Hydrogenase/Ferredoxin Charge-Transfer Complexes: Effect of Hydrogenase Mutations on the Complex Association†

Hai Long, Paul W. King, Maria L. Ghirardi, and Kwiseon Kim*

National Renewable Energy Laboratory, 1617 Cole Blvd., Golden, Colorado 80401

Received: November 26, 2008; Revised Manuscript Received: February 18, 2009

The [FeFe]-hydrogenases in the green alga *Chlamydomonas reinhardtii* utilize photogenerated electrons to reduce protons into hydrogen gas. The electrons are supplied from photosystem I and transferred to the [FeFe]-hydrogenase through specific hydrogenase−ferredoxin association. To understand how structural and kinetic factors control the association better, we used Brownian dynamics simulation methods to simulate the charge-transfer complex formation between both native and in silico mutants of the [FeFe]-hydrogenase HYDA2 and the [2Fe2S]-ferredoxin FDX1 from *C. reinhardtii*. The changes in binding free energy between different HYDA2 mutants and the native FDX1 were calculated by the free-energy perturbation method. Within the limits of our current models, we found that two HYDA2 mutations, $T99K_H$ and $D102K_H$, led to lower binding free energies and higher association rate with FDX1 and are thus promising targets for improving hydrogen production rates in engineered organisms.

Introduction

A hydrogen economy has been proposed as an alternative to today's hydrocarbon economy, but most of the hydrogen gas (H_2) currently produced in the United States is obtained by nonrenewable steam reforming of hydrocarbons.[1] Several renewable methods for the production of H_2 have been proposed, one of which is photobiological H_2 production using the green alga *Chlamydomonas reinhardtii*.[2−7] The H_2 production process in *C. reinhardtii* is highly sensitive to O_2,[8] an obligatory byproduct of photosynthesis. However, by attenuating the O_2-evolving activity of photosystem II (PSII) through sulfur deprivation or reduced gene expression, cellular respiration rapidly consumes the photogenerated O_2 and creates an anaerobic environment that is suitable for H_2 production in sealed cultures.[9−14] Under these conditions, electrons from the anaerobic oxidation of starch, the water oxidation by residual PSII activity, or both are fed into the plastoquinone pool and donated to photosystem I (PSI) where photoexcitation and charge separation lead to the reduction of the electron-carrier ferredoxin.[15−17] This reduced ferredoxin pool is able to provide low-potential electrons to [FeFe]-hydrogenase for catalytic H_2 production.[18,19]

C. reinhardtii biosynthesizes two nuclear-encoded [FeFe]-hydrogenases, HYDA1 and HYDA2, from the nuclear-encoded genes *HYDA1* and *HYDA2*, respectively.[20] The nuclear genome of *C. reinhardtii* also encodes six ferredoxin genes (*FDX1*−*FDX6*).[21] The physiological regulation and metabolic function of each of the six ferredoxins has not yet been established, although it is assumed that both hydrogenases may associate with the *FDX1* gene product. Purification of *C. reinhardtii* hydrogenase and ferredoxin led to an estimated Michaelis−Menten constant (K_M) of 10−35 μM[18,19] and a k_{cat} for H_2 production from reduced ferredoxin of 600 s^{-1}.[19] When an electron is transported from PSI to ferredoxin, ferredoxin distributes electrons to not only hydrogenase but also other electron-dependent enzymes, for example, ferredoxin-$NADP^+$ reductase (FNR)[22−24] for CO_2 fixation. A more detailed understanding of the hydrogenase−ferredoxin association mechanism and kinetics is necessary for engineering hydrogenase to divert a greater electron flux to the H_2 production pathway to optimize photobiological H_2 production.

In previous reports, we modeled the structures of the *C. reinhardtii* [FeFe]-hydrogenase (HYDA2)[20] and ferredoxin (PETF1, currently known as FDX1)[25] and obtained two candidate binding-complex structures,[25] which are denoted as complexes 16 and 42 (Figure 1, coordinates attached in the Supporting Information). In both complexes, the HYDA2 catalytic site, which lies near a positively charged surface, was oriented toward the negatively charged surface and the $[2Fe2S]_F$[26] cluster of FDX1. The main difference between complexes 16 and 42 is the orientation of FDX1 with respect to HYDA2. The two complex structures are related by a ∼180° rotation of FDX1 around the axis formed by the $[4Fe4S]_H$ subcluster of HYDA2 and the $[2Fe2S]_F$ cluster of FDX1. HYDA2−FDX1 binding free-energy calculations and encounter-complex studies identified complex 16 as having the lowest binding free energy[25] and encounter-complex free energy,[27] implying that it is most likely to represent an in vivo interaction configuration. In this study, we used Brownian dynamics (BD) to simulate the association kinetics of native and mutant HYDA2 with FDX1. Mutations of HYDA2 were selected on the basis of the potential to enhance complex formation with FDX1, created in silico and investigated using BD simulations. Finally, the binding free-energy changes of HYDA2 mutants with native FDX1 were calculated by the free-energy perturbation (FEP) method on the basis of molecular dynamics (MD) simulations.

Methods

Reference Structures. BD and MD simulations were conducted using HYDA2[20] and FDX1[25] homology structures and the structures of HYDA2−FDX1 binding complexes 16 and 42.[25] We simulated four systems using BD: (i) complex 16_R, using the atomic coordinates of HYDA2 and FDX1 from the energy-minimized structure of binding complex 16 where FDX1 is modeled in the reduced state; (ii) complex 16_O, based on the

† Part of the special issue "George C. Schatz Festschrift".

* Corresponding author. E-mail: kwiseon_kim@nrel.gov.

