J|A|C|S ARTICLES

Published on Web 06/11/2009

On the Contributions of Diffusion and Thermal Activation to Electron Transfer between *Phormidium laminosum* Plastocyanin and Cytochrome *f*: Brownian Dynamics Simulations with Explicit Modeling of Nonpolar Desolvation Interactions and Electron Transfer Events

Razif R. Gabdoulline*,†,‡,§ and Rebecca C. Wade*,†

Molecular and Cellular Modeling Group, EML Research gGmbH, Schloss-Wolfsbrunnenweg 33, D-69118 Heidelberg, Germany, and Center for Modeling and Simulation in the Biosciences (BIOMS), University of Heidelberg, Im Neuenheimer Feld 367, D-69120 Heidelberg, Germany

Received December 8, 2008; E-mail: razif.gabdoulline@biobase-international.com

Abstract: The factors that determine the extent to which diffusion and thermal activation processes govern electron transfer (ET) between proteins are debated. The process of ET between plastocyanin (PC) and cytochrome *f* (CytF) from the cyanobacterium *Phormidium laminosum* was initially thought to be diffusion-controlled but later was found to be under activation control (Schlarb-Ridley, B. G.; et al. *Biochemistry* **2005**, *44*, 6232). Here we describe Brownian dynamics simulations of the diffusional association of PC and CytF, from which ET rates were computed using a detailed model of ET events that was applied to all of the generated protein configurations. The proteins were modeled as rigid bodies represented in atomic detail. In addition to electrostatic forces, which were modeled as in our previous simulations of protein−protein association, the proteins interacted by a nonpolar desolvation (hydrophobic) force whose derivation is described here. The simulations yielded close to realistic residence times of transient protein−protein encounter complexes of up to tens of microseconds. The activation barrier for individual ET events derived from the simulations was positive. Whereas the electrostatic interactions between *P. laminosum* PC and CytF are weak, simulations for a second cyanobacterial PC−CytF pair, that from *Nostoc* sp. PCC 7119, revealed ET rates influenced by stronger electrostatic interactions. In both cases, the simulations imply significant contributions to ET from both diffusion and thermal activation processes.

Introduction

Cytochrome *f* (CytF) is a heme protein that is the largest subunit of the Cytochrome b_6f complex.[1,2] It accepts electrons from the Rieske protein and passes them to P_{700}^+ of photosystem I via plastocyanin (PC), a lumenal electron carrier of the cupredoxin family. The electron transfer (ET) between PC and CytF is a process involving at least two steps:

$$PC_{ox} + CytF_{red} \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} [PC_{ox}CytF_{red}] \underset{k_{-ET}}{\overset{k_{ET}}{\rightleftharpoons}} [PC_{red}CytF_{ox}]$$

In this study, we aimed to compute the rate constant for ET by treating the association and dissociation of PC and CytF as a continuous diffusional process (step 1) during which ET can occur (step 2). We neglected back-ET from PC to CytF. The mutual relative diffusion of the proteins was simulated, and during this diffusion, ET probabilities were calculated and accumulated to derive the total ET rate at the end of the simulation.

The application of Brownian dynamics simulation methods to ET proteins is well-established; the models used have become increasingly accurate[3−6] and have been applied to study PC and CytF.[6−8] In our previous studies of plant PC and CytF, we simulated only the diffusional association step using the Brownian dynamics methodology[9] and from these simulations computed the diffusional association rate constant.[10] The proteins were modeled as rigid bodies represented in atomic detail and interacting by electrostatic forces. The computed diffusional rate constant was correlated with the ET rate constant, and the simulations quantitatively showed how mutations of charged residues affect the ET rate. The electrostatic interactions between the plant proteins are strong, and the ET is apparently diffusion-controlled.

† EML Research gGmbH.
‡ University of Heidelberg.
§ Current address: BIOBASE GmbH, Halchterschestr. 33, D-38304 Wolfenbüttel, Germany.

(1) Hope, A. B. *Biochim. Biophys. Acta* **2000**, *1456*, 5.
(2) Hervás, M.; Navarro, J. A.; De la Rosa, M. A. *Acc. Chem. Res.* **2003**, *36*, 798.
(3) Allison, S. A.; Northrup, S. H.; McCammon, J. A. *Biophys. J.* **1986**, *49*, 167.
(4) Northrup, S. H.; Boles, J. O.; Reynolds, J. C. L. *Science* **1988**, *241*, 67.
(5) Northrup, S. H.; Thomasson, K. A.; Miller, C. M.; Barker, P. D.; Eltis, L. D.; Guillemette, J. G.; Inglis, S. C.; Mauk, A. G. *Biochemistry* **1993**, *32*, 6613.
(6) Pearson, D. C.; Gross, E. L. *Biophys. J.* **1998**, *75*, 2698.
(7) Gross, E. L. *Biophys. J.* **2004**, *87*, 2043s.
(8) Haddadian, E. J.; Gross, E. L. *Biophys. J.* **2005**, *88*, 2323.
(9) Gabdoulline, R. R.; Wade, R. C. *Methods* **1998**, *14*, 329.
(10) De Rienzo, F.; Gabdoulline, R. R.; Menziani, M. C.; DeBenedetti, P. G.; Wade, R. C. *Biophys. J.* **2001**, *81*, 3090.

 10.1021/ja809567k CCC: $40.75

In the cyanobacterial PC-CytF pairs that have been characterized to date, from *P. luminosum* and from *Nostoc* sp. PCC 7119, the electrostatic interactions are weaker, and the contributions of diffusion and thermal activation to their ET is debated.[11] To compute ET rates for these cyanobacterial PC−CytF pairs and provide an understanding of the factors affecting these rates, we require a more extensive computational model than used previously. First, we introduce an explicit treatment of the ET probability for each configuration of the two proteins generated during the simulations (step 2 above). Second, a fast, accurate calculation of short-range nonpolar forces between proteins is introduced. This model of nonpolar forces should be of general value for the simulation of other types of protein−protein association.

In our established model for the simulation of the diffusional association of proteins, the forces between them are computed as the sum of electrostatic interaction and electrostatic desolvation terms derived from a finite-difference Poisson−Boltzmann treatment.[10,12−14] This model has been applied to a wide variety of protein−protein complexes, including enzyme−inhibitor, antigen−antibody, and ET complexes.[10,12−14] Desolvation interactions of proteins (treated as spheres) were implemented previously in Brownian dynamics simulations of proteins.[15] That study showed that despite their short-range nature, the nonelectrostatic desolvation interactions could make significant contributions to protein association rates. Here we add nonpolar desolvation interaction terms to the electrostatic terms described previously[10,12−14] in a formulation that ensures model consistency.

