Computational Simulation and Analysis of Dynamic Association between Plastocyanin and Cytochrome *f*. Consequences for the Electron-Transfer Reaction

G. Matthias Ullmann,*,† Ernst-Walter Knapp,† and Nenad M. Kostić*,‡

Contribution from the Institut für Kristallographie, Freie Universität Berlin, Takustrasse 6, 14195 Berlin, Germany, and Department of Chemistry, Iowa State University, Ames, Iowa 50011

Received July 1, 1996[⊗]

Abstract: This work is a theoretical analysis in four stages of association between the blue copper protein plastocyanin and the heme protein cytochrome *f*, which are physiological partners in the photosynthetic electron-transfer chain. In the first stage, 32 000 trajectories of approach by plastocyanin to cytochrome *f* were generated with implicit consideration of hydration and with gradual cooling of the system from 300 to 0 K. Approximately 2000 trajectories resulted in local minima of energy, i.e., in docking. The molecular configurations having relatively low energies were grouped, by structural similarity, into six families. In the second stage, six configurations having the lowest energies, one from each family, were subjected to thorough molecular dynamics simulation, for 260 ps. Extensive hydration of the proteins was treated explicitly. The whole plastocyanin molecule and the relevant parts of the cytochrome *f* molecule were given conformational freedom. In the third stage, the following three contributions to the energy of binding were calculated: polarization of water by the proteins, determined from numerical solutions of the Poisson−Boltzmann equation; nonelectrostatic (van der Waals and other) interactions involving the proteins and water; and the Coulombic interactions within and between the protein molecules. Total energy of association was calculated with a thermodynamic cycle; several realistic sets of parameters gave consistent results. The configuration having the most favorable Coulombic interactions turned out to have the second highest total energy. This finding exemplifies the importance of allowing for hydration and for conformational flexibility in docking calculations and perils of neglecting these factors. In the fourth stage, electronic coupling between the copper and heme sites in the six configurations was analyzed and compared by the Pathways method. The configuration providing the most efficient path for electron tunneling turned out to be different from the most stable configuration. There are indications that the evident interaction between Lys65 in cytochrome *f* and Tyr83 in plastocyanin may involve the ammonium group of the former and the aromatic ring of the latter. These surprisingly strong noncovalent interactions, so-called charge-π interactions, have recently been discovered and are important for molecular recognition. Modeling and structural optimization of these interactions are beyond the state of the art in molecular mechanics, but these studies should become possible with improved force fields. The electron-transfer reaction between cupriplastocyanin and ferrocytochrome *f* is fast in the noncovalent complex and undetectably slow in the covalent complex. This contrast is explained in terms of our theoretical analysis.

Introduction

Protein association (docking) plays a central role in metabolism, flow of genetic information, and molecular physiology. Therefore, association has been studied by experimental and theoretical methods, at an ever increasing pace. Numerous theories and computational methods have been applied to simulation of the docking process and of the resulting diprotein complexes. In most of these studies the protein molecules are treated as rigid bodies, and their surfaces are matched. One general method is to seek complementarity of surface patterns.[1,2] In some cases the energy of the resulting complex is calculated;[3,4] in the latest studies energy is being estimated from numerical solutions of the Poisson−Boltzmann equation.[5,6] Another general method is to seek complementarity of Coulombic electrostatic fields by maximizing the number of paired opposite charges or by minimizing the electrostatic