10.1021/jp810409z CCC: $40.75
Published on Web 03/24/2009

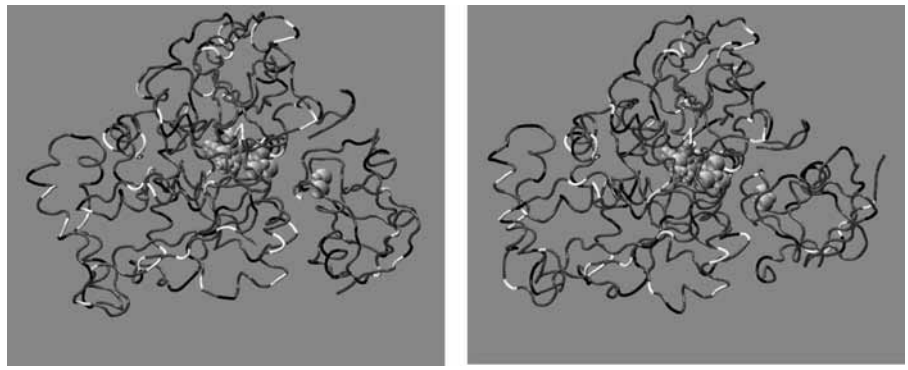


Figure 1. Structures of complex 16 (left) and complex 42 (right). For each structure, the protein on the left is hydrogenase HYDA2 and the one on the right is ferredoxin FDX1. The metalloclusters shown are the $[2Fe2S]_F$ cluster of FDX1 and the H-cluster of HYDA2, which includes a $[4Fe4S]_H$ subcluster and a $[2Fe]_H$ subunit. Positively charged residues on both proteins are shown in white, and negatively charged residues are shown in black.

complex 16 structure with an oxidized FDX1; (iii) complex 42_R, based on the complex 42 structure with a reduced FDX1; and (iv) complex 42_O, based on the complex 42 structure with an oxidized FDX1. HYDA2 was kept in the oxidized state in all systems. The CHARMM22 atomic charges and atom radii[28] were used for the amino acid residues of HYDA2 and FDX1 during the electrostatic potential grid calculations and MD simulations. The partial atomic charges and radii for the metalloclusters were derived from geometry-optimized model clusters, using a BLYP/6-31+G(d) model and the natural population analysis charges method[29] (manuscript in review). In the reduced state of FDX1, the $[2Fe2S]_F$ cluster, along with the four cysteine sulfur atoms ($Cys39_F$, $Cys44_F$, $Cys47_F$, and $Cys77_F$) attached to it, has a total partial charge of $-3e$, whereas in the oxidized state, the total partial charge is $-2e$. The conformations of the reduced and oxidized FDX1 are identical, except for the peptide bond conformation between residues $Cys44_F$ and $Ser45_F$.[27,30]

In Silico Mutagenesis. To identify candidate HYDA2 mutation sites with the potential to affect the HYDA2−FDX1 association, the residues on the HYDA2 binding face were inspected by computational alanine scanning using the Robetta web server.[31,32] Alanine scanning systematically mutates each target residue to alanine, and the change of binding free energy, $\Delta\Delta G$, of the association with FDX1 is calculated. A negative $\Delta\Delta G$ implies that the specific mutation decreases the binding free energy and stabilizes the binding complex, indicating that the original residue is less favorable for binding. On the other hand, a positive $\Delta\Delta G$ shows that the original residue is energetically more favorable for establishing a binding complex. Residues with $\Delta\Delta G > 1$ kcal/mol are called "hotspots" and are considered to be key binding sites; residues with $0 < \Delta\Delta G < 1$ kcal/mol are called "warm" residues and contribute only moderately to binding; and residues with $\Delta\Delta G < 0$ are the unfavorable residues for binding.

The four potential mutation sites on HYDA2 identified by alanine scanning are listed in Table 2: $T99_H$, $D102_H$, $M214_H$, and $E221_H$. Three of them are listed in Table 2. The kinetics and thermodynamics of mutant HYDA2−FDX1 complexes were investigated by MD and BD simulations. FEP-MD simulations of HYDA2, complex 16_R, and complex 42_R were performed for each of the four HYDA2 mutants, $T99K_H$, $D102K_H$, $M214K_H$, and $E221K_H$. BD simulations included the same four mutants and two additional double mutants, $D102K_H$/$E221K_H$ and $D102K_H$/$T99K_H$, incorporated into each of the four simulation systems: 16_R, 16_O, 42_R, and 42_O. The mutant structures were generated using the program psfgen included in the NAMD 2.6 package[33] and were then energy minimized for 1000 steps using the conjugate gradient method with NAMD. During the minimization, the coordinates of the unchanged residues were fixed to relieve any local stresses introduced by the mutation and to prevent altering the global structure of HYDA2. The minimized structures were then subjected to BD or MD simulation runs.

TABLE 1: Brownian Dynamics Simulation Parameters

	$16_{R/O}$	$42_{R/O}$
number of donor−acceptor pairs[a]	36 (8)	43 (6)
relative translational diffusion constant (Å^2/ps)	2.160×10^{-2}	2.165×10^{-2}
rotational diffusion constant of FDX1 (rad^2/ps)	3.016×10^{-5}	2.937×10^{-5}
rotational diffusion constant of HYDA2 (rad^2/ps)	5.918×10^{-6}	6.289×10^{-6}

[a] Number of intermolecular donor−acceptor pairs (not necessarily forming hydrogen bonds) within 5 Å in the homology models of the complexes. The number of independent donor−acceptor pairs is given in parentheses.

Brownian Dynamics Simulations. The simulation of diffusional association (SDA) package[34] was used to perform the BD simulations. HYDA2 and FDX1 were treated as rigid bodies in the BD simulations, with the center of mass (COM) of HYDA2 fixed at the center of the simulation sphere. Only rotational movement of HYDA2 was allowed, whereas FDX1 could move stochastically around HYDA2 both translationally and rotationally. The BD trajectory propagates by solving the diffusion equations

$$\Delta\vec{r}(t) = \frac{D\Delta t}{k_B T}\vec{F}(t) + \vec{R}(t) \quad (1)$$

where D is the translational (or rotational) diffusion constant, Δt is the time step, $\vec{F}(t)$ is the instantaneous systematic potential force (or torque), k_B and T are the Boltzmann constant and temperature, respectively ($T = 298$ K in this work), and $\vec{R}(t)$ is the random displacement due to the collisions between the protein and solvent molecules. The values of $\vec{R}(t)$ satisfy the boundary conditions

$$<\vec{R}(t)> = 0 \quad \text{and} \quad <\vec{R}(t)\cdot\vec{R}(0)> = 6D\Delta t \quad (2)$$

TABLE 2: Warm or Unfavorable Residues on the HYDA2 Binding Face in Complexes 16 and 42

	ΔΔG (kcal/mol)		
residue[a]	complex 16[b]	complex 42[b]	note
$T99_H$	**0.18**	**0.07**	
$D102_H$	**N/A**	**0.53**	
$V104_H$	0.28	1.07	"hotspot" in complex 42
$K108_H$	−0.05	0.09	positively charged residue
$M214_H$	**0.10**	**0.38**	
$S218_H$	0.53	1.34	"hotspot" in complex 42
$Y219_H$	N/A	0.56	salt bridge with $K442_H$
$V264_H$	0.37	0.46	close to the $[4Fe4S]_H$ cluster
$R265_H$	0.66	0.97	positively charged residue
$L438_H$	0.80	1.05	"hotspot" in complex 42
$K443_H$	0.37	N/A	positively charged residue
$K447_H$	0.51	N/A	positively charged residue
$C463_H$	−0.02	−0.05	attached to the $[4Fe4S]_H$ cluster
$K479_H$	N/A	0.06	positively charged residue

[a] Residues in bold are the ones selected as potential mutation sites. [b] N/A indicates that the residue does not appear on the HYDA2 binding face in this complex.

