THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

Single Mutation Induced H3N2 Hemagglutinin Antibody Neutralization: A Free Energy Perturbation Study

Ruhong Zhou, Payel Das, and Ajay K. Royyuru

J. Phys. Chem. B, **2008**, 112 (49), 15813-15820 • Publication Date (Web): 01 November 2008

Single Mutation Induced H3N2 Hemagglutinin Antibody Neutralization: A Free Energy Perturbation Study

Ruhong Zhou,*,†,‡ Payel Das,† and Ajay K. Royyuru†

IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598, and Department of Chemistry, Columbia University, New York, New York 10027

Received: June 23, 2008; Revised Manuscript Received: September 15, 2008

The single mutation effect on the binding affinity of H3N2 viral protein hemagglutinin (HA) with the monoclonical antibody fragment (Fab) is studied in this paper using the free energy perturbation (FEP) simulations. An all-atom protein model with explicit solvents is used to perform an aggregate of several microsecond FEP molecular dynamics simulations. A recent experiment shows that a single mutation in H3N2 HA, T131I, increases the antibody−antigen dissociation constant K_d by a factor of ∼4000 (equivalent to a binding affinity decrease of ∼5 kcal/mol), thus introducing an escape of the antibody (Ab) neutralization. Our FEP result confirms this experimental finding by estimating the HA−Ab binding affinity decrease of 5.2 ± 0.9 kcal/mol but with a somewhat different molecular mechanism from the experimental findings. Detailed analysis reveals that this large binding affinity decrease in the T131I mutant is mainly due to the displacement of two bridge water molecules otherwise present in the wild-type HA/Ab interface. The decomposition of the binding free energy supports this observation, as the major contribution to the binding affinity is from the electrostatic interactions. In addition, we find that the loss of the binding affinity is also related to the large conformational distortion of one loop (loop 155-161) in the unbound state of the mutant. We then simulate all other possible mutations for this specific mutation site T131, and predict a few more mutations with even larger decreases in the binding affinity (i.e., better candidates for antibody neutralization), such as T131W, T131Y, and T131F. As for further validation, we have also modeled another mutation, S157L, with experimental binding affinity available (K_d increasing ∼500 times), and found a binding affinity decrease of 4.1 ± 1.0 kcal/mol, which is again in excellent agreement with experiment. These large scale simulations might provide new insights into the detailed physical interaction, possible future escape mutation, and antibody−antigen coevolution relationship between influenza virus and human antibodies.

Introduction

The wide spread of avian flu, including the 1918 H1N1, 1968 H3N2, and more recent H5N1, has caused a great public health concern due to the emergence of potential pandemic threats.[1−7] Though new vaccines are in development against both H3N2 and H5N1, it is unclear if they will be effective against future strains due to the high mutation rates of influenza virus.[2,7−12] Thus, new techniques that allow for both the prediction of future mutations and the development of appropriate vaccines (antibodies) are in great need for better preparation of future pandemics. The escape of antibody (Ab) neutralization by influenza virus is accompanied by mutations in the viral surface glycoprotein hemagglutinin (HA), the principal antigen on the influenza virus surface. HA has emerged as a critical target for both vaccine and drug development in recent years, in addition to neuraminidase (NA). Given the current lack of H5N1 HA/Ab complex structures,[10,11] many of the studies have focused on H3N2,[2,9,13] for which an X-ray complex structure is available along with some limited binding affinity data.[2] In a recent experiment, Skehel and co-workers[2] have shown that a single mutation T131I in H3N2 HA can lead to an escape of the monoclonal antibody fragment (Fab) neutralization, i.e., a significant decrease in the antibody−antigen binding affinity. A 4000-fold increase in the dissociation constant K_d has been reported, which is equivalent to an ∼5 kcal/mol decrease in binding affinity for this single mutation T131I. Skehel and co-workers have also proposed a mechanism for this binding affinity decrease which involves the burial of LYS156's carbonyl group as well as the distortion of one loop structure (loop 155-161) in the unbound (free) state due to the mutation.[2] However, the exact molecular mechanism remains somewhat unclear. It is also unclear which physical interactions dominate the antibody−antigen binding affinity, which might be useful for the prediction of emerging mutations and future antibody design. Furthermore, very few data points are available for the mutation induced binding affinity (or dissociation constant) change.

In this study, we use a large scale free energy perturbation (FEP) method to model this H3N2 antibody−antigen complex system. The goal is to provide a better understanding of the molecular mechanism behind the single mutation induced antibody neutralization, and to further predict other possible mutations with even greater capability in escaping antibody neutralization. The FEP method has been widely used to calculate binding affinities for a variety of biophysical phenomena, such as solvation free energies, enzyme catalysis, redox, pK_a, ion conductance, ligand−receptor binding, protein−protein interaction, and protein−DNA (RNA) binding, etc.[14−29] Among several available computational methods developed in past years, FEP using an all-atom explicit solvent model probably provides the most accurate approach for our current

* To whom correspondence may be addressed. E-mail: ruhongz@us.ibm.com. Phone: 914-945-3591. Fax: 914-945-4104.
† IBM Thomas J. Watson Research Center.
‡ Columbia University.

