antiviral oseltamivir is not removed or degraded in normal sewage water treatment: implications for development of resistance by influenza A virus. *PLoS ONE.* 2:e986.
30. Bright, R. A., D. K. Shay, …, A. I. Klimov. 2006. Adamantane resistance among influenza A viruses isolated early during the 2005-2006 influenza season in the United States. *JAMA.* 295:891–894.
31. de Jong, M. D., T. T. Tran, …, J. Farrar. 2005. Oseltamivir resistance during treatment of influenza A (H5N1) infection. *N. Engl. J. Med.* 353:2667–2672.
32. Kiso, M., K. Mitamura, …, Y. Kawaoka. 2004. Resistant influenza A viruses in children treated with oseltamivir: descriptive study. *Lancet.* 364:759–765.
33. Jorgensen, W. L. 1989. Free-energy calculations—a breakthrough for modeling organic-chemistry in solution. *Acc. Chem. Res.* 22:184–189.
34. Deng, Y., and B. Roux. 2006. Calculation of standard binding free energies: aromatic molecules in the T4 lysozyme L99A mutant. *J. Chem. Theory Comput.* 2:1255–1273.
35. Kollman, P. 1993. Free-energy calculations—applications to chemical and biochemical phenomena. *Chem. Rev.* 93:2395–2417.
36. Pathiaseril, A., and R. J. Woods. 2000. Relative energies of binding for antibody-carbohydrate-antigen complexes computed from free-energy simulations. *J. Am. Chem. Soc.* 122:331–338.
37. Simonson, T., G. Archontis, and M. Karplus. 2002. Free-energy simulations come of age: protein-ligand recognition. *Acc. Chem. Res.* 35:430–437.
38. Tembe, B. L., and J. A. McCammon. 1984. Ligand receptor interactions. *Comput. Chem.* 8:281–283.
39. Warshel, A. 1984. Simulating the energetics and dynamics of enzymatic reactions. *Pontif. Acad. Sci. Scripta Varia.* 55:60–81.
40. Warshel, A., P. K. Sharma, …, W. W. Parson. 2006. Modeling electrostatic effects in proteins. *Biochim. Biophys. Acta.* 1764:1647–1676.
41. Xia, Z., Z. Zhu, …, R. Zhou. 2009. Recognition mechanism of siRNA by viral p19 suppressor of RNA silencing: a molecular dynamics study. *Biophys. J.* 96:1761–1769.
42. Zheng, L., M. Chen, and W. Yang. 2008. Random walk in orthogonal space to achieve efficient free-energy simulation of complex systems. *Proc. Natl. Acad. Sci. USA.* 105:20227–20232.
43. Jiao, D., P. A. Golubkov, …, P. Ren. 2008. Calculation of protein-ligand binding free-energy by using a polarizable potential. *Proc. Natl. Acad. Sci. USA.* 105:6290–6295.
44. Chodera, J. D., D. L. Mobley, …, V. S. Pande. 2011. Alchemical free-energy methods for drug discovery: progress and challenges. *Curr. Opin. Struct. Biol.* 21:150–160.
45. Almlöf, M., J. Aqvist, …, B. O. Brandsdal. 2006. Probing the effect of point mutations at protein-protein interfaces with free-energy calculations. *Biophys. J.* 90:433–442.
46. Brandsdal, B. O., and A. O. Smalås. 2000. Evaluation of protein-protein association energies by free-energy perturbation calculations. *Protein Eng.* 13:239–245.
47. Wang, J., Y. Deng, and B. Roux. 2006. Absolute binding free-energy calculations using molecular dynamics simulations with restraining potentials. *Biophys. J.* 91:2798–2814.
48. Darden, T. A., D. M. York, and L. G. Pedersen. 1993. Particle mesh Ewald: an NlogN method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.
49. Eleftheriou, M., R. S. Germain, …, R. Zhou. 2006. Thermal denaturing of mutant lysozyme with both the OPLSAA and the CHARMM force fields. *J. Am. Chem. Soc.* 128:13388–13395.
50. Liu, P., X. Huang, …, B. J. Berne. 2005. Observation of a dewetting transition in the collapse of the melittin tetramer. *Nature.* 437:159–162.
51. Zhou, R., M. Eleftheriou, …, B. J. Berne. 2007. Destruction of long-range interactions by a single mutation in lysozyme. *Proc. Natl. Acad. Sci. USA.* 104:5824–5829.
52. Zhou, R., X. Huang, …, B. J. Berne. 2004. Hydrophobic collapse in multidomain protein folding. *Science.* 305:1605–1609.
53. Gao, Y. Q., W. Yang, and M. Karplus. 2005. A structure-based model for the synthesis and hydrolysis of ATP by F_1-ATPase. *Cell.* 123:195–205.
54. Hummer, G., J. C. Rasaiah, and J. P. Noworyta. 2001. Water conduction through the hydrophobic channel of a carbon nanotube. *Nature.* 414:188–190.
55. Kamberaj, H., and A. van der Vaart. 2009. An optimized replica exchange molecular dynamics method. *J. Chem. Phys.* 130:074906.
56. Karplus, M., Y. Q. Gao, …, W. Yang. 2005. Protein structural transitions and their functional role. *Philos. Transact. A Math. Phys. Eng. Sci.* 363:331–355, discussion 355–356.
57. Hua, L., X. Huang, …, B. J. Berne. 2007. Nanoscale dewetting transition in protein complex folding. *J. Phys. Chem. B.* 111:9069–9077.
58. Li, X., J. Li, …, R. Zhou. 2006. Hydration and dewetting near fluorinated superhydrophobic plates. *J. Am. Chem. Soc.* 128:12439–12447.
59. Kumar, S., C. Huang, …, L. V. Kale. 2008. Scalable molecular dynamics with NAMD on Blue Gene/L. *IBM J. Res. Dev.* 52:177–188.
60. MacKerell, A. D., D. Bashford, …, M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102:3586–3616.

61. Jorgensen, W. L., J. Chandrasekhar, …, M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

62. Beutler, T. C., A. E. Mark, …, W. F. v. Gunsteren. 1994. Avoiding singularities and numerical instabilities in free-energy calculations based on molecular simulations. *Chem. Phys. Lett.* 222:529–539.

63. Zacharias, M., T. P. Straatsma, and J. A. McCammon. 1994. Separation-shifted scaling, a new scaling method for Lennard-Jones interactions in thermodynamic integration. *J. Chem. Phys.* 100:9025–9031.

64. Boresch, S., and M. Karplus. 1995. The meaning of component analysis: decomposition of the free-energy in terms of specific interactions. *J. Mol. Biol.* 254:801–807.

65. Brady, G. P., and K. A. Sharp. 1995. Decomposition of interaction free energies in proteins and other complex systems. *J. Mol. Biol.* 254:77–85.

66. Bren, M., J. Florian, …, U. Bren. 2007. Do all pieces make a whole? Thiele cumulants and the free-energy decomposition. *Theor. Chem. Acc.* 117:535–540.

67. Mark, A. E., and W. F. van Gunsteren. 1994. Decomposition of the free-energy of a system in terms of specific interactions. Implications for theoretical and experimental studies. *J. Mol. Biol.* 240:167–176.

68. Golubkov, P. A., and P. Y. Ren. 2006. Generalized coarse-grained model based on point multipole and Gay-Berne potentials. *J. Chem. Phys.* 125:64103.