Nonpolar desolvation terms make an energetically favorable contribution to protein−protein binding, and therefore, their addition to the interprotein forces increases the lifetime of diffusional encounter complexes. In this work, we concentrated on short-lived protein complexes, namely, those between the ET proteins CytF and PC. The interaction between PC and CytF is weak and fast, which ensures that the dissociation of the proteins does not limit turnover of the ET chain. At the same time, the interaction is specific enough to allow a close approach of the redox centers. The expected interaction times of these proteins are much shorter than for many other protein complexes (for example, enzyme−inhibitor complexes). The simulation of the association and dissociation of high-affinity proteins with long-lived complexes using atomically detailed models is not computationally feasible. On the other hand, the lifetimes of the ET complexes studied here are expected to be in the microsecond range,[11,16] and thus, the simulation of their association and dissociation, while computationally demanding, is feasible. We show here that it is possible to generate many trajectories of the relative diffusion of cyanobacterial CytF and PC and use these trajectories to derive accurate estimates of ET rates.

In the Methods section, we describe how the interaction forces were calculated in the Brownian dynamics (BD) simulations. The nonpolar desolvation term was calibrated by fitting the buried solvent accessible surface area (SASA), which was calculated by the "grid-site"-based method implemented in the BD simulations, to the buried SASA calculated by conventional methods for an ensemble of encounter complexes.[17] The details of the ET rate calculation procedure for BD simulations are described next. The method implemented here permitted the calculation and accumulation of ET probabilities along many possible pathways at each BD trajectory step.

In the Results section, we describe the outcome of BD simulations for the *P. laminosum* PC−CytF pair, based on the implemented interaction energy and individual ET rate calculation protocols. We found that ET rates can be predicted for a set of mutant proteins when the rate for the wild-type (WT) proteins is known. We also found that realistic activation free energy values and protein interaction times can be derived from these simulations. We then describe the application of our method to predict ET rates for the cyanobacterial *Nostoc* sp. PCC 7119 PC−CytF pair.

In the Discussion section, we compare our findings with previous calculations of ET rates, discuss the relative contributions of diffusion and activation to the ET process, and point out the current limitations of the simulation model.

Methods

Interaction Energies and Forces. The interprotein forces are computed at each step of a BD simulation as finite-difference derivatives of the following free energy of interaction between the proteins:

$$\Delta G = \frac{1}{2}\sum_{i2}\Phi_{\mathrm{el}1}q_{i2} + \frac{1}{2}\sum_{j1}\Phi_{\mathrm{el}2}q_{j1} + \sum_{i2}\Phi_{\mathrm{edesolv}1}{q_{i2}}^2 + \sum_{j1}\Phi_{\mathrm{edesolv}2}{q_{j1}}^2 + \sum_{m2}\Phi_{\mathrm{npdesolv}1}\mathrm{SASA}_{m2} + \sum_{n1}\Phi_{\mathrm{npdesolv}2}\mathrm{SASA}_{n1} \quad (1)$$

The first four terms in eq 1 are electrostatic terms and the last two terms are nonpolar terms. For the electrostatic terms, the summations are over all charges q in protein 1 with $j1 = 1, 2,...$ (terms 2 and 4) and in protein 2 with $i2 = 1, 2,...$ (terms 1 and 3). The first two terms correspond to the energy for the interaction of the charges of one protein (q_{i2} or q_{j1}) with the electrostatic potential of the other protein ($\Phi_{\mathrm{el}1}$ or $\Phi_{\mathrm{el}2}$, respectively). The third and fourth terms correspond to the energy due to the interaction of the charges of one protein (q_{i2} or q_{j1}) with the electrostatic desolvation potential of the other protein ($\Phi_{\mathrm{edesolv}1}$ or $\Phi_{\mathrm{edesolv}2}$, respectively) resulting from the approach of the low-dielectric cavity of the latter protein. The last two terms describe the nonpolar interactions due to the burial of the SASAs of the surface atoms (SASA_{m2}, SASA_{n1}) of one protein by the other protein, which are calculated using the respective "burial potentials" ($\Phi_{\mathrm{npdesolv}1}$, $\Phi_{\mathrm{npdesolv}2}$).

The first two terms are identical to the first term of the effective charge approximation to the electrostatic interaction free energy given by the Poisson−Boltzmann equation.[12] It is the asymptotic approximation to this free energy for electrostatic interactions between the proteins at large separations.

The third and fourth terms correspond to the previously introduced electrostatic desolvation penalty.[12,18] We found that when using the formula from ref 18,

(11) Schlarb-Ridley, B. G.; Mi, H.; Teale, W. D.; Meyer, V. S.; Howe, C. J.; Bendall, D. S. *Biochemistry* **2005**, *44*, 6232.
(12) Gabdoulline, R. R.; Wade, R. C. *J. Phys. Chem.* **1996**, *100*, 3868.
(13) Gabdoulline, R. R.; Wade, R. C. *Biophys. J.* **1997**, *72*, 1917.
(14) Gabdoulline, R. R.; Wade, R. C. *J. Mol. Biol.* **2001**, *306*, 1139.
(15) Camacho, C. J.; Kimura, S. R.; DeLisi, C.; Vajda, S. *Biophys. J.* **2000**, *78*, 1094.
(16) Crowley, P. B.; Ubbink, M. *Acc. Chem. Res.* **2003**, *36*, 723.
(17) Gabdoulline, R. R.; Wade, R. C. *J. Mol. Recognit.* **1999**, *12*, 226.
(18) Elcock, A. H.; Gabdoulline, R. R.; Wade, R. C.; McCammon, J. A. *J. Mol. Biol.* **1999**, *291*, 149.

$$\Delta G_{\text{edesolv}} = \sum_i \Phi_{\text{edesolv}}(i)\, q_i^2$$

$$= \alpha \frac{\varepsilon_s - \varepsilon_p}{\varepsilon_s(2\varepsilon_s + \varepsilon_p)} \sum_i \sum_j q_i^2 a_j^3 \frac{(1 + \kappa r_{ij})^2}{r_{ij}^4} e^{-2\kappa r_{ij}} \quad (2)$$

one should use different values of the scaling factor α at different ionic strengths in order to reproduce the desolvation penalties calculated by finite-difference methods. In eq 2, κ is the Debye−Hückel parameter and ε_s and ε_p are the relative solvent and protein dielectric permittivities, respectively. The desolvation penalty is the sum of all the interactions of charges q_i of one protein with the corresponding spherical low-dielectric cavities j (of radius a_j) of the other protein; r_{ij} is the distance between the ith charge and the jth cavity.

To simplify the calculation protocol, and in view of the fact that the electrostatic desolvation penalty terms are small (less than 1 kcal/mol), we used electrostatic desolvation forces computed at zero ionic strength to represent the forces at all ionic strengths. Electrostatic desolvation forces were thus calculated by setting $\kappa = 0$ in eq 2 (which corresponds to eq B1 in ref 12). The scaling factor α was assigned a value of 0.36 to reproduce the electrostatic desolvation energies calculated by finite-difference methods at zero ionic strength. Since we used all of the atoms of the protein as the induced dipole sites (indexed by j in eq 2), a value less than 1 for this scaling factor reflects the overlap of these atoms.