The translational and rotational diffusion constants for HYDA2 and FDX1, calculated by the HYDROPRO program,[35] are presented in Table 1. The BD simulation time step, Δt, was chosen to be 1.0 ps when the COM−COM distance of the two proteins was within 80 Å. At larger separation distances, Δt was increased linearly with a slope of 0.5 ps/Å.

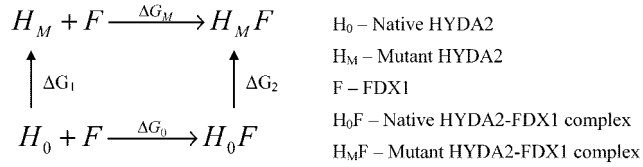
During the BD simulations, the potential force (or torque) $\vec{F}(t)$ on protein 1 due to the presence of protein 2 was computed using the effective charges on protein 1 and the electrostatic potential grid and charge desolvation penalty grid around protein 2. The effective charges on each protein were calculated by the ECM (effective charges for macromolecules in solvent) module[36] in the SDA package. The electrostatic potential grid around each protein was calculated using the adaptive Poisson−Boltzmann solver (APBS) program[37] to solve the full Poisson−Boltzmann equation at the given ionic strength. The charge desolvation penalty grids were calculated using the mk_ds_grd program in the SDA package[34] with a scaling factor $\alpha = 1.67$. For a more detailed description, please refer to the method section of our previous report.[27]

The BD simulations started at $b = 100$ Å of COM−COM distance between HYDA2 and FDX1, and they were terminated if the two proteins achieved a COM−COM distance of $c > b$. To better estimate the protein association rates, the cutoff boundary c should be large enough so that electrostatic interactions between the proteins can be neglected. We chose $c = 5b = 500$ Å, which equals 36 times the Debye length at 298 K and 50 mM ionic strength. This distance is large enough to screen out electrostatic interactions between HYDA2 and FDX1. BD simulations were performed at 50, 150, 300, and 500 mM and infinite ionic strengths for native HYDA2−FDX1 associations and at 150 mM ionic strength for the associations of HYDA2 mutants and FDX1. The infinite ionic strength was achieved by artificially setting the protein effective charges as well as the grid point values of the electrostatic potential grids and charge desolvation penalty grids to zero. A total of 10 000 trajectories were generated at each ionic strength.

To calculate association rates, distances between the hydrogen bond donors and acceptors on the binding faces of HYDA2 and FDX1 were monitored during the BD simulations. The choice of the hydrogen bond donors and acceptors and the atom contact criteria is similar to those used by Gabdoulline et al.[38] A donor−acceptor pair separated from other pairs by at least $d_{min} = 5$ Å was considered to be an independent contact.

SCHEME 1: Scheme Used for Calculating Binding Free Energy Change

Thermodynamic Cycle and Free-Energy Perturbation



Thermodynamic Cycle and Free-Energy Perturbation. We used the FEP method to calculate the binding free-energy changes for the mutated HYDA2 on the basis of the thermodynamic cycle, as shown in Scheme 1. The free-energy changes ΔG_1 and ΔG_2 were estimated using the FEP method, and the binding free-energy change, $\Delta\Delta G$, was obtained using eq 3

$$\Delta\Delta G = \Delta G_M - \Delta G_0 = \Delta G_2 - \Delta G_1 \quad (3)$$

During the FEP calculations, the native state is denoted as $\lambda = 0$, and the mutated state is denoted as $\lambda = 1$, where λ is the coupling parameter linked to the alchemical disappearance of the original amino acid residue atoms and the concomitant appearance of the mutation. The two states are thus connected by a series of intermediate states. A total of 13 intermediate states, $\lambda = 0.01$, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, and 0.99, were used, and the total free-energy change was calculated by adding up the free-energy changes obtained at each integration step. To avoid artifacts due to the sudden appearance or disappearance of an atom at end points, λ changes were small at both the beginning and end of the calculations. In addition, the square of the interatomic distance of FEP atoms was shifted by 5 Å^2 for the van der Waals (vdW) interaction calculation to ensure that the vdW potential was finite near the FEP end points. Finally, the vdW potential and the short-range electrostatic potential were decoupled for FEP atoms. The vdW potential FEP was fully grown by $\lambda = 0.5$, whereas the electrostatic potential FEP was turned off during $\lambda = 0$ to 0.5. The electrostatic potential was then turned on and began to grow linearly. In this way, the FEP atoms had enough vdW potential force to repel other FEP atoms before turning on the electrostatic potential FEP calculation. These FEP features were added in the CVS version of NAMD.[39]

The HYDA2 and FDX1 proteins are highly charged. Normally, counterions are required to neutralize simulation systems for MD simulations. However, it was reported that the calculated free-energy change by the FEP method was highly sensitive to the presence of counterions.[40] This problem could be avoided by running lengthy FEP simulations (20−100 ns per λ) to obtain a better sampling of the ions, but we considered such a time-consuming method to be beyond the focus of this study. More importantly, our interest is to determine $\Delta\Delta G$, not ΔG. The systematic errors arising in the FEP calculations are expected to cancel out when using eq 3. Therefore, no counterions were added, and the simulation systems were neutralized by a uniform charge atmosphere throughout the simulation box.