10.1021/jp805529z CCC: $40.75
Published on Web 11/01/2008

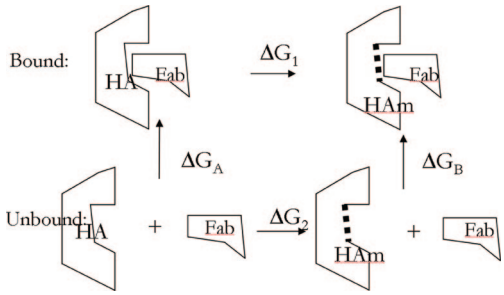


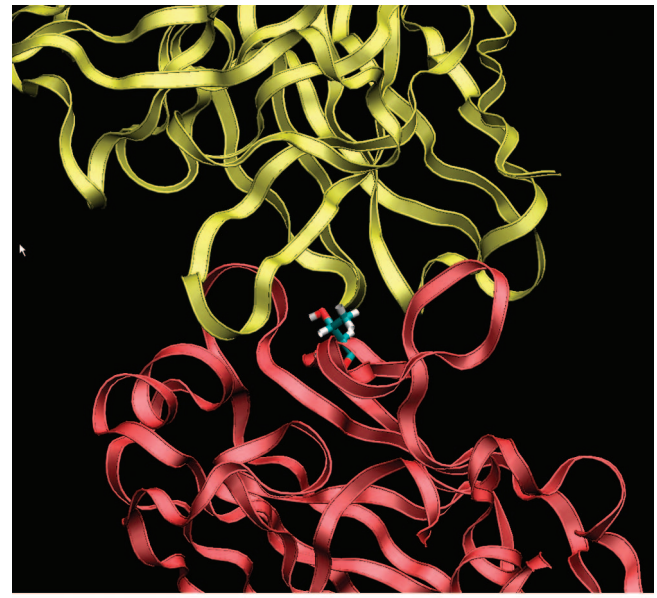
Figure 1. Scheme of the thermodynamical cycle for the calculation of the binding affinity change due to a mutation (HA, hemagglutinin; Ham, mutant hemagglutinin; Fab, antibody fragment).

needs in estimating the relative antibody−antigen binding affinity due to a single mutation.[29,30] However, such simulations for realistic biological systems often require extensive computational resources. Here, we have utilized the massively parallel molecular dynamics (MD) software developed on IBM BlueGene[31−36] to perform these FEP calculations. We have obtained a 5.2 ± 0.9 kcal/mol decrease in the HA/Ab binding affinity due to this T131I single mutation, in excellent agreement with experiment, but with a somewhat different molecular mechanism. We have also predicted a few more mutations on the T131 site which show even larger decreases in the binding affinity (even better candidates for antibody neutralization). These predictions from the current FEP calculations can provide a testing ground for the experimental validation of the antibody reactivity, which might ultimately help us to chart a path toward effective influenza virus neutralization. The approach we have developed in this study can also be applied to H5N1, which is of even greater concern due to its high mortality rate, once the H5N1 HA/Ab complex structures become available.

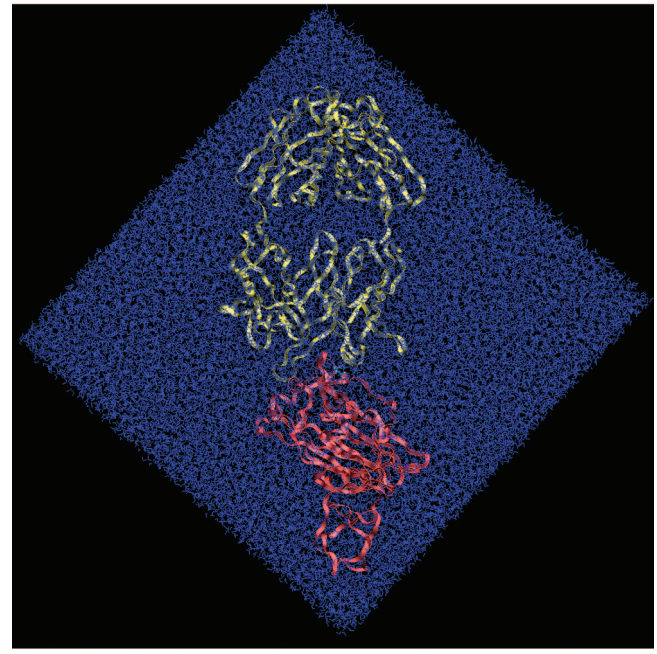
Results and Discussion

The rigorous FEP method[24−26] is used to estimate the binding affinities between HA and Abs, as mentioned above. In order to calculate this binding affinity change from one residue (e.g., THR131, named state A) to another (e.g., ILE131, named state B), we designed a thermodynamical cycle in FEP, as shown in Figure 1. This binding free energy change $\Delta\Delta G_{AB}$ can then be obtained from the difference between the free energy changes caused by the particular mutation in the bound state (ΔG_1) and the unbound or free state (ΔG_2)[24−26] (see the Method and System section for more details). In a typical FEP calculation for a mutation from state A to state B, many perturbation windows are needed in order to obtain a smooth transition from the initial state A to the final state B, usually with more windows near the two ends to enhance the sampling statistics. In our current simulations, a 22-window scheme has been adopted to achieve a high precision. Figure 2 shows the simulation system setup for the H3N2 HA/Fab binding complex, with a realistic all-atom protein model in explicit solvent. The antigen/antibody complex has one monomer of HA and one Fab antibody fragment solvated in a water box, as shown in Figure 2b (see more details in the Method and System section). With 1024 processors of BlueGene/L (half-rack running in virtual-node-mode), we can propagate the ~113 000 total atom system by approximately 25 ns/day in molecular dynamics (MD) using a 1.5 fs time step.

Table 1 lists some representative FEP simulations on the binding affinity changes (ΔG_1, ΔG_2, and $\Delta\Delta G$) due to the T131I mutation. Two different mutation directions have been performed: one forward from THR131 to ILE131 and the other



(a)



(b)

Figure 2. Molecular modeling system for the H3N2 hemagglutinin (HA, colored red) binding with the antibody fragment Fab (colored yellow). (a) A closer view of the binding site (epitope) by zooming into the mutation site THR131 (represented by sticks). (b) The simulation box of the full system, with the HA and Fab represented by ribbons and water molecules by wires (colored blue). The total system size is 112 711 atoms.

backward from ILE131 to THR131, each with two different-length single runs, one 0.3 ns/window (total 6.6 ns) and the other 3 ns/window (total 66 ns). The backward mutation I131T presumably should display the same magnitude of the free energy change but with an opposite sign. Indeed, Table 1 shows a similar magnitude of the binding affinity changes for both the bound (ΔG_1) and free (ΔG_2) states in the forward and backward mutations. For example, in the forward T131I longer simulation, the FEP shows a free energy change of 31.4 kcal/mol for the bound state, and 26.1 kcal/mol for the free state, resulting in $\Delta\Delta G = 5.3$ kcal/mol, while the backward I131T

TABLE 1: FEP Simulation Results for the H3N2 HA/Fab Binding Free Energy Change due to the T131I Mutation[a]

mutation	MD steps/ window	total MD time (ns)	ΔG_1 (kcal/mol)	ΔG_2 (kcal/mol)	$\Delta\Delta G$ (kcal/mol)
T131I	200 000	6.6	31.38	24.97	6.41
	2 000 000	66	31.41	26.07	5.34
I131T	200 000	6.6	−31.37	−26.39	−4.98
(reverse)	2 000 000	66	−30.73	−25.03	−5.70

[a] A reverse mutation I131T is also being performed to check the consistency of results. A total of 22 windows are used in the FEP calculations for both the bound and the free states, with a time step of 1.5 fs in the MD simulation.