For the last two terms, precalculated grids of a piecewise linear function $\Phi_{\text{npdesolv}}(r)$ that depends on the distance r between the van der Waals surface of one protein and each grid point are used. The energy terms in eq 1 are then calculated from the positions of the centers of the atoms of one protein on the grid of $\Phi_{\text{npdesolv}}(r)$ values computed for the other protein. The function $\Phi_{\text{npdesolv}}(r)$ is given by

$$\Phi_{\text{npdesolv}}(r) = \beta c \begin{cases} 1 & r < a \\ \dfrac{b - r}{b - a} & a < r < b \\ 0 & r > b \end{cases} \quad (3)$$

where β (in units of kcal mol^{-1} Å^{-2}) is a factor for converting the calculated buried area to hydrophobic desolvation energy and c is a factor introduced to obtain the buried area after the function $\Phi_{\text{npdesolv}}(r)/\beta$ (calculated for one protein) is multiplied by atomic accessibilities (of the other protein), as in eq 1. The function in eq 3 depends on two distance parameters, a and b. When the distance r between the van der Waals surface of one protein and the center of the atom of another protein is less than a, complete burial of the atom is assumed. The burial is fractional for distances between a and b and zero for distances larger than b. By varying the parameters a and b, one can achieve the best approximation to the buried SASA calculated by standard methods (here, NACCESS[19]). With $a = 3.1$ Å, $b = 4.35$ Å, and $c = 0.5$, the buried areas were approximated to within 10% root-mean-square deviation (rmsd), or 20 Å^2, for 50 encounter complexes of PC and CytF generated in BD simulations (see Materials and Calculation Details). In deriving values of the parameters a and b, the calculated areas were first correlated with NACCESS areas to derive a constant scaling factor c, and the rmsd values between the NACCESS areas and the scaled areas obtained using the grid-site method were calculated. The best performance of the grid-site method was obtained with $a = 3.1$ Å, $b = 4.35$ Å and $c = 0.5$ (see Figure 1). The values $a = 2.6$ Å and $b = 3.35$ Å approximated the buried area at $c = 1.0$ but with less accuracy. This method is similar to that described by Elcock and McCammon[20] but appears to be much more accurate: the correlation coefficient for the energies in Figure 1 is $r^2 = 0.988$, compared with $r^2 = 0.79$ in ref 20.

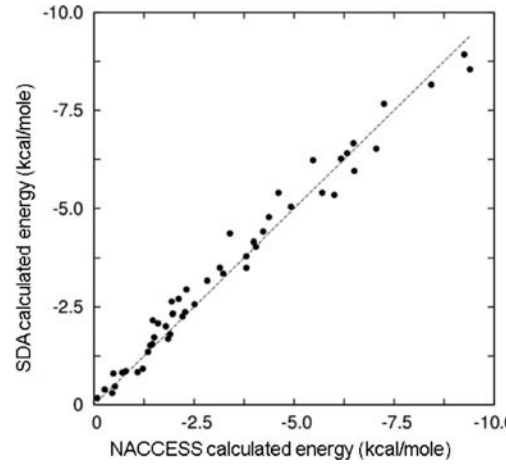

Figure 1. Correlation between the buried-area energies calculated using NACCESS (x axis) and using the last two terms in eq 1 as implemented in the SDA software (y axis) with $a = 3.1$ Å, $b = 4.35$ Å, and $c = 0.5$. The buried area is multiplied by the factor $\beta = -0.025$ kcal mol^{-1} Å^{-2} [35] to obtain interaction energy estimates in kcal/mol.

Electron Transfer Rate Calculations. At each step of a BD trajectory, it is supposed that the nonadiabatic ET rate constant k_{ET} at temperature T is given by Marcus's formula:[21]

$$k_{\text{ET}} = \frac{4\pi^2 H_{\text{AD}}^2}{h\sqrt{4\pi\lambda k_B T}} \exp\left[-\frac{(\Delta G^\circ + \lambda)^2}{4\lambda k_B T}\right] \quad (4)$$

where ΔG° is the free energy of an ET reaction and λ is the reorganization energy (i.e., the energy required to reorganize the system from the initial to the final coordinates without changing the electronic state) of the configuration at this BD step. We assume that neither of these energies varies during a trajectory. Their sum in eq 4 can therefore be considered to have a constant value during all of the BD simulations, and its exact value can be defined by comparing the final computed rate with an experimentally measured rate (only one comparison is needed).

H_{AD} is the electronic-coupling matrix element between the acceptor (A) and donor (D) sites. Through-bond terms in this coupling are computed for each possible pathway from the electron donor or acceptor to each surface atom. The through-space factor is assumed to be $\exp[-1.7(d - 1.4)]$,[22] where d is the distance (in Å) between the surface-accessible atoms of the donor and acceptor. H_{AD} was computed at each BD step for each possible pathway from the donor to the acceptor. For computational efficiency, the number of pathways recorded was restricted to the 500 having the largest through-bond coupling constants.

We calculate the escape probability (i.e., the probability that no ET occurs) during a BD step as the following product over all possible ET pathways between the surface atoms i of the donor protein and the surface atoms j of the acceptor protein:

$$P_{\text{escape}} = \prod_{ij} (1 - k_{\text{ET},ij}\Delta t)$$

where Δt is the time step and $k_{\text{ET},ij}$ is the ET rate constant (eq 4) along the path going through surface atoms i and j. If the energies

(19) Hubbard, S. J.; Thornton, J. M. *NACCESS*; Department of Biochemistry and Molecular Biology, University College London: London, 1993.

(20) Elcock, A. H.; McCammon, J. A. *Biophys. J.* **2001**, *80*, 613.
(21) Marcus, R. A.; Sutin, N. *Biochim. Biophys. Acta* **1985**, *811*, 265.
(22) Liang, Z.-X.; Kurnukov, I. V.; Nocek, J. M.; Mauk, A. G.; Beratan, D. N.; Hoffman, B. M. *J. Am. Chem. Soc.* **2004**, *126*, 2785.

are written in units of k_BT, the distances in units of angstroms, and time in picoseconds, then this escape probability is:

$$P_{\text{escape}} = \prod_{ij}\left(1 - 10^5 \exp\left[-\Delta G^{\#} - \frac{\ln(\lambda)}{2}\right] H_{\text{D}i}{}^2 H_{\text{A}j}{}^2 \times \{0.6 \exp[-1.7(d_{ij} - 1.4)]\}^2 \Delta t\right) \quad (5)$$

where d_{ij} is the distance between the surface atoms coupled to the donor and the acceptor and $\Delta G^{\#} = (\Delta G^{\circ} + \lambda)^2/4\lambda$ is the activation free energy. In deriving eq 5, we have converted all of the constants to the above-mentioned units and moved the factor of $\lambda^{-1/2}$ in eq 4 to the exponent. $H_{\text{D}i}$ and $H_{\text{A}j}$ are the couplings between surface atom i of the donor protein and the donor site and between surface atom j of the acceptor protein and the acceptor site, respectively. They are precalculated before carrying out the BD simulations by using the HARLEM methodology.[22] The d_{ij} are calculated during the BD simulations. The composite term (starting with 10^5 and ending with Δt) in eq 5 can rewritten in a simple form as:

$$H_{\text{D}i}{}^2 H_{\text{A}j}{}^2 \exp\left[15.25 - \Delta G^{\#} - \frac{\ln(\lambda)}{2} - 3.4 d_{ij}\right] \Delta t$$

This expression contains two unknown quantities, $\Delta G^{\#}$ and λ, but it is only necessary to know their combination in the form $W = 15.25 - \Delta G^{\#} - \ln(\lambda)/2$. Thus, one can do simulations with different values of W and select the most appropriate value by comparing the final calculated rates with the rates expected on the basis of experiments. Matching only one experimental ET rate is sufficient to derive W, and this value can be used to calculate ET rates for similar systems. To avoid repeating the simulations with different values of W, we calculated rates simultaneously for a set of different values of W during a single BD simulation.