Before the FEP calculation, the native HYDA2 (or the native HYDA2−FDX1 binding complexes 16_R and 42_R) was solvated in a TIP3P water box. The water box was made to be large enough so that the protein atoms were at least 10 Å away from the box surfaces. The MD simulations were performed using NAMD[33] with the CHARMM22 protein force field.[28] The time step used in the simulations was 2 fs. Electrostatic energies were

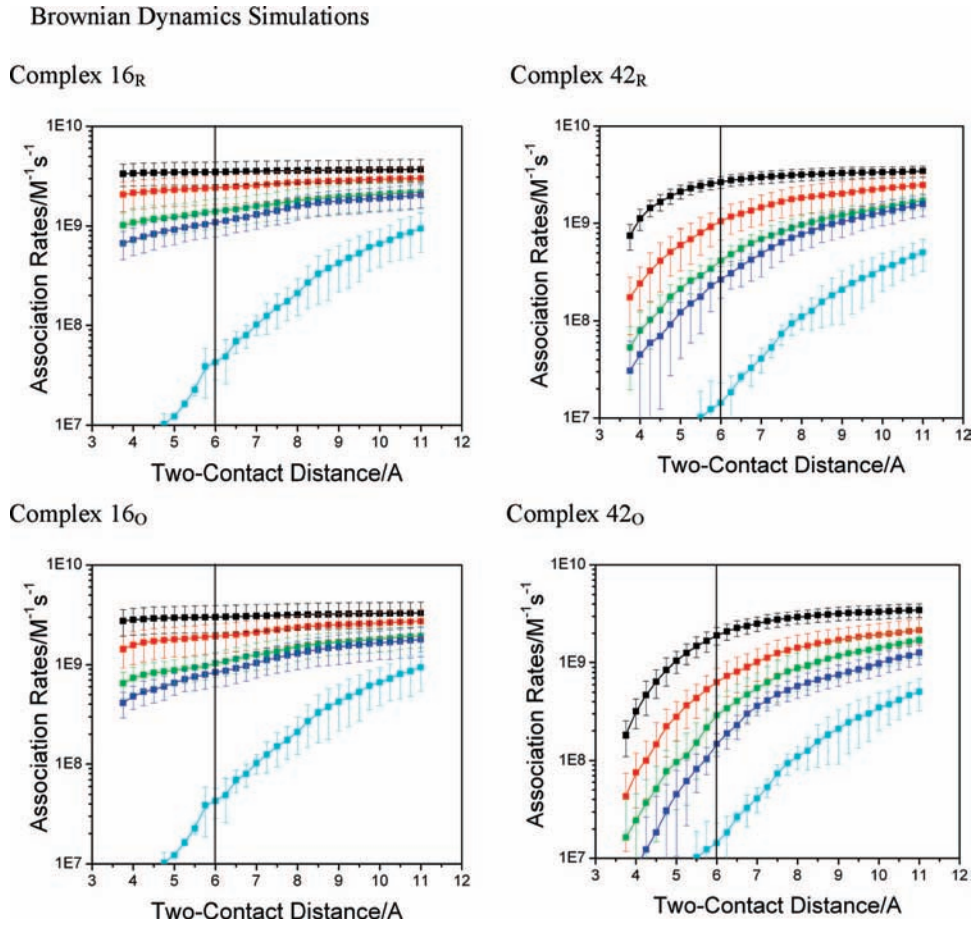


Figure 2. Association rates as a function of the two-contact distance at 50 mM (black curve), 150 mM (red), 300 mM (green), 500 mM (blue), and infinite (cyan) ionic strengths.

calculated by the particle mesh ewald (PME) summation method with ~1 Å mesh density. The simulation was performed in the NPT ensemble with temperature coupled to a 298 K Langevin bath and pressure maintained at 1 atm with a Langevin piston.[41,42] The simulated system was first minimized for 1000 steps using the conjugate gradient method. It was then equilibrated for 25 ps with all protein atoms fixed, followed by another 1000-step conjugate gradient minimization. The system was equilibrated for another 25 ps with only the protein backbone atoms fixed, followed by another minimization. Finally, the whole system was allowed to move and equilibrate for 3 ns. The equilibrated protein was then used for the FEP mutation simulations. In each FEP step, we equilibrated for 50 ps and then sampled data for 150 ps to calculate the free-energy change. Two FEP calculations were performed for each mutant: in the forward run, λ growing from 0 to 1; in the backward run, λ from 1 to 0. Therefore, we simulated 5.6 ns of dynamics for each mutant.

Results

Brownian Dynamics Simulations. On the basis of the BD simulation trajectories, plots of association rates as a function of "two-contact distance" at different ionic strengths are calculated and presented in Figure 2. The two-contact distance means that there are two pairs of independent hydrogen-bond donors and acceptors on the binding faces within a given distance.[34] For all simulation systems, the association rates increase when the two-contact distance increases. When the two-contact distance between the proteins is less than a critical distance, d_C, the two proteins form a so-called "diffusional encounter complex,"[38] and the proteins can reorient themselves rapidly and form the bound complex. The formation of a diffusional encounter complex is usually the rate-limiting step for a protein−protein association process. Therefore, the association rate at d_C can be treated as the diffusion-controlled protein association rate constant, k_1. The distance d_C is usually obtained by comparing the experimentally measured k_1 with the association rates at different two-contact distances computed by BD simulations. By fitting the computed rates and measured rates of a series of protein association complexes, Gabdoulline et al. concluded that the optimal value of d_C is 6 Å for the two-contact distance.[38] Because the measured k_1 values for the *C. reinhardtii* ferredoxin−hydrogenase system are still unavailable, we will adopt $d_C = 6$ Å in this study as the criterion to estimate k_1.

The computed k_1 values at different ionic strengths are plotted in Figure 3. The magnitudes of k_1 are 10^8-10^9, which is typical for protein associations.[38,43] The $16_{R/O}$ complexes have higher k_1 values compared with the $42_{R/O}$ complexes when simulated at the same ionic strength. For all of the simulations, k_1 decreases when the ionic strength increases, which is similar to the BD results from other protein association pairs.[38] In addition, k_1 at infinite ionic strength is one to two orders of magnitude lower than k_1 at 500 mM ionic strength. This suggests that a major driving force for the HYDA2−FDX1 association process is the electrostatic interaction between the binding faces. Although the total charges of HYDA2 and FDX1 are both very negative ($-17e$ and $-11e$, respectively), a majority of the charged

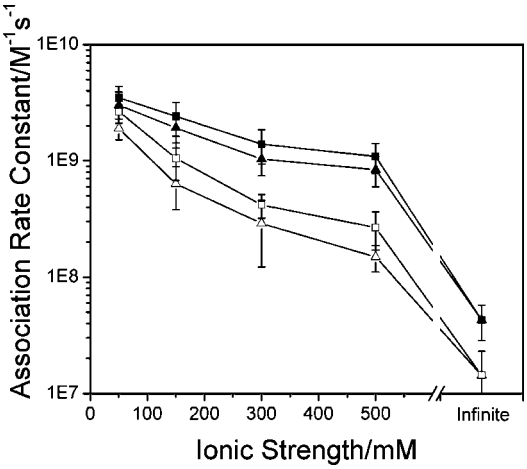


Figure 3. Association rate constant, k_1, as a function of ionic strength: ■, complex 16_R; ▲, complex 16_O; □, complex 42_R; △, complex 42_O.