TABLE 2: FEP Simulation Results for the H3N2 HA/Fab Binding Free Energy Change due to Various Mutations of T131[a]

mutation	calcd $\Delta\Delta G$ (kcal/mol)	exptl $\Delta\Delta G$ (kcal/mol)	mutation	calcd $\Delta\Delta G$ (kcal/mol)	exptl $\Delta\Delta G$ (kcal/mol)
T131I	5.20 ± 0.94	5.0	T131F	5.68 ± 1.48	
T131G	-3.72 ± 0.69		T131W	7.46 ± 1.91	
T131A	-2.81 ± 0.91		T131L	3.15 ± 1.19	
T131C	0.117 ± 1.24		T131H	3.84 ± 1.17	
T131V	2.58 ± 0.89		T131Y	6.01 ± 1.31	
T131M	0.57 ± 1.63		T131N	2.92 ± 1.16	
T131Q	1.22 ± 1.20				
T131S	-0.48 ± 1.57		S157L	4.09 ± 1.03	3.7

[a] A total of five independent runs are performed for both the bound and free states for the standard deviation calculations with each running 6.6 ns.

mutation shows −30.7 kcal/mol for the bound state and −25.0 kcal/mol for the free state, resulting in $\Delta\Delta G = -5.7$ kcal/mol. Overall, the two longer (66 ns), the forward and backward, FEP calculations show an average of 5.5 kcal/mol decrease in the binding affinity upon the mutation from the wild types THR131 to ILE131. This data is in excellent agreement with the experimental results, a 4000-fold increase in the dissociation constant K_d value, which is equivalent to an ~5 kcal/mol decrease in binding affinity.[2] The shorter runs (6.6 ns) show a slightly larger deviation from the average value, with the forward T131I mutation of 6.41 kcal/mol and the backward I131T mutation of −4.98 kcal/mol, indicating that more conformational space sampling is indeed helpful in determining the binding affinity (more below). It is widely recognized that, for effective FEP calculations, a sufficient conformational space sampling is often critical.[25,26,37−39] Many groups have developed efficient sampling methods to tackle this problem.[25,37−39] Warshel and co-workers also pointed out that a satisfactory convergence in FEP is not only related to the extensive sampling but also to the proper electrostatic boundary condition and long-range treatment (such as local reaction field, LRF).[15,16,40−42] Our current approach is to develop massively parallel MD softwares[31,36] on supercomputers such as IBM BlueGene/L to perform these large scale FEP simulations. The other significant effort along these lines is from Pande's group with tens of thousands PC clusters at Folding@Home.[26]

A decomposition of the total binding affinity into its van der Waals and electrostatic components might offer useful information about the molecular mechanism involved in the antibody−antigen binding. In this study, we use a straightforward decomposition in FEP by collecting van der Waals and electrostatic interaction contributions separately, i.e., $V(\lambda) = V(\lambda)_{elec} + V(\lambda)_{vdW}$, in the same ensemble with full interactions in eq 2 (see the Method and System section). Due to the nonlinearity of the FEP formulation, there might be a small coupling term in this approach.[43] Again, using the forward T131I simulation as an example, out of 31.4 kcal/mol total bound state free energy change ΔG_1, 26.2 kcal/mol is from the electrostatic interactions, and 4.9 kcal/mol from van der Waals, and the coupling term is about 0.3 kcal/mol. Therefore, the electrostatic interactions dominate the free energy change. Similarly, out of the total 26.1 kcal/mol total free state free energy change ΔG_2, 22.6 kcal/mol is from the electrostatic interactions, 3.3 kcal/mol from van der Waals, and 0.2 kcal/mol from the coupling. In the final binding affinity change $\Delta\Delta G$ for the mutation T131I, the electrostatic interactions also dominate the contribution, with about 70% from electrostatic and 30% from van der Waals interactions. It should be noted that there are still controversies in the literature about the meaningfulness of breaking the total free energy into components.[43−46] Also, the free energy decomposition might be path-dependent, i.e., turning on van der Waals interactions first or electrostatic interactions first in some FEP approaches. Nevertheless, we find such decomposition offers useful insights into the various physical interactions involved in the HA−Ab molecular binding (more later). Another way to decompose free energy components is the linear response approach (LRA) by Warshel and co-workers.[47]

These representative free energy calculations also reveal that larger fluctuations in ΔG_1 and ΔG_2 can be seen in some λ windows (not all of them) but not always the same windows. A detailed convergence analysis shows that more runs, with 6.6 ns each, give a slightly better convergence than single longer runs (as shown in the final $\Delta\Delta G$ when compared to experiment). Therefore, in the following binding affinity calculations, we use many 6.6 ns shorter runs for $\Delta\Delta G$ average and standard deviations. At least five independent runs starting from different initial configurations are performed for average binding affinity and standard deviation calculations. The simulation length for each mutation is more than 66 ns (22 windows × 0.3 ns × 5+ runs × 2 states), and the total aggregate simulation time for this study is more than 2 μs, which is significantly longer than most FEP calculations currently reported in the literature[24−26,37,44,48] (larger window sizes and longer simulation durations have also been tested, and a good convergence has been found with the current settings). These extensive simulations show a binding affinity change of 5.20 ± 0.94 kcal/mol for the T131I mutation (see Table 2), in excellent agreement with the experiment.