Materials and Calculation Details. The structure of the complex of *P. laminosum* PC and CytF[23] was determined by NMR and provided by M. Ubbink (Leiden University, The Netherlands). The *Nostoc* sp. PCC 7119 PC−CytF complex structure[24] was determined by NMR and obtained from the RCSB (PDB entry code 1tu2). Each coordinate file contains an ensemble of complexes, and in each of these, the respective conformations of PC and CytF are the same and are those determined for the unbound forms of each protein. The first structural models in the coordinate files were used in the simulations. It should be noted that the choice of the structural model of the complex influences only the definition of the encounter complexes used to calibrate the interaction energies, not the BD simulation results. This is true because the conformations of the individual proteins are identical in all of the structural models in the ensemble of structures in each coordinate file and also because the rate calculations in this work do not use information about the specific configuration of the complex structure. The coordinates of mutants of *P. laminosum* PC were constructed using Swiss-PdbViewer.[25]

BD simulations were carried out with a modified version of the SDA software[9] capable of computing nonpolar desolvation interactions and calculating ET probabilities during BD simulations. The diffusion constants, time steps, and starting and ending distances for trajectories were the same those as used previously for simulations of plant PC and CytF.[10] The translational and rotational diffusion coefficients of PC were assigned to be 1.37×10^{-2} Å^2/ps and 4.07×10^{-5} ps^{-1}, respectively, whereas the corresponding values for CytF were 1.27×10^{-2} Å^2/ps and 3.27×10^{-5} ps^{-1}. The minimum time step for BD moves was 1 ps. At the beginning of each trajectory, PC was positioned with a randomly chosen orientation at a randomly chosen point on the surface of a sphere of radius $b = 100$ Å centered on the center of mass of CytF. BD simulations were then performed until PC diffused outside a sphere of radius $q = 500$ Å. The radius b was chosen because at this distance the forces between PC and CytF were centrosymmetric and isotropic; the radius q was chosen to ensure that the diffusive flux of PC was centrosymmetric. The total ET rates were computed by accumulating probabilities for escape (i.e., a move without a reaction, ET in this case) during the BD trajectory and then converting these escape probabilities into a reaction (ET) probability and deriving the total ET rate, as previously implemented in generic protein−protein association simulations.[9]

For the WT *P. laminosum* PC−CytF pair at 100 mM ionic strength, 2000 BD trajectories were generated using a β value of -0.013 kcal mol^{-1} Å^{-2} (see eq 3) assigned to describe the strength of the nonpolar desolvation forces. These trajectories took ∼32 CPU hours to run on a single 2.4 GHz Linux cluster node. The number of trajectories was set to be inversely proportional to the expected rate constant for other values of β in order to optimize the ratio of calculated rate accuracy to required CPU time. The longest simulations were run with $\beta = -0.019$ kcal mol^{-1} Å^{-2} and required ∼1200 h for 240 trajectories. The rate constants were computed as averages over eight separate runs; the error in the rates was estimated from the rmsd of the eight separate rates, with 50% being a tolerated limit.

Polar hydrogen atoms were added with Whatif.[26] Protein atoms were assigned partial charges and radii in accordance with the OPLS force field.[27] The partial charges of copper and iron atoms and their ligands in PC and CytF were assigned as described in an earlier publication,[10] taking into account the differences in sequence but assuming the same metal cofactor ligation. Electrostatic calculations were done with UHBD.[28] Electrostatic potential grids for BD simulations with SDA, with a spacing of 1 Å and with 150^3 grid points, were calculated by solving the finite-difference linearized Poisson−Boltzmann equation for each simulated protein. We used the van der Waals surface to define the protein interior, which was assigned a low relative dielectric constant of 4.0. Solvent (water) outside the protein was assigned a relative dielectric constant of 78.5. An ion radius of 2.0 Å was used to define the Stern layer.[28] Electrostatic desolvation grids, with a spacing of 1 Å and with 110^3 grid points, were calculated as described previously[14] but with the changes described in the section 'Interaction Energies and Forces'. The nonpolar desolvation grids also had a spacing of 1 Å and 110^3 grid points.

Encounter complexes for calibrating the interaction energies were generated using our previous BD simulation protocol.[10] We carried out BD simulations (using SDA[9] without nonpolar desolvation forces) of diffusional association of the proteins to form their bound complexes. Encounter complexes were defined as configurations with two native polar contacts (observed in the bound complex structure) with contact distances between 6.0 and 6.5 Å. The number of complexes was reduced to 50, keeping only those with a pairwise rmsd larger than 2 Å.

Electrostatic energy calculations on encounter complexes were done with UHBD. The two different types of term, the grid−charge interaction energy (ΔE_c) and the electrostatic desolvation penalty ($\Delta E_{\text{edesolv}}$) were calculated separately. First, the total electrostatic interaction free energy (ΔE_t) was calculated following the standard protocol $\Delta E_t = E_{t,\text{bound}}(1 + 2) - E_{t,\text{unbound}}(1) - E_{t,\text{unbound}}(2)$, where the E_t's are the total energies of bound and unbound (separated) proteins, which are calculated by UHBD as

(23) Crowley, P. B.; Otting, G.; Schlarb-Ridley, B. G.; Canters, G. W.; Ubbink, M. *J. Am. Chem. Soc.* **2001**, *123*, 10444.
(24) Diaz-Moreno, I.; Diaz-Quintana, A.; De la Rosa, M. A.; Ubbink, M. *J. Biol. Chem.* **2005**, *280*, 18908.
(25) Guex, N.; Peitsch, M. C. *Electrophoresis* **1997**, *18*, 2614.
(26) Vriend, G. *J. Mol. Graphics* **1990**, *8*, 52.
(27) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.
(28) Madura, J. D.; Briggs, J. M.; Wade, R. C.; Davis, M. E.; Luty, B. A.; Ilin, A.; Antosiewicz, J.; Gilson, M. K.; Bagheri, B.; Scott, L. R.; McCammon, J. A. *Comput. Phys. Commun.* **1995**, 1995.

$$E_t = \frac{1}{2}\sum_i \Phi(r_i)\, q(r_i)$$

in which the summation of charges q in the system multiplied by the electrostatic potential Φ at the positions of these charges runs over all of the positions r_i. The electrostatic desolvation penalty was calculated as

$$\Delta E_{\text{edesolv}} = E_{\text{t,bound}}(1+0) - E_{\text{t,unbound}}(1) + E_{\text{t,bound}}(0+2) - E_{\text{t,unbound}}(2)$$

where $E_{\text{t,bound}}(1+0)$ and $E_{\text{t,bound}}(0+2)$ are the total energies of the bound states with protein 1 with standard charges and protein 2 with all charges set to zero and vice versa, and $E_{\text{t,unbound}}(1)$ and $E_{\text{t,unbound}}(2)$ are the total energies of the separated proteins 1 and 2, respectively. A three-step grid focusing was used, and the grid size was varied from 110 to 150 to ensure that the values were not affected by discretization errors (deviation threshold of 1%).