residues on the binding face of HYDA2 are positive, and most of the charged residues on the binding face of FDX1 are negative. Consequently, the positively charged binding face of HYDA2 and the negatively charged binding face of FDX1 attract each other, which facilitates an orientated association that is conducive for charge-transfer and catalysis. Weaker electrostatic interactions at higher ionic strengths reduce k_1. In addition, increasing the ionic strength influences k_1 values of $16_{R/O}$ complexes less than k_1 values of $42_{R/O}$ complexes. The higher k_1 values of complexes $16_{R/O}$ indicate that they may have better diffusional encounter complex orientation than complexes $42_{R/O}$ do at d_C, thus resulting in shorter distances between the interacting charges on the binding faces. The k_1 values of $16_{R/O}$ complexes weakly depend on the ionic strength due to the smaller ionic screening between the closer interacting charges. Because reduced FDX1 has one more negative charge on the $[2Fe2S]_F$ cluster than the oxidized one does, reduced FDX1 has a greater attraction to the positively charged HYDA2 binding face. Therefore, the reduced complexes 16_R and 42_R have higher k_1 values than the oxidized complexes 16_O and 42_O, respectively, under the same conditions.

Mutation Site Identification. The "warm" or unfavorable residues on the HYDA2 binding face identified by the Robetta alanine scanning server are listed in Table 2. These residues were examined for mutations that might improve the HYDA2−FDX1 association. To do that, we initially did not mutate "hotspot" residues in either complex structure; those residues are $V104_H$, $S218_H$, and $L438_H$. We then found three "warm" residues ($Y219_H$, $V264_H$, and $C463_H$) that might be significant for HYDA2. $Y219_H$ forms a salt bridge with $K442_H$, which may be important for maintaining HYDA2 structure. $V264_H$ and $C463_H$ are closely located or attached to the $[4Fe4S]_H$ subcluster of HYDA2, which is the electron acceptor site of HYDA2. Therefore, mutations of those residues could dramatically affect the enzymatic activity, and thus they were excluded from the mutation study. Moreover, we excluded all positively charged residues from our study. Because the binding face of HYDA2 is positively charged and the binding face of FDX1 is negatively charged, a more positively charged HYDA2 binding face might lead to improved association kinetics. Therefore, positively charged residues identified by alanine scanning results, $K108_H$, $R265_H$, $K443_H$, $K447_H$, and $K479_H$ were left unchanged. This resulted in the selection of three sites for the mutation studies: $T99_H$, $D102_H$, and $M214_H$. To improve association, each site was mutated to the positively charged residue, lysine (K), because between the two positively charged amino acids, lysine and arginine (R), lysine's side chain is less bulky and is expected to have smaller perturbations on the HYDA2 structure.

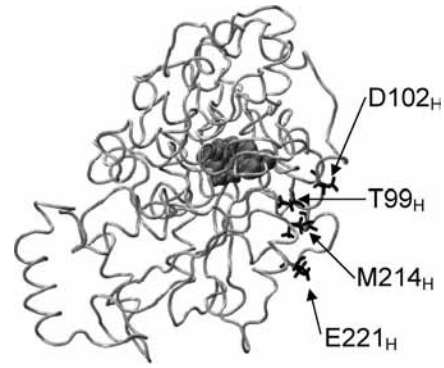


Figure 4. Potential mutation sites identified on HYDA2.

In addition to the residues identified by alanine scanning, the negatively charged residues away from the HYDA2 binding face might limit protein association because of long-range electrostatic repulsion interactions with the negatively charged FDX1 binding face. The electrostatic interaction between two unit charges becomes significant when their separation distance is smaller than the Bjerrum length l_B

$$l_B = \frac{e^2}{\varepsilon k_B T} \tag{4}$$

where e is the elementary charge and ε is the dielectric constant. In an aqueous solution at $T = 298$ K, $\varepsilon = 78.4$ and the calculated value of l_B is 7.2 Å. This means that even though a charged residue is not on the binding face, it can have considerable contribution to the association via electrostatic interactions if it is within 7.2 Å from the binding face. We found three negatively charged residues ($E216_H$, $E221_H$, and $D478_H$) on HYDA2 within 7.2 Å of the HYDA2−FDX1 binding face. Because residues $E216_H$ and $D478_H$ are expected to form an intramolecular salt bridge with $R272_H$ and $K479_H$, respectively, only $E221_H$ was selected for in silico mutation to lysine.

In summary, four potential mutation sites on HYDA2 (Figure 4) were identified and mutated into lysine in silico. The resulting change in total charge (Δe) for each HYDA2 mutant was: $T99K_H$, $+1e$; $D102K_H$, $+2e$; $M214K_H$, $+1e$; and $E221K_H$, $+2e$.

Mutation Study. Table 3 lists the calculated association rate constants, k_1, between HYDA2 mutants and native FDX1 in either $16_{R/O}$ or $42_{R/O}$ on the basis of the BD simulation trajectories at an ionic strength of 150 mM. A d_C value of 6 Å was used as the criterion to estimate k_1 as in the studies of complexes containing native HYDA2. The HYDA2 mutant $D102K_H$ enhanced k_1 the most among all single mutant systems simulated, whereas $M214K_H$ has the least enhancement. The large enhancement in k_1, observed with the $D102K_H$ mutant, may be due to the greater increase in charge ($+2e$) and to the residue's closeness to the HYDA2 binding face. The mutant $M214K_H$ modestly enhanced k_1 values for complexes $42_{R/O}$ but did not affect k_1 values for complexes $16_{R/O}$, presumably because $M214_H$ is close to the positively charged $R40_F$ in complexes $16_{R/O}$. Two additional mutants, $E221K_H$ and $T99K_H$, showed comparable enhancements in k_1 for complexes $16_{R/O}$. However, $E221K_H$ enhanced k_1 much less than $T99K_H$ did for complexes $42_{R/O}$ because of $E221_H$ being farther from the HYDA2 binding face in complexes $42_{R/O}$. Combining $D102K_H$ with $E221K_H$ or $T99K_H$ in a double mutant enhanced k_1 even more. The double

TABLE 3: Association Rate Constants for the Complexes Containing HYDA2 Mutants, as Calculated Using BD Simulations[a]