Next, we investigate the underlying physicochemical factors causing this relatively large 5.2 kcal/mol binding affinity decrease. The previous experimental work referred to the burial of a carbonyl group in the LYS156 backbone in the loop of 155-161 for the possible explanation.[2] The experiment also found that the unbound (free) state, i.e., HA with the T131I

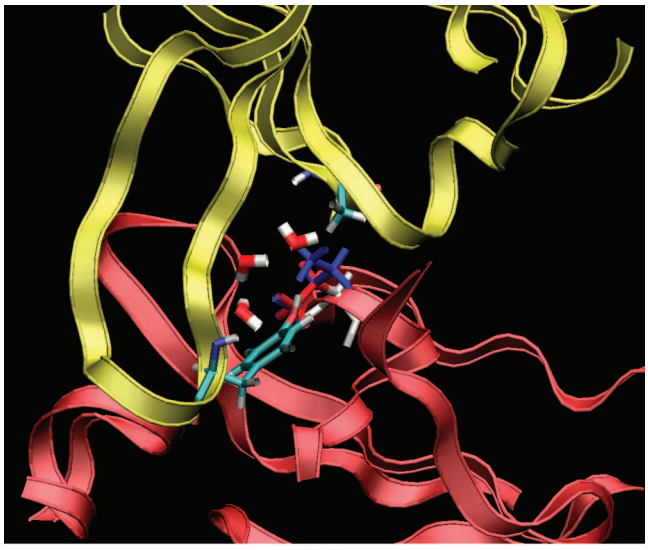


Figure 3. Bridging water between T131 with the antibody: (a) the initial bridging water; (b) the bridging water disappeared at the end of the T131I mutation. The bridge water and hydrogen bonding residues ALA53 and TYR107 from Fab are shown with sticks. The T131I residue is colored by dual topology (red, T131; blue, I131).

mutation but no antibody binding, shows a large conformational distortion in the loop 155-161 (backbone rmsd 1.1 Å).[2] However, our detailed trajectory analysis shows a slightly different molecular picture. We constantly find two or more bridging water molecules in between the wild-type HA and the antibody Fab near residue T131. These bridge water molecules form hydrogen bonds between the HA THR131 side chain's hydroxyl group (−OH) and Fab heavy chain residues, such as TYR107 (hydroxyl group in the side chain), SER31 (hydroxyl group and backbone), and ALA53 (backbone). These water molecules can also be viewed as lubricants between the HA and Fab interface; they are not necessarily fixed in space but rather fairly mobile with hydrogen bonds forming and breaking all the time. Upon the gradual mutation of the HA THR131 residue to ILE131 during FEP simulation, the much bulkier hydrophobic side chain of ILE131 displaces these two or three bridging water molecules, as shown in Figure 3. For further validation, we have repeated the same length simulation with the THR131 residue intact (i.e., no mutation) as a control run, and found about two bridge water molecules maintained during the entire simulation near residue THR131. On the other hand, we do not observe the induced burial of the carbonyl group of LYS156, as speculated in the experiment.[2] The carbonyl group of LYS156 makes hydrogen bonds with nearby water molecules during the entire simulation. Therefore, it seems that the displacement of bridge water molecules from the binding site has contributed to the loss of the binding affinity due to the T131I mutation but not the burial of the carbonyl group of LYS156. The displacement of bridging water molecules is also consistent with the earlier free energy decomposition results which indicate that the electrostatic interactions (hydrogen bonds with bridging waters) dominate the binding.

Interestingly, the unbound state conformational distortion in the loop of 155-161 due to the T131I mutation has also been observed in our simulation. Figure 4a shows the structural change of the 155-161 loop during the MD simulation (green, starting wild-type structure; red, final structure). In this comparison, we have aligned the initial and final structures by superimposing the nearby loop 127-132, same as in the previous experimental work[2] to show the conformational distortion of the loop 155-161 with regard to the 127-132 loop. Indeed, a noticeable distortion has been found for the loop 155-161, as shown in Figure 4a. This distortion has been observed throughout the simulation once the mutation is done (last stage in FEP λ space). The average backbone rmsd for the loop 155-161 in the last FEP stage is 1.41 Å (maximum 1.72 Å), when the other loop 127-132 is superimposed. Figure 4b further shows the comparison of the final simulated structure (in red) with the X-ray experimental structure for the mutated free state (pdb id 2VIU, in blue), where the simulated structure displays a slightly more distorted loop than the experimental structure. This can be further seen from the backbone rmsd for the loop 155-161, which is 1.1 Å in the experiment[2] but about 1.4 Å in our current simulation, as mentioned above.

Another interesting finding from the experiment is the movement of Fab's heavy chain ALA53 residue. This residue has been found to be slightly shifted after the mutation.[2] Figure 5 shows the movement for the ALA53 residue in Fab before and after the mutation, as observed during the FEP simulation. The mutation trajectory shows that the bulkier ILE131 side chain is shifted toward Fab's ALA53 residue due to the loss of the bridging waters and the "steric pushes" from nearby residues of Fab to accommodate the bulkier ILE, such as TYR107, ASN32, and SER31 in the heavy chain. This accommodation of ILE131 has resulted in a shift in ALA53's side chain. All of the trajectories show a similar effect. A closer look shows a movement of about 1.7 Å (rmsd) for this ALA53 residue when the initial and final antibody structures are aligned with all backbone atoms. There are no detailed X-ray structures given for the antibody before or after the mutation, so a direct comparison with experimental data is not feasible; however, from the description of this residue movement in the previous work, it seems consistent with our current findings.

The next interesting question thus becomes "Are there any other residues that T131 can mutate to, which will show an even stronger neutralization escape (larger binding affinity decrease)?" To address this question, we have run all of the possible mutations to other residues, except for proline (PRO), which has a special ring structure in the side chain and is currently not supported in the hybrid dual topology in NAMD2 (Not Another Molecular Dynamics Package).[31,36] A total of 14 neutral mutations as well as 4 charged ones (see below for