Results

Electron Transfer Rate Constants. ET rate contants were calculated for *P. laminosum* PC and CytF at various ionic strengths, as described in Methods. Calculations were done for the WT and two mutants of PC. The mutants of PC chosen have the highest (D44A) and lowest (R93E) ET rates of all mutants studied.[29]

A β value of -0.013 kcal mol^{-1} Å^{-2} was used in these rate calculations. This value of β was chosen because it results in sufficiently strong nonpolar desolvation interactions to ensure physically meaningful ET process properties with a positive activation barrier. We will show later on that a β value of -0.019 kcal mol^{-1} Å^{-2} is more consistent with experiments in our model; however, simulations with this value of β require significantly more computational resources, making repeat control calculations of rates impractical.

Simulations were performed first for the WT proteins at 100 mM. It was found that in order to reproduce the experimental ET rate constant (4.4×10^7 M^{-1} s^{-1}), the parameter W defined above had to be set to $10.7k_BT$ (at $T = 25$ °C; see Electron Transfer Rate Calculations). This value of W was then used to calculate the rate constants for ET between the two PC mutants (D44A and R93E) and CytF as well as the PC−CytF ET rates at different ionic strengths. With this value of W, there was good agreement between calculated and experimentally measured ET rates (see Table 1 and Figure 2). This implies that neither these mutations of PC nor variations in the ionic strength significantly alter the activation and reorganization energies of the ET steps. Therefore, the calculated ET rates can be considered as predictions based on one experimental value. These simulations correctly calculated not only the rates that decrease with increasing ionic strength (WT and the D44A mutant) but also those that increase with ionic strength (the R93E mutant) when electrostatic interactions hamper the ET reaction. This suggests that with this approach, it should be possible to reproduce complex, nonmonotonic ionic-strength dependences of ET rates for which a specific combination of nonpolar desolvation and electrostatic interactions results in a maximal ET rate at an intermediate value of the ionic strength.

Activation Energies. The value of the activation energy in each case was derived by the following procedure. Initially, the ET rate constants were computed for different values of the unknown parameter W. After the simulations, the value of W that reproduced the experimental ET rate constant for the WT proteins at 100 mM ionic strength was selected (as described in the previous section).

From the results of the simulations, the value $W = 10.7k_BT$ was chosen (see Figure 3). This means that the equality $\Delta G^{\#} + \ln(\lambda)/2 = 4.55$ should be satisfied for activation and reorganization energies, restricting the reorganization and activation energies to a limited set of values. This relationship approximately defines the value of $\Delta G^{\#}$ at almost any value of the other unknown parameter λ, because the dependence of $\Delta G^{\#}$ on λ (at fixed W) is weak (see Figure 4). In our energy estimates, we assumed $\lambda \approx 24k_BT$, corresponding to the estimates for electrostatic reorganization energies derived from encounter-complex configurations; this value is in accordance with estimates for similar systems, for example, cytochrome c.[30] If we were to take into account the fact that the driving force for ET reactions between PC and CytF is small (<20 mV[11,29,31]),

Table 1. Experimental[29] and Calculated ET Rate Constants for *P. laminosum* PC (WT and Two Mutants) and CytF at Various Ionic Strengths

		k_{ET} (10^7 M^{-1} s^{-1})		
case	ionic strength (mM)	exptl	calcd	σ_{calcd} (10^7 M^{-1} s^{-1})[a]
WT	0	−	11.4	2.7
	15	6.0	11.9	4.6
	100	4.4	4.4[b]	2.2
	500	3.5	2.8	1.0
D44A	0	−	86.7	44.6
	15	22	50	25
	100	7.3	8.9	3.2
	500	4.4	3.9	2.1
R93E	15	0.4	0.39	0.19
	100	1.0	0.7	0.3
	500	1.6	1.5	0.4

[a] Statistical error of the calculated k_{ET} value, defined as the rmsd of the rate constants computed in eight separate BD simulations with respect to their average. The experimental values have an error of ~5%.[29] [b] The value of W was chosen to reproduce the k_{ET} value for WT proteins at 100 mM. The same value of W was used to compute k_{ET} values for all the other cases.

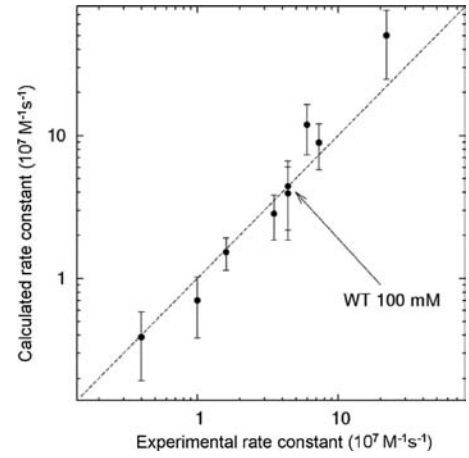

Figure 2. Experimental and calculated ET rate constants for *P. laminosum* PC (WT and two mutants) with CytF at ionic strengths of 15, 100, and 500 mM (see Table 1).

(29) Schlarb-Ridley, B. G.; Bendall, D. S.; Howe, C. J. *Biochemistry* **2002**, *41*, 3279.

(30) Muegge, I.; Qi, P.; Wand, A.; Chu, Z.; Warshel, A. *J. Phys. Chem. B* **1997**, *101*, 825.

(31) Ubbink, M.; Ejdebaeck, M.; Karlsson, B. G.; Bendall, D. S. *Structure* **1998**, *6*, 323.

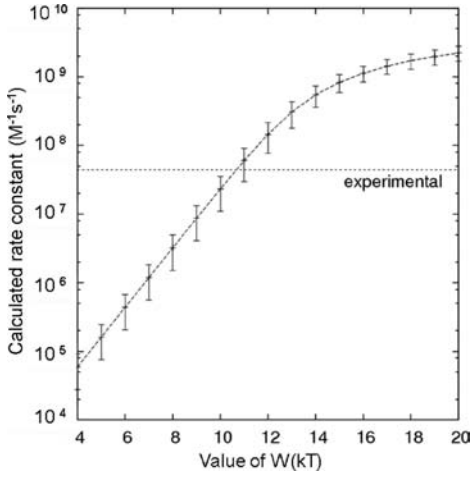

Figure 3. Procedure for deriving the value of the unknown parameter W from the simulations. The ET rate constants were computed for a set of W values, and then the value of W at which the calculated rate matches the experimentally measured rate was chosen (see the text for details).