	16_R	16_O	42_R	42_O
native	2.4 ± 0.8	1.9 ± 0.6	1.2 ± 0.5	0.6 ± 0.3
$T99K_H$	2.7 ± 0.6 (+13%)	2.6 ± 0.6 (+37%)	2.4 ± 0.6 (+100%)	1.5 ± 0.4 (+150%)
$D102K_H$	3.4 ± 0.9 (+42%)	3.0 ± 0.8 (+58%)	2.6 ± 0.9 (+117%)	1.8 ± 0.7 (+200%)
$M214K_H$	2.3 ± 0.7 (−4%)	2.0 ± 0.5 (+5%)	1.6 ± 0.2 (+33%)	1.0 ± 0.2 (+67%)
$E221K_H$	3.1 ± 1.2 (+29%)	2.5 ± 0.5 (+32%)	1.5 ± 0.3 (+25%)	1.1 ± 0.3 (+83%)
$D102K_H/T99K_H$	4.1 ± 0.8 (+71%)	3.4 ± 0.7 (+79%)	4.0 ± 0.8 (+233%)	3.1 ± 0.7 (+417%)
$D102K_H/E221K_H$	4.0 ± 1.0 (+67%)	3.2 ± 1.0 (+68%)	3.6 ± 1.2 (+200%)	3.0 ± 0.7 (+400%)

[a] Unit: $10^9\ M^{-1}\cdot s^{-1}$. Numbers shown in parentheses are the percentages of enhancement relative to the native values.

TABLE 4: FEP Calculations for the Mutated HYDA2 and Complexes[a]

		$T99K_H$	$D102K_H$	$M214K_H$	$E221K_H$
HYDA2	ΔG_1 (0 → 1)	−17.1	90.6	−35.8	65.2
	ΔG_1 (1 → 0)	−14.0	86.1	−39.4	61.6
	ΔG_1 (average)	−15.6 ± 1.6	88.4 ± 2.2	−37.6 ± 1.8	63.4 ± 1.8
16_R	ΔG_2 (0 → 1)	−24.7	87.0	−34.0	61.5
	ΔG_2 (1 → 0)	−29.2	84.0	−39.5	59.3
	ΔG_2 (average)	−27.0 ± 2.2	85.5 ± 1.5	−36.8 ± 2.7	60.4 ± 1.1
	$\Delta\Delta G$	−11.4 ± 2.7	−2.9 ± 2.7	0.8 ± 3.2	−3.0 ± 2.1
42_R	ΔG_2 (0 → 1)	−22.1	87.1	−31.7	66.7
	ΔG_2 (1 → 0)	−25.4	83.3	−34.9	63.1
	ΔG_2 (average)	−23.8 ± 1.7	85.2 ± 1.9	−33.3 ± 1.6	64.9 ± 1.8
	$\Delta\Delta G$	−8.2 ± 2.3	−3.2 ± 2.9	4.3 ± 2.4	1.5 ± 2.5

[a] Unit: kcal/mol.

mutants, $D102K_H/T99K_H$ and $D102K_H/E221K_H$, resulted in similar levels of enhancement in k_1 for all complexes. The relative level of enhancement was highest for complexes 42_O which also possessed the lowest k_1 values among the complexes containing the native HYDA2. This suggests that unfavorable electrostatic interactions limit k_1 in the diffusional encounter complexes with low k_1 values. Therefore, HYDA2 mutations improve association dynamics and have a greater impact on k_1 for the encounter complexes with low native k_1 values.

Table 4 presents the FEP calculation results for the mutated HYDA2 and complexes. The free-energy change, ΔG, for each mutant was obtained as the average of the forward and backward simulation runs, and binding free-energy changes, $\Delta\Delta G$, were calculated using eq 3. The standard errors of $\Delta\Delta G$, estimated by a simple propagation of error, are in the range of 1 to 3 kcal/mol. Among the four HYDA2 mutants, $T99K_H$ had the lowest $\Delta\Delta G$ value, suggesting a greater binding affinity of this mutant for FDX1. Its $\Delta\Delta G$ was −11.4 kcal/mol for 16_R and −8.2 kcal/mol for 42_R, corresponding to a 10^6-10^8-fold increase in the binding equilibrium constant, K_a. The $D102K_H$ mutants and 16_R-$E221K_H$ mutant stabilized binding modestly, with $\Delta\Delta G$ around −3 kcal/mol, or about a 100-fold increase in K_a. The complexes formed with $M214K_H$ and 42_R-$E221K_H$ resulted in positive $\Delta\Delta G$ values, indicating that these HYDA2 mutations destabilize rather than stabilize the binding complexes with FDX1. Nevertheless, their $\Delta\Delta G$ values were not prohibitively large. Interestingly, for the complexes containing the $M214K_H$ mutant, the $\Delta\Delta G$ value for 16_R was lower than that for 42_R. Because the HYDA2 residue $M214_H$ in the 16_R structure is close to the positively charged FDX1 residue $R40_F$, replacing methionine (M) with the positively charged lysine is expected to result in electrostatic repulsion within complex 16_R. However, a comparison of 16_R-$M214K_H$ and 42_R-$M214K_H$ complexes showed that substituting methionine with lysine stabilizes complex 16_R by forming two hydrogen bonds with $Y23_F$ and $Y80_F$ of FDX1, thus reducing $\Delta\Delta G$.

Concluding Remarks

In this study, we used BD simulations to study the kinetics of the HYDA2−FDX1 association. The kinetics results showed that the complexes $16_{R/O}$ are favored over complexes $42_{R/O}$, which is mainly due to the much larger association rate constants of the former, especially at high ionic strengths. Because complexes $16_{R/O}$ have higher association rate constants, lower binding free energies,[25] and lower free energies for the encounter complexes,[27] they are more likely to represent HYDA2−FDX1 association and electron transfer processes in vivo than complexes $42_{R/O}$. Moreover, because of the larger negative charge on the $[2Fe2S]_F$ cluster, reduced FDX1 is more negatively charged at the binding face than oxidized FDX1, resulting in a higher association rate with HYDA2. Therefore, we have shown that the underlying chemical and physical nature of the interaction face in HYDA2−FDX1 complexes kinetically favors association of the redox state that promotes charge transfer from FDX1 to HYDA2 and ultimately H_2 production.

We identified four sites on *C. reinhardtii* HYDA2, $T99_H$, $D102_H$, $M214_H$, and $E221_H$, where mutations have the potential to enhance complex formation with FDX1 and improve overall H_2 production efficiency. The BD simulations showed that $D102K_H$ had the most enhancements in the overall association rate constant, followed by $T99K_H$ or $E221K_H$ depending on the complex configuration and the redox state. The $M214K_H$ mutation resulted in the least enhancements. In addition to the

BD simulations, FEP calculations showed that mutations $T99K_H$ and $D102K_H$ stabilized binding complexes, whereas mutation $M214K_H$ destabilized them. By considering both the kinetic and thermodynamic results, we conclude that $T99K_H$ and $D102K_H$ represent the best candidates for future experimental studies.