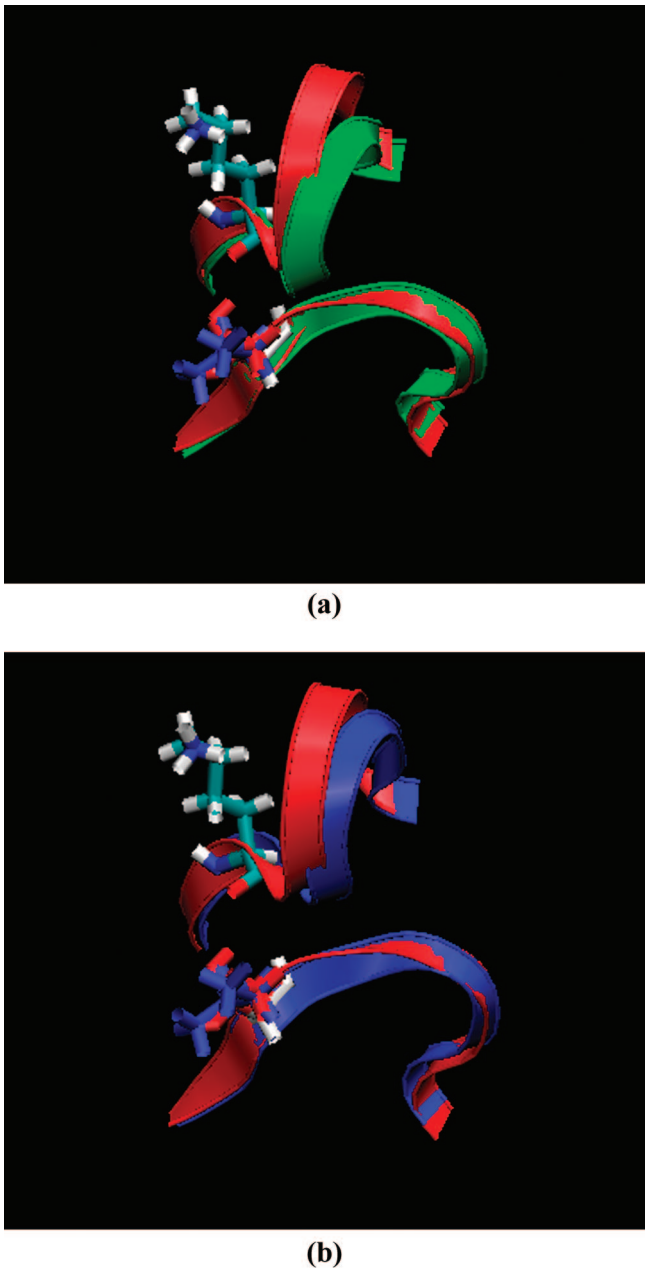


Figure 4. Hemagglutinin local loop structure in the free state due to T131I mutation: (a) structural change during the MD simulation (green, starting WT structure; red, final structure); (b) comparison with the X-ray experimental structure (blue, X-ray structure; red, structure from simulation).

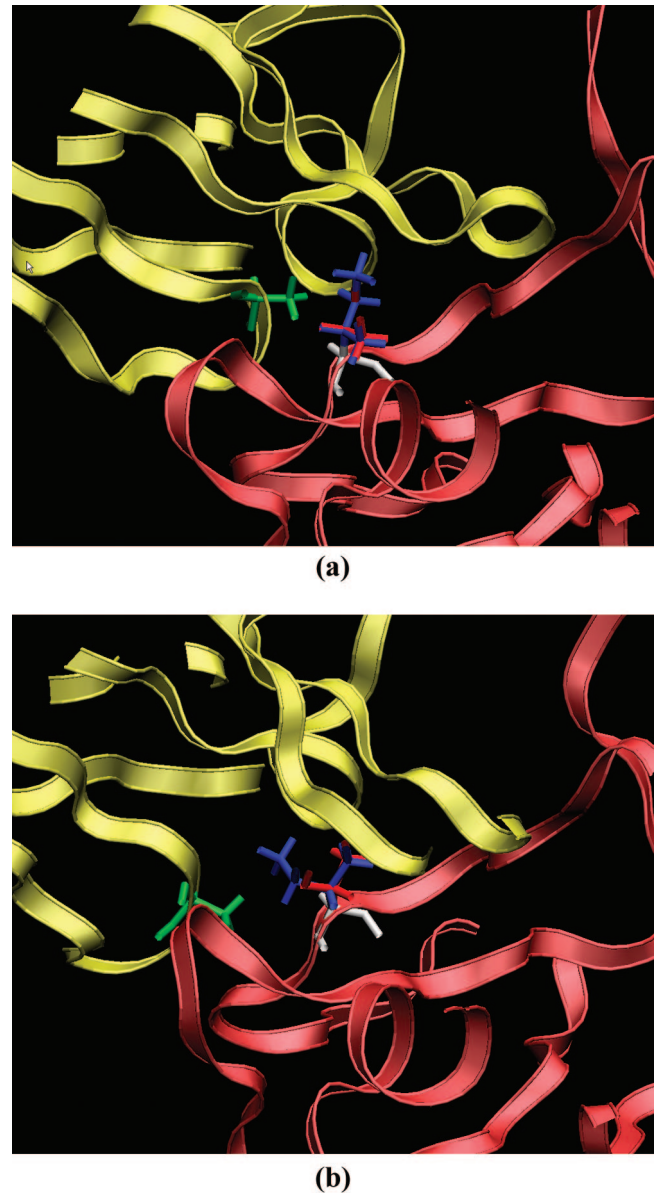


Figure 5. Shift of the antibody Fab heavy chain ALA53 residue due to the bulkier ILE131 side chain: (a) the initial postion in the wild type; (b) the shifted position of ALA53 after the mutation (due to the bulkier ILE131 mutant). The ALA53 is colored green, and T131I is colored by dual topology (red, T131; blue, I131).

charged mutations) are performed. Table 2 lists all of the binding affinity changes with standard deviations for these neutral residue mutations. Interestingly, the mutations T131W, T131Y, and T131F are found to have even larger binding affinity decreases, with $\Delta\Delta G$ values of 7.46 ± 1.91, 6.01 ± 1.31, and 5.68 ± 1.48 kcal/mol, respectively. A closer look at these mutations also reveals a significant displacement of bridging water molecules. The larger hydrophobic side chains such as TRP131 display the bound water molecules otherwise present in the wild-type HA/Fab interface. Many other residues also show a decrease in the binding affinity, such as T131H (3.84 kcal/mol), T131L (3.15 kcal/mol), T131N (2.92 kcal/mol), T131V (2.58 kcal/mol), and T131Q (1.22 kcal/mol), while a few others show an increase in the binding affinity (more negative), such as T131G (−3.72 kcal/mol), T131A (−2.81 kcal/mol), and T131S (−0.48 kcal/mol). It is interesting to notice that T131A mutation appeared in 1990 and circulated for some years, and then, it mutated back to THR131 (A131T) in 2004.[9] Apparently, the H3N2 virus somehow figured that the T131A mutation was not in its favor for circulation (due to antibody neutralization and/or other causes), which seems to be consistent with the current binding affinity predictions—the T131A mutation increases the binding affinity by 2.81 kcal/mol, thus enhancing the neutralization by the monoclonical antibody.