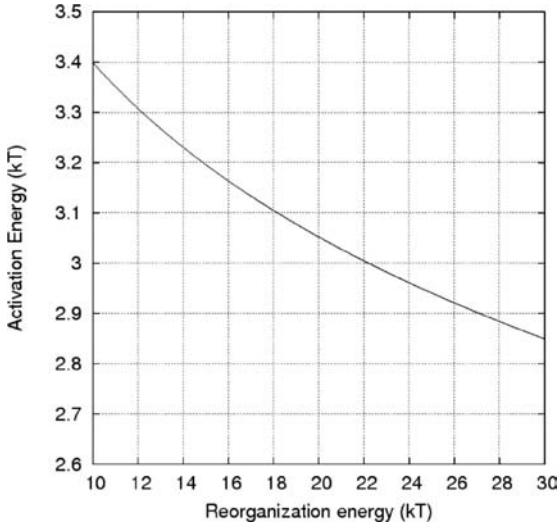

Figure 4. Possible values of the reorganization energy λ (x axis) and activation energy $\Delta G^{\#}$ (y axis) derived from BD simulations (with $W = 10.7k_BT$, i.e., $\Delta G^{\#} + \ln(\lambda)/2 = 4.55k_BT$).

which means that $\Delta G^{\circ} \approx 0$, we would obtain an estimate for the reorganization energy of $\lambda \approx 4\Delta G^{\#}$. This relationship would imply a smaller reorganization energy ($\lambda \approx 13k_BT$) for this particular value of β (-0.013 kcal mol^{-1} Å^{-2}) but larger values of λ (up to $\sim 50k_BT$) for other values of β investigated (see below).

We also carried out simulations with the nonpolar desolvation force factor β assigned a larger magnitude (i.e., $\beta = -0.015$, -0.017, or -0.019 kcal mol^{-1} Å^{-2}). The calculated rate contant increases as the magnitude of β increases. Therefore, W and the activation energy should be assigned higher values in order to reproduce the experimental ET rate constant for the WT proteins at 100 mM ionic strength. There is a clear linear trend between the activation free energy and the nonpolar desolvation force factor (see Figure 5).

We estimated the average interaction time between PC and CytF as the difference between the average length of the BD trajectories simulated with and without interaction forces.

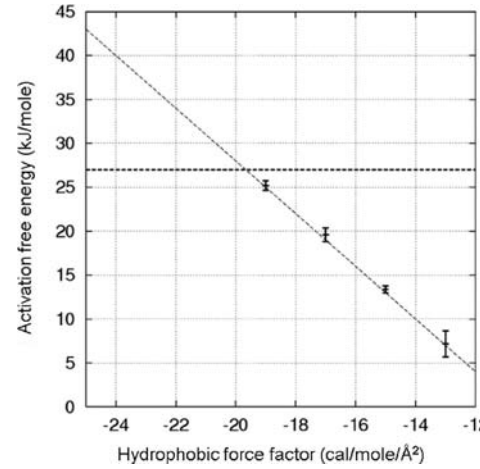

Figure 5. Dependence of the activation energy assigned to reproduce the ET rate constant for WT *P. laminosum* PC−CytF at 100 mM ionic strength upon the value of the nonpolar desolvation force factor used in the simulations. The horizontal line is an estimate of the activation energy, defined as the difference between the activation enthalpy for total ET and the activation enthalpy attributed to diffusion[11] (see the Discussion).

Average interaction times increased from 0.2 ± 0.04 to 40 ± 10 μs as β was changed from -0.013 to -0.019 kcal mol^{-1} Å^{-2}.

Prediction of the ET Rate Constant for Another Cyanobacterial PC−CytF Pair. We performed additional calculations for the recently investigated *Nostoc* sp. PCC 7119 PC and CytF pair.[24,32] An inspection of the redox-center geometry showed no detectable difference from that in *P. laminosum*. Therefore, in the simulations we assumed the same reorganization and activation energies and used the same value of W ($10.7k_BT$) for this case. The simulations showed that in this case the electrostatic interactions contribute significantly to the ET rate constant, which was found to be as high as $(12.5 \pm 5.6) \times 10^7$ M^{-1} s^{-1} at 100 mM ionic strength and $(0.48 \pm 0.2) \times 10^7$ M^{-1} s^{-1} at 500 mM ionic strength, indicating rate enhancement by at least a factor of 20 due to electrostatic interactions. These results are in agreement with the experimental results available to date, namely, the strong dependence on the electrostatic interactions of the pseudo-first-order observed rate constants for this system.[32]

Discussion

Calculation of Relative and Absolute ET Rates. Previously performed simulations of protein ET systems[7,10,33] neglected nonpolar (and hydrophobic) interactions between ET proteins as well as the details of individual ET transfer events along different pathways. Nevertheless, these earlier simulations gave reasonably good results for ET rates. To reproduce these earlier simulations, we carried out BD simulations without nonpolar desolvation forces and calculated ET rate constants on the basis of redox center-to-center distances. The simulation setup as well as the results were similar to those from previous simulations of protein ET systems.[7,10,33] From these simulations, it was possible to accurately calculate the *relative* rates of the mutant proteins compared with the WT proteins. This shows that the dominant contribution to the relative ET rates comes from the difference in electrostatic interactions and that simple methods

(32) Albarrán, C.; Navarro, J. A.; Molina-Heredia, F. P.; Murdoch, P. d. S.; De la Rosa, M. A.; Hervás, M. *Biochemistry* **2005**, *44*, 11601.
(33) Lin, J.; Beratan, D. N. *J. Phys. Chem. B* **2005**, *109*, 7529.

can be successful in predicting the rate ratios in some cases. However, with these simplified approaches, it is not possible to gain meaningful insights into the magnitudes of the energies relevant for the absolute rates and into the details of the ET process.

Activation or Diffusion Control of ET in *P. laminosum* PC−CytF. Overall, the ET between proteins in the present simulations results from a combination of two steps. The first is the diffusion of the proteins to some bimolecular configuration, and the second is an ET event at that configuration. The total ET rate depends significantly on the parameters of these two steps. When the second step is very fast at some accessible configurations, then the ET rate is defined only by the diffusion to those configurations. In this case, the ET probability for at least one configuration should be close to 1 during its lifetime, which is the 1 ps time step in the BD simulations. We found, however, that for the values of W accepted after the simulations, the maximal ET probability in 1 ps was $\sim 10^{-4}$. This type of behavior is called "activation-controlled" in ref 33 because the total ET rate depends markedly on the activation barriers to ET at the individual protein configurations. Inspection of all the cases simulated in this work showed that they can all be classified as activation-controlled, because the total ET rate constant depends exponentially on W near the selected value of W (similar to the situation in Figure 3).