We need to address three issues related to the accuracy of our work. First, we address the accuracy of our model structures. Because there is no experimental structure available for the *C. reinhardtii* hydrogenase, ferredoxin, or the binding complex, we used the homology-modeled structures obtained in our previous works. These structures were stable during the MD simulations in this and our previous research,[27] and our estimated association rate constants based on these structures are reasonable for protein associations. Therefore, we believe that these structures are one of the representatives for the physical structures and that the results estimated using these structures would provide valuable information for future protein-engineering research to improve the H_2 production efficiency of *C. reinhardtii*.

Second, we estimated the association rate constants by the criterion proposed by Gabdoulline et al.[38] It is important to note that even in Gabdoulline's work, the estimated association rate constants of some protein pairs had as large as ~30-fold differences compared with experimentally derived values. Moreover, we computed the association rate constants for the HYDA2 mutants using rigid body BD simulations without considering the influence of mutation on the conformations of other residues. Therefore, our association rate constants might only generally approximate the actual values. However, because in the current work we calculated the association rate constants between the same pair of proteins, that is, HYDA2 and FDX1, and because we assumed that one or two mutations on HYDA2 or different redox states of FDX1 would only slightly influence the kinetic properties, we expect that the systematic errors associated with our rate constants will be independent of the system analyzed. As a consequence, the calculated association rates can be meaningfully compared to each other.

Third, we address the FEP calculations. For most of the FEP calculation results, the forward and backward runs converged with a standard deviation of ~2 kcal/mol, which suggests that mutations did not result in significant protein conformation changes during the 5.6 ns simulation time. A possible source of systematic error is the lack of counterions in the simulations, where the simulation system was neutralized by a uniform charge atmosphere throughout the simulation box. Our simulations were thus biased from the reality. Moreover, two of the mutants ($T99K_H$ and $M214K_H$) have $\Delta e = +1e$ and the other two mutants ($D102K_H$ and $E221K_H$) have $\Delta e = +2e$. As a result, the density of the neutralizing charge atmosphere for mutants with $\Delta e = +1e$ is different from that for mutants with $\Delta e = +2e$, which causes significant differences in the calculated ΔG values. For example, for HYDA2, ΔG were negative for $T99K_H$ and $M214K_H$ mutants (−15.6 and −37.6 kcal/mol, respectively), but ΔG for $D102K_H$ and $E221K_H$ mutants had large positive values, +88.4 and +63.2 kcal/mol, respectively. However, it is expected that these systematic errors will be canceled when $\Delta\Delta G$ is calculated by eq 3 because each mutation has the same Δe for HYDA2 and HYDA2−FDX1 complexes.

On the basis of our BD simulation results, we propose the following detailed kinetics mechanism for H_2 production by hydrogenase−ferredoxin

$$[H]_{ox} + [F]_{red} + H^+ \underset{k_1,k_{-1}}{\leftrightarrow} [H\cdots F] \xrightarrow[k_2]{} [H]_{ox} + [F]_{ox} + \frac{1}{2}H_2 \quad (5)$$

where k_1 is the association rate constant, k_{-1} is the dissociation rate constant, and k_2 (or k_{cat}) is the H_2 catalytic reaction rate constant. The latter may involve transferring electrons and protons to the catalytic center and releasing H_2. A more rigorous discussion is presented in the Supporting Information. According to the Michaelis−Menten kinetics model, the overall H_2 production rate will be

$$\text{Rate} = \frac{1}{2}\frac{k_2[H_0]}{1 + \dfrac{K_M}{[F]_{red}[H^+]}} \quad (6)$$

where

$$K_M = \frac{k_{-1} + k_2}{k_1} = \frac{1}{K_a} + \frac{k_2}{k_1} \quad (7)$$

and K_a is the hydrogenase−ferredoxin binding equilibrium constant. In this research, we considered only the mutation sites on the hydrogenase binding face far from the $[4Fe4S]_H$ subcluster so that mutations would have little influence on k_2. Therefore, according to eqs 6 and 7, hydrogenase mutants with larger k_1, K_a, or both ($T99K_H$ and $D102K_H$) than those of native will lead to an improved overall H_2 production rate.

Finally, we can deduce the rate-limiting step for the kinetics mechanism of eq 5 using our BD simulation results. Equation 7 can be rewritten as

$$k_{-1} = K_M k_1 - k_2 \quad (8)$$

It is experimentally known that K_M is (1.0 to 3.5) × 10^{-5} M[18,19] and k_2 is ~600 s^{-1} measured at ~80 mM ionic strength.[19] Using these conditions, k_1 for the optimal complex 16_R can be calculated to be ~3 × 10^9 $M^{-1}\cdot s^{-1}$, which results in $K_M k_1$ = (3 to 10) × 10^4 s^{-1}, a value that is much larger than k_2. Therefore, according to eq 8, k_{-1} is estimated to be around (3 to 10) × 10^4 $s^{-1} \gg k_2$, and K_M will be

$$K_M \approx \frac{k_{-1}}{k_1} = \frac{1}{K_a} \quad (9)$$

In this case, the overall H_2 production rate can be approximated by

$$\text{rate} \approx \frac{1}{2}\frac{k_2[H_0]}{1 + \dfrac{1}{K_a[F]_{red}[H^+]}} \quad (10)$$

Equation 10 indicates that within the limits of our current models, the rate-limiting step for H_2 production might be the H_2 catalytic reaction, not protein association. Moreover, the binding constant, K_a, is more important for the H_2 production rate than the association rate constant, k_1.

In summary, we studied the HYDA2−FDX1 association process using BD and FEP-MD simulations. By analyzing the kinetics data obtained from BD simulations, we concluded that complex 16 represents a kinetically and thermodynamically preferred binding complex. On the basis of the structures of the binding complexes and on the alanine scanning results, we

identified four potential mutation sites on HYDA2 for improving its association with FDX1. The kinetic and thermodynamic properties of these mutants were studied using BD simulations and the MD-based FEP method. Within the limits of our current models, we found that two mutants, $T99K_H$ and $D102K_H$, might lead to diverting more electrons to the HYDA2 pathway and would be good candidates for future protein engineering studies to improve the H_2 production efficiency. Moreover, our results also implied that the H_2 catalytic reaction, not protein diffusion, might be the rate-limiting step for H_2 production.