The FEP simulations for mutations to charged residues, on the other hand, were found to be much more challenging. Previous studies have pointed out that FEP calculations for neutral to charged (or charged to neutral) residues are nontrivial.[49,50] The problem arises largely from the fact that, in order to neutralize the system during the simulation under the periodic boundary condition, a counterion must be created (mutated) at the same time. Meanwhile, the free energy change of creation of a free ion, i.e., the ion solvation free energy, can be as large as −100 kcal/mol.[49,50] Thus, even a small error in

TABLE 3: Different Approaches for the FEP Mutation from Neutral to Charged Residues[a]

mutations	free ions	fixed ions	no ions	counter-mutations
T131K	-3.82 ± 7.36	3.22 ± 2.02	2.38 ± 1.07	2.01 ± 2.81
T131R	$2.72 \pm 8..97$	-0.54 ± 1.66	3.86 ± 1.37	2.20 ± 1.73
T131D	-2.93 ± 7.25	12.90 ± 1.92	12.69 ± 2.18	12.95 ± 2.34
T131E	-0.77 ± 9.57	6.11 ± 2.58	8.20 ± 1.95	11.71 ± 2.17

[a] Four different approaches are used for THR131 mutation to possible charged residues, i.e., T131K, T131R, T131D, and T131E.

the ion solvation free energy can cause a relatively large error in the HA/Fab binding free energy.[49−52] Here, we have used four different approaches to address this problem: (i) add a free counterion (mutating with the same λ); (ii) add a counterion but constrained in space (far away from the binding site); (iii) incorporate no counterions (i.e., "do nothing"); and (iv) counter-mutate another residue far away from the binding site. In approach iv, we chose a residue near the C-terminus, a polar residue TYR297, to reduce the possible effects on the HA/Ab binding. Table 3 lists all of the binding affinity results for the four different approaches. Our results show that approaches iii and iv behave most reliably. The addition of a free ion, even with its position fixed or constrained in space, creates more problems than fixing the "non-zero net-charge problem" with the periodic boundary condition. The "free ion" approach (method i) also shows a significantly larger standard deviation in the binding free energy due to the mobility of the free ion. It is somewhat surprising that approach iii, "doing nothing" (incorporating no counterions), behaves reasonably well. However, previous studies on the free energy of hydration of two triosephosphate isomerase inhibitors[53] also showed a similar conclusion. Approach iv, by counter-mutating another distant residue, seems to work best overall. The mutations to charged residues all seem to show a decrease in binding affinity, with the negatively charged residues (T131D and T131E) showing significantly more decrease. This is related to the fact that burying a charge inside a hydrophobic environment is very costly. The less unfavorable binding for the two basic residues (T131K and T131R) is because there is a nearby acidic residue ASP98 from the antibody heavy chain, within 7−8 Å of the mutation site, which contributes favorably to the electrostatic interactions with T131K and T131R (but contributes unfavorably for the T131D and T131E mutations, meanwhile). On the other hand, there is no basic residue from either the antibody heavy chain or light chain within 10 Å of the mutation site. It should be pointed out that Warshel and co-workers have also achieved a high convergence in FEP calculations for neutral to charged residue mutations by introducing physically consistent polarization constraints and local reaction field (LRF) long-range treatment for electrostatics[15,16,40−42] (interested readers are referred to the above references for more details).

Finally, as a further validation, we calculated the binding affinity for another mutation S157L, whose binding affinity change was also measured in the previous experiment.[2] For this mutation, the experiment reported an ~500 times increase in the dissociation constant K_d, indicating a binding affinity increase of ~3 kcal/mol. From the FEP simulations, we estimate a binding affinity change of 4.1 ± 1.0 kcal/mol upon the S157L mutation, again in excellent agreement with the experiment. We have not explored this mutation as extensively as the T131I mutation partly due to its smaller binding affinity change (thus not as effective in escaping the antibody neutralization) and partly due to the heavy computational resource requirements. As mentioned above, our simulation length is about 1 order of magnitude longer than most current state-of-the-art FEP calculations, and even with supercomputers like BlueGene/L, it is nontrivial to run these extensive simulations.

These results from current massively parallel FEP calculations provide some testing ground for further experimental validation, such as using X-ray crystallography to uncover key changes in the HA structure, and immunological assays to determine the antibody−antigen reactivity, due to these single mutations. The approach we have developed in this study for H3N2 can be readily applied to H5N1, once the H5N1 HA/Ab complex structures become available. The same methodology can also be used to assess the HA−receptor binding specificity switch due to single or double mutations.[3,4,7] Thus, these large scale computer simulations might provide a first step toward our ultimate goal that to keep ahead of the virus and offer surveillance in silico which can be correlated with in vitro as well as in vivo evolution.

Conclusions

In this paper, we have studied the single mutation effect on the antibody−antigen binding affinity of H3N2 hemagglutinin (HA) and a monoclonical antibody fragment (Fab) using the free energy perturbation (FEP) method. With an extensive FEP sampling (an aggregate of several microsecond molecular dynamics simulations), we have obtained the following major findings:

A single mutation T131I in H3N2 HA is found to cause a 5.2 ± 0.9 kcal/mol decrease in the HA/Ab binding affinity, which is in excellent agreement with the recent experiment (where the dissociation constant K_d increases by a factor of ~4000), but with a somewhat different molecular mechanism. Detailed analysis reveals that this large binding affinity decrease is largely due to the displacement of two bridge water molecules otherwise present in the wild-type HA/Fab interface but not so much due to the burial of the backbone carbonyl group of LYS156 in the loop 155-161 as suggested by the experiment. The lack of the net binding is also related to the relatively large structural changes in the free state (without Fab) of the mutant.