Nevertheless, the contributions of diffusion and long-range electrostatic steering are also significant, because the rate curves are lower or higher depending on electrostatic steering and the variation of the ET rate constants was found to be modulated by electrostatic interactions. Therefore, one must consider the total ET process as a two-step process with two barriers to overcome: a diffusional barrier (as described in ref 34) and an activation barrier for ET at given configurations. The diffusional barrier, E_{diff}, can be derived by expecting the diffusion-controlled rate constant for association (k_{diff}) to exhibit Arrhenius behavior: $k_{\text{diff}} \propto \exp(-E_{\text{diff}}/k_{\text{B}}T)$.[34] This barrier corresponds to an energy cost for diffusional transport and is defined by the temperature dependence of the diffusion constants, which in turn is defined by the dependence of the solvent viscosity on temperature. We will neglect entropic contributions to the activation free energies because in our model these contributions are very small, being mainly due to the temperature dependence of the water dielectric permittivity. This allows us to relate the derived activation free energies to the activation enthalpies for total ET derived in ref 11 from the temperature dependence of the ET rate. Specifically, for the WT proteins at 100 mM ionic strength, we can assume that the activation enthalpy for total ET, $\Delta H^{\text{tot}} \approx 47$ kJ/mol (see ref 11; here the units are those used in the cited papers, i.e., 1 kJ/mol = 0.239 kcal/mol, 1 kcal/mol = 4.184 kJ/mol, $k_{\text{B}}T \approx 0.6$ kcal/mol at $T = 25$ °C), is the sum of the diffusional part of the enthalpy, $\Delta H^{\text{diff}} \approx 20$ kJ/mol (derived from data on the temperature dependence of the viscosity of water; see ref 34), and our estimate of the activation free energy for an individual ET step, $\Delta G^{\#}$. Thus, an estimate of $\Delta G^{\#}$ is given by the difference $\Delta H^{\text{tot}} - \Delta H^{\text{diff}}$, which is ~ 27 kJ/mol. This is larger than the value of $\sim 3k_{\text{B}}T$ (~ 7.5 kJ/mol) obtained from the simulations with $\beta = -0.013$ kcal mol^{-1} Å^{-2}. Simulations with the larger-magnitude nonpolar desolvation forces may result in a better consistency with experiment, because they assume a larger value of $\Delta G^{\#}$. Indeed, the simulations with $\beta = -0.019$ kcal mol^{-1} Å^{-2} gave an activation free energy for an individual ET step of $\Delta G^{\#} \approx 24$ kJ/mol, which is close to the difference of the two enthalpies. Therefore, one probably can assume that a β value of -0.019 kcal mol^{-1} Å^{-2} describes the magnitude of the nonpolar interactions realistically. It is also close to the magnitude of the β value of -0.025 kcal mol^{-1} Å^{-2} defined by fitting calculated and experimental data to weak interactions of proteins (second virial coefficient).[35] Independent of its exact value, the contribution of the ET step to the nondiffusional activation barrier should be considered significant. This is in agreement with experimental evidence that, in the case of protein association not involving ET, the activation barrier can be almost fully attributed to diffusion (see ref 36), although this is not always the case.[37]

Interaction and ET Patterns for Different PC−CytF Pairs. Comparing three different cases (see Figure 6), one can see that the interactions between PC and CytF not only alter the sizes of regions with high ET probability but can also shift them. We calculated interaction energies (E) when the centers of PC were placed on a regular grid around CytF with 1 Å spacing, and 4000 orientations were generated randomly at each grid point. Strong interaction regions were defined as having a Boltzmann factor $\exp(-E/k_{\text{B}}T) > 10^5$ (Figure 6, top, white regions). The ET probability was also calculated at all configurations using eq 5, and the regions of ET-capable configurations (shown in red in Figure 6) were identified as those with ET probabilities of $>10^{-15}$ ps^{-1}. The regions of ET-capable configurations in the presence of interactions were defined as those for which the product of the ET probability and the Boltzmann factor $\exp(-E/k_{\text{B}}T)$ was $>10^{-10}$ ps^{-1}. These regions are shown in Figure 6, (bottom, white regions). Completely different interaction patterns can be seen for the three cases, namely, PC−CytF from *P. laminosum*, *Nostoc* sp. PCC 7119, and plant sources. In addition, the region of strong interactions does not directly correspond to the ET-capable positions.

Limitations of the Simulation Model. Configurational reorganization processes during association of ET proteins were suggested to be important[16] when the proteins first associate electrostatically to a complex that is not ET-competent and then rearrange to an ET-competent complex because of hydrophobic interactions. This type of reorganization may have taken place in our simulations, but it was difficult to extract any distinguishable contribution to the ET rates from this type of dynamics. In our model, which includes only diffusional dynamics, the rate depends mainly on the interaction energy in the vicinity of the ET-capable conformations.[38] The situation may differ if nondiffusional processes occur between different states, but this type of nondiffusional dynamic event is not present in the model.

An interesting observation of a nonlinear viscosity dependence of ET rates for *P. laminosum* PC−CytF[11] cannot be reproduced in these simulations unless the ET step is assumed to depend on interactions with viscogens that are not modeled explicitly in our simulations. It is argued in ref 11 on the basis of studies with various viscogens that specific interactions between proteins and viscogens cannot explain the observed viscosity dependence, thus suggesting that this viscosity de-

(34) van Holde, K. E. *Biophys. Chem* **2002**, *101*−*102*, 249.
(35) Elcock, A. H.; Sept, D.; McCammon, J. A. *J. Phys. Chem. B* **2001**, *105*, 1504.
(36) Frisch, C.; Fersht, A. R.; Schreiber, G. *J. Mol. Biol.* **2001**, *308*, 69.
(37) Schreiber, G. *Curr. Opin. Struct. Biol.* **2002**, *12*, 41.
(38) Zhou, H. X.; Briggs, J. M.; Tara, S.; McCammon, J. A. *Biopolymers* **1998**, *45*, 355.