Acknowledgment. We thank Jordi Cohen and Professor Klaus Schulten for the modified NAMD FEP codes and Dr. Christopher Chang for providing us with the charge and force field parameters of the metalloclusters. This work was supported by the Laboratory-Directed Research and Development Program of the U.S. Department of Energy's National Renewable Energy Laboratory (NREL). Computing resources at the NREL Scientific Computing Center were used in this work.

Supporting Information Available: Discussion of the H_2 production mechanism and coordinates for complexes 16 and 42. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) United States Department of Energy. Today's Hydrogen Production Industry. http://fossil.energy.gov/programs/fuels/hydrogen/currenttechnology.html.

(2) Happe, T.; Hemschemeier, A.; Winkler, M.; Kaminski, A. *Trends Plant Sci.* **2002**, *7*, 246–250.

(3) Melis, A.; Seibert, M.; Happe, T. *Photosynth. Res.* **2004**, *82*, 277–288.

(4) Prince, R. C.; Kheshgi, H. S. *Crit. Rev. Microbiol.* **2005**, *31*, 19–31.

(5) Esper, B.; Badura, A.; Rogner, M. *Trends Plant Sci.* **2006**, *11*, 543–549.

(6) Rupprecht, J.; Hankamer, B.; Mussgnug, J. H.; Ananyev, G.; Dismukes, C.; Kruse, O. *Appl. Microbiol. Biotechnol.* **2006**, *72*, 442–449.

(7) Ghirardi, M. L.; Posewitz, M. C.; Maness, P. C.; Dubini, A.; Yu, J.; Seibert, M. *Annu. Rev. Plant Biol.* **2007**, *58*, 71–91.

(8) Ghirardi, M. L.; Togasaki, R. K.; Seibert, M *Appl. Biochem. Biotechnol.* **1997**, *63*−*5*, 141–151.

(9) Ghirardi, M. L.; Zhang, J. P.; Lee, J. W.; Flynn, T.; Seibert, M.; Greenbaum, E.; Melis, A. *Trends Biotechnol.* **2000**, *18*, 506–511.

(10) Melis, A.; Happe, T. *Plant Physiol.* **2001**, *127*, 740–748.

(11) Tsygankov, A.; Kosourov, S.; Seibert, M.; Ghirardi, M. L *Int. J. Hydrogen Energy* **2002**, *27*, 1239–1244.

(12) Zhang, L. P.; Happe, T.; Melis, A. *Planta* **2002**, *214*, 552–561.

(13) Melis, A.; Zhang, L. P.; Forestier, M.; Ghirardi, M. L.; Seibert, M. *Plant Physiol.* **2000**, *122*, 127–135.

(14) Surzycki, R.; Cournac, L.; Peltiert, G.; Rochaix, J. D. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 17548–17553.

(15) Kosourov, S.; Seibert, M.; Ghirardi, M. L. *Plant Cell Physiol.* **2003**, *44*, 146–155.

(16) Mus, F.; Cournac, L.; Cardettini, W.; Caruana, A.; Peltier, G. *Biochim. Biophys. Acta* **2005**, *1708*, 322–332.

(17) Fouchard, S.; Hemschemeier, A.; Caruana, A.; Pruvost, K.; Legrand, J.; Happe, T.; Peltier, G.; Cournac, L. *Appl. Environ. Microbiol.* **2005**, *71*, 6199–6205.

(18) Happe, T.; Naber, J. D. *Eur. J. Biochem.* **1993**, *214*, 475–481.

(19) Roessler, P. G.; Lien, S. *Plant Physiol.* **1984**, *75*, 705–709.

(20) Forestier, M.; King, P.; Zhang, L. P.; Posewitz, M.; Schwarzer, S.; Happe, T.; Ghirardi, M. L.; Seibert, M. *Eur. J. Biochem.* **2003**, *270*, 2750–2758.

(21) Mus, F.; Dubini, A.; Seibert, M.; Posewitz, M. C.; Grossman, A. R. *J. Biol. Chem.* **2007**, *282*, 25475–25486.

(22) Jacquot, J. P.; Stein, M.; Suzuki, K.; Liottet, S.; Sandoz, G.; Miginiac-Maslow, M. *FEBS Lett.* **1997**, *400*, 293–296.

(23) Carrillo, N.; Ceccarelli, E. A. *Eur. J. Biochem.* **2003**, *270*, 1900–1915.

(24) Medina, M.; Abagyan, R.; Gomez-Moreno, C.; Fernandez-Recio, J. *Proteins: Struct. Funct. Bioinf.* **2008**, *72*, 848–862.

(25) Chang, C. H.; King, P. W.; Ghirardi, M. L.; Kim, K. *Biophys. J.* **2007**, *93*, 3034–3045.

(26) The subscript F denotes that this is a ferredoxin residue. Similarly, a subscript H is added for the residues on hydrogenase.

(27) Long, H.; Chang, C. H.; King, P. W.; Ghirardi, M. L.; Kim, K. *Biophys. J.* **2008**, *95*, 3753–3766.

(28) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

(29) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735–746.

(30) Morales, R.; Chron, M. H.; Hudry-Clergeon, G.; Petillot, Y.; Norager, S.; Medina, M.; Frey, M. *Biochemistry* **1999**, *38*, 15764–15773.

(31) Kortemme, T.; Baker, D. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 14116–14121.

(32) Robetta Web Server. http://www.robetta.org/.

(33) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

(34) Gabdoulline, R. R.; Wade, R. C. *Methods* **1998**, *14*, 329–341.

(35) de la Torre, J. G.; Huertas, M. L.; Carrasco, B. *Biophys. J.* **2000**, *78*, 719–730.

(36) Gabdoulline, R. R.; Wade, R. C. *J. Phys. Chem.* **1996**, *100*, 3868–3878.

(37) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10037–10041.

(38) Gabdoulline, R. R.; Wade, R. C. *J. Mol. Biol.* **2001**, *306*, 1139–1155.

(39) CVS version of NAMD source code can be downloaded at http://www.ks.uiuc.edu/Research/namd/development.html.

(40) Donnini, S.; Mark, A. E.; Juffer, A. H.; Villa, A. *J. Comput. Chem.* **2005**, *26*, 115–122.

(41) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Chem. Phys.* **1994**, *101*, 4177–4189.

(42) Feller, S. E.; Zhang, Y. H.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995**, *103*, 4613–4621.

(43) Pearson, D. C.; Gross, E. L. *Biophys. J.* **1998**, *75*, 2698–2711.

JP810409Z