We then modeled all other possible mutations for this specific residue site T131 and predicted a few more mutations, such as T131W, T131Y, T131F, and charged ones like T131D and T131E, which might display an even greater capability in escaping the antibody neutralization. As for further validation, we also simulated another experimentally reported mutation S157L, for which the binding affinity results also agree with experiment quite well.

Finally, we examined in detail various FEP approaches for addressing the long-standing challenge in mutations involving net-charge changes, and found the approaches using counterions, either free or fixed in space, result in large errors due to the solvation free energy of the counterions. A better approach is to counter-mutate (with an opposite charge) a far-distant polar residue from the active site, or simply do nothing special as others had previously advocated.

Method and System

For the HA/Ab molecular system, we used the H3N2 HA/Fab complex (pdb id 2VIR) for the antigen/antibody binding, which has one monomer of HA with two subunits (HA1 and HA2) and one Fab antibody fragment with two chains (heavy chain and light chain). The HA/Fab complex is solvated in a 88 Å × 115 Å × 113 Å water box (see Figure 2b), with a total of ~34 000 water molecules. The counterions (2 Cl^-) are then added to neutralize the simulation system, which gives a total of 112 711 atoms. The complex system is then minimized by 5000 steps, followed by a two-stage, 10 ns each, equilibration. In the first stage, the protein backbones are constrained, and in the second stage, all atoms are relaxed. The unbound (free) state is modeled with the HA only, solvated in water, and equilibrated with a similar process. The particle-mesh Ewald (PME) method is used for the long-range electrostatic interactions.[54] All molecular dynamics (MD) simulations are performed with a specially optimized version of NAMD2[31,36] (for IBM BlueGene) using NPT ensemble at 1 atm pressure and 310 K temperature. The CHARMM22 force field[55] and TIP3P water model[56] are used.

When an antigenic variation occurs, the change in the HA binding affinity to a neutralizing antibody can be calculated by the free energy perturbation (FEP) method.[14−29] The Helmholtz free energy of a system can be expressed as

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

where Z is the partition function and $H(p,q)$ is the Hamiltonian of the system. The binding free energy change ΔG due to a mutation in hemagglutinin 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 $V(\lambda) = (1 - \lambda)V_1 + \lambda V_2$, V_1 represents the potential energy of the wild type, and V_2 represents the potential energy of the mutant. 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 $\langle ... \rangle_\lambda$ represents the ensemble average at potential $V(\lambda)$. In typical FEP calculations, for a single mutation from residue A to residue B, many perturbation windows have to be used in order to have a "smooth" transition from state A to B, with typically more windows near the two ends in order to enhance the sampling statistics.

In general, it is difficult and extremely time-consuming to directly calculate the binding affinity change ΔG_A (see Figure 1) for the binding process between a protein and an antibody, due to the complicated binding process. However, we can avoid this problem by designing a thermodynamical cycle in FEP calculations, as shown in Figure 1. In order to obtain the binding free energy change from residue A (wild type) to residue B (mutant), $\Delta\Delta G_{AB}$, we can instead calculate the difference in free energy changes for the bound state (ΔG_1) and the free state (ΔG_2) with the same mutation. Within a complete thermodynamical cycle, the total free energy change should be zero

$$\Delta G_A + \Delta G_1 - \Delta G_B - \Delta G_2 = 0 \quad (4)$$

which gives the binding affinity change due to the mutation from A to B as

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

In the current setup, a 22-window scheme has been adopted for both the bound and free states to achieve a reasonably high accuracy ($\lambda = 1.0^{-7}$, 1.0^{-6}, 1.0^{-5}, 1.0^{-4}, 1.0^{-3}, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.99, 0.999, 0.9999, 0.99999, 0.999999, 0.9999999). At least five independent runs starting from different initial configurations (picked evenly from the second stage of equilibration) are performed for average binding affinity and standard deviation calculations. The simulation length for each mutation is at least 66 ns (22 windows × 0.3 ns × 5+ runs × 2 states), and the total aggregate simulation time for this study is more than 2 μs, which is significantly longer than most FEP calculations currently reported in the literature.[24−26,37] In each FEP λ-window, the first 4000 steps are for further equilibration. Larger window sizes and longer simulation durations have also been tested, and we found that the current setup gives us a reasonable convergence in the final binding affinities.

Acknowledgment. We would like to thank Ian Wilson, Dennis Burton, Peter Palese, Jingyuan Li, Bruce Berne, and Isidore Rigoutsos for helpful discussions. We also 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.

Note Added after ASAP Publication. This paper was published ASAP on November 1, 2008. THE131 was changed to THR131 throughout the text. The revised paper was reposted on November 6, 2008.

References and Notes

(1) Stevens, J.; Blixt, O.; Glaser, L.; Taubenberger, J. K.; Palese, P.; Paulson, J. C.; Wilson, I. A. *J. Mol. Biol.* **2006**, *355*, 1143.
(2) Fleury, D.; Wharton, S. A.; Skehel, J. J.; Knossow, M.; Bizebard, T. *Nat. Struct. Biol.* **1998**, *5*, 119.
(3) Yang, Z. Y.; Wei, C. J.; Kong, W. P.; Wu, L.; Xu, L.; Smith, D. F.; Nabel, G. J. *Science* **2007**, *317*, 825.
(4) 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.
(5) Stevens, J.; Corper, A. L.; Basler, C. F.; Taubenberger, J. K.; Palese, P.; Wilson, I. A. *Science* **2004**, *303*, 1866.
(6) 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.
(7) Stevens, J.; Blixt, O.; Tumpey, T. M.; Taubenberger, J. K.; Paulson, J. C.; Wilson, I. A. *Science* **2006**, *312*, 404.
(8) 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.
(9) Shih, A. C.; Hsiao, T. C.; Ho, M. S.; Li, W. H. *Proc. Natl. Acad. Sci.* **2007**, *104*, 6283.
(10) Apisarnthanarak, A.; Erb, S.; Stephenson, I.; Katz, J. M.; Chittaganpitch, M.; Sangkitporn, S.; Kitphati, R.; Thawatsupha, P.; Waicharoen, S.; Pinitchai, U.; Apisarnthanarak, P.; Fraser, V. J.; Mundy, L. M. *Clin. Infect. Dis.* **2005**, *40*, e16.
(11) Sawai, T.; Itoh, Y.; Ozaki, H.; Isoda, N.; Okamoto, K.; Kashima, Y.; Kawaoka, Y.; Takeuchi, Y.; Kida, H.; Ogasawara, K. *Immunology* **2008**,
(12) Burton, D. R.; Williamson, R. A.; Parren, P. W. *Virology* **2000**, *270*, 1.