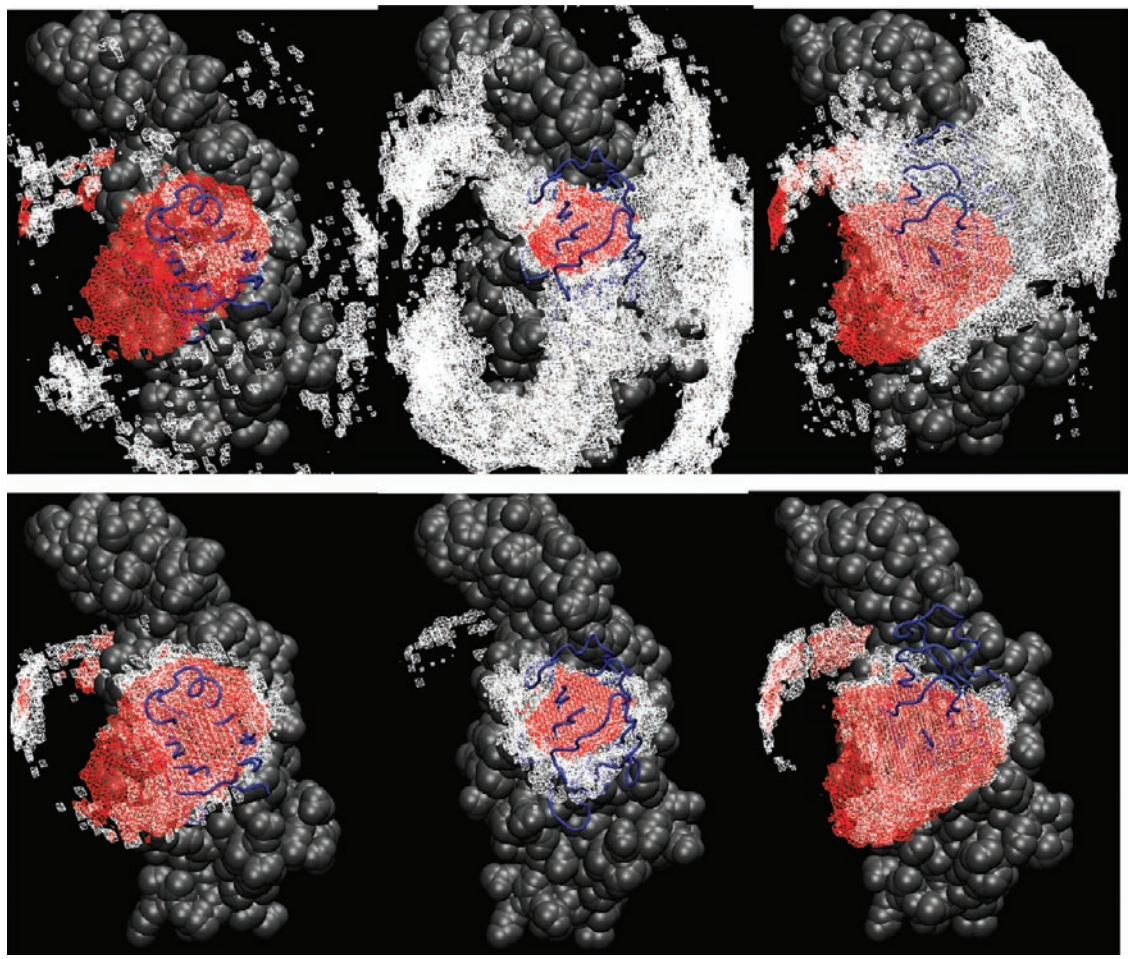

Figure 6. Comparison of the ET properties of three PC−CytF pairs. PC in the NMR structure of its complex with CytF (shown as gray van der Waals spheres) is depicted by the blue tubes. Red regions show ET probability ($>10^{-15}$ ps^{-1}) positions of the center of PC around CytF for (from left to right) *P. laminosum*, *Nostoc* sp. PCC 7119, and plant sources. White regions are regions of (top) strong energetic interactions [$\exp(-E/k_BT) > 10^5$] and (bottom) ET-capable configurations in the presence of interactions (ET probability * $\exp(-E/kT) > 10^{-10}$ ps^{-1}).

pendence could be caused by crowding effects[39] or the osmotic pressure effect of viscogens.[40] Simulations taking into account crowding, osmotic pressure effects, and specific interactions between proteins and viscogens would be needed in order to capture phenomena other than the decrease in diffusion constants due to the increase in the concentration of viscogens.

In the simulation method, the internal dynamics of the proteins are not modeled, and the proteins are treated as rigid. As the proteins spend more time in close proximity, this approximation may have a greater effect in these simulations than in simulations done without a nonpolar desolvation term.[13] In addition, hydrodynamic interactions[41] could have an influence on the course of the interaction.

The magnitude of the nonpolar interaction term, as well as that of the electrostatic term, can be questioned. Electrostatics can be treated using the molecular surface as a low−high dielectric interface between the protein and the solvent. Typically, this results in higher electrostatic desolvation penalties than in the model used in the present work, where the van der Waals surface serves as the dielectric boundary. The nonpolar desolvation factor used here, $\beta = -0.013$ kcal mol^{-1} Å^{-2}, can also be replaced by lower or higher values, for example, $\beta = -0.019$ kcal mol^{-1} Å^{-2}. The latter results in higher and more realistic values of the activation free energies of the ET step at the individual configurations of the proteins. The β values -0.013 and -0.019 kcal mol^{-1} Å^{-2} give total interaction energies (which do not include translational, rotational, and internal vibrational entropy changes during protein interactions) of -7.5 and -11.4 kcal/mol, respectively, for the *P. laminosum* PC−Cyt F complex. This complex has a buried interface area of ~600 Å^2 and lacks appreciable electrostatic interactions. These computed interaction energies are in a reasonable agreement with the observed millimolar-to-micromolar affinity of PC−CytF.[11]

Conclusions

We have implemented atomic-detail nonpolar desolvation interactions in BD simulations of proteins and applied the simulation method to ET between plastocyanin and cytochrome *f*. We have also implemented a detailed model of ET events at each configuration of the proteins generated. The main advantage of the resultant model is that it captures the energetic balance of the different contributions to protein interactions. With a suitable choice of the weight of the nonpolar desolvation term, the model gives positive values of the activation energy

(39) Kuttner, Y. Y.; Kozer, N.; Segal, E.; Schreiber, G.; Haran, G. *J. Am. Chem. Soc.* **2005**, *127*, 15138.
(40) Kozer, N.; Kuttner, Y. Y.; Haran, G.; Schreiber, G. *Biophys. J.* **2007**, *92*, 2139.
(41) Antosiewicz, J.; McCammon, J. A. *Biophys. J.* **1995**, *69*, 57.

barrier for ET events at the individual configurations of the proteins. The model without nonpolar desolvation interactions has a negative value of this barrier (-35 ± 2 kJ/mol, which can be derived by extrapolation of the linear relationship in Figure 5 to $\beta = 0$). The interaction times of the proteins obtained with the model including nonpolar desolvation forces are realistic, being on the microsecond time scale, in contrast to the nanosecond time scale obtained in the model without nonpolar desolvation forces.

In these simulations, one does not need to know which contacts are formed between the two proteins, so the simulations can be performed using the 3D structures of unbound proteins. The calculation of bimolecular ET rate constants is dependent on a single parameter, which is a combination of reorganization and activation energies for ET at individual configurations of the two proteins. We have shown that this parameter can be derived using a single experimentally measured ET rate constant and can then be used to calculate and predict ET rate constant values in other similar cases. Formulated in another way, this type of simulation can shed light on the value of the (otherwise difficult to measure) activation barrier for ET when the total ET rate has been measured. For the cyanobacterial PC and CytF pairs studied here, we have found that both the activation and diffusional barriers contribute to the ET rates.

Acknowledgment. This work was supported by the Klaus Tschira Foundation and the Center for Modelling and Simulation in the Biosciences (BIOMS) in Heidelberg. We thank Prof. Marcellus Ubbink for providing the NMR structure of the *P. laminosum* PC−CytF complex and Daria Kokh, Georgi Pachov, and Anna Feldman-Salit for critical reading of the manuscript.

JA809567K