(13) Stephenson, I.; Das, R. G.; Wood, J. M.; Katz, J. M. *Vaccine* **2007**, *25*, 4056.
(14) Wang, W.; Donini, O.; Reyes, C. M.; Kollman, P. A. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 211.
(15) Warshel, A. *Specificity in Biological Interactions, Pontificiae Academiae Scientiarum Scripta Varia* **1984**, *55*, 60.
(16) Warshel, A.; Sharma, P. K.; Kato, M.; Parson, W. W. *Biochim. Biophys. Acta* **2006**, *1764*, 1647.
(17) Simonson, T.; Archontis, G.; Karplus, M. *Acc. Chem. Res.* **2002**, *35*, 430.
(18) Tembe, B. L.; McCammon, J. A. *Comput. Chem.* **1984**, *8*, 281.
(19) Jorgensen, W. L. *Acc. Chem. Res.* **1989**, *22*, 184.
(20) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395.
(21) Rami Reddy, M. M. D. E. A. A. Free Energy Calculations: Use and Limitations in Predicting Ligand Binding Affinities. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Ed.; 2007; p 217.
(22) Pathiaseril, A.; Woods, R. J. *J. Am. Chem. Soc.* **2000**, *122*, 331.
(23) Laederach, A.; Reilly, P. J. *Proteins: Struct., Funct., Bioinf.* **2005**, *60*, 591.
(24) Tirado-Rives, J.; Jorgensen, W. L. *J. Med. Chem.* **2006**, *49*, 5880.
(25) Deng, Y.; Roux, B. *J. Chem. Theory Comput.* **2006**, *2*, 1255.
(26) Jayachandran, G.; Shirts, M. R.; Park, S.; Pande, V. S. *J. Chem. Phys.* **2006**, *125*, 084901.
(27) Almlof, M.; Aqvist, J.; Smalas, A. O.; Brandsdal, B. O. *Biophys. J.* **2006**, *90*, 433.
(28) Brandsdal, B. O.; Osterberg, F.; Almlof, M.; Feierberg, I.; Luzhkov, V. B.; Aqvist, J. *Adv. Protein Chem.* **2003**, *66*, 123.
(29) Brandsdal, B. O.; Smalas, A. O. *Protein Eng.* **2000**, *13*, 239.
(30) Almlof, M.; Aqvist, J.; Smalas, A. O.; Brandsdal, B. O. *Biophys. J.* **2006**, *90*, 433.
(31) Kumar, S.; Huang, C.; Almasi, G.; Kale, L. V. "Achieving Strong Scaling with NAMD on Blue Gene/L"; *Proceedings of IEEE International Parallel and Distributed Processing Symposium*, 2006.
(32) Eleftheriou, M.; Germain, R. S.; Royyuru, A. K.; Zhou, R. *J. Am. Chem. Soc.* **2006**, *128*, 13388.
(33) Liu, P.; Huang, X.; Zhou, R.; Berne, B. J. *Nature* **2005**, *437*, 159.
(34) Zhou, R.; Eleftheriou, M.; Royyuru, A. K.; Berne, B. J. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 5824.
(35) Zhou, R.; Huang, X.; Margulis, C. J.; Berne, B. J. *Science* **2004**, *305*, 1605.
(36) 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.
(37) Thrope, I. F.; Brooks, C. L. *Proc. Natl. Acad. Sci.* **2007**, *104*, 8821.
(38) Pan, Y.; Gao, D.; Yang, W.; Cho, H.; Zhan, C. G. *J. Am. Chem. Soc.* **2007**, *129*, 13537.
(39) Yang, W.; Bitetti-Putzer, R.; Karplus, M. *J. Chem. Phys.* **2004**, *120*, 2618.
(40) Warshel, A.; Sussman, F.; King, G. *Biochemistry* **1986**, *25*, 8368.
(41) Aqvist, J.; Warshel, A. *Biophys. J.* **1989**, *56*, 171.
(42) Warshel, A.; Hwang, J. K.; Aqvist, J. *Faraday Discuss.* **1992**, *93*, 225.
(43) Bren, M.; Florian, J.; Mavri, J.; Bren, U. *Theor. Chem. Acc.* **2007**, *117*, 535.
(44) Mark, A. E.; Vangunsteren, W. F. *J. Mol. Biol.* **1994**, *240*, 167.
(45) Boressch, S.; Karplus, M. *J. Mol. Biol.* **1995**, *254*, 801.
(46) Brady, G. P.; Sharp, K. A. *J. Mol. Biol.* **1995**, *254*, 77.
(47) Warshel, A.; Sharma, P. K.; Chu, Z. T.; Aqvist, J. *Biochemistry* **2007**, *46*, 1466.
(48) Pathiaseril, A.; Woods, R. J. *J. Am. Chem. Soc.* **2000**, *122*, 331.
(49) Kastenholz, M. A.; Hunenberger, P. H. *J. Chem. Phys.* **2006**, *124*, 224501.
(50) Kastenholz, M. A.; Hunenberger, P. H. *J. Chem. Phys.* **2006**, *124*, 124106.
(51) Koculi, E.; Hyeon, C.; Thirumalai, D.; Woodson, S. A. *J. Am. Chem. Soc.* **2007**, *129*, 2676.
(52) Vaitheeswaran, S.; Thirumalai, D. *J. Am. Chem. Soc.* **2006**, *128*, 13490.
(53) Donnini, S.; Mark, A. E.; Juffer, A. H.; Villa, A. *J. Comput. Chem.* **2005**, *26*, 115.
(54) Darden, T. A.; York, D. M.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *98*, 10089.
(55) MacKerell, A. D.; Bashford, D.; Bellott, M.; Evanseck, R. L. D. J.; 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.
(56) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

JP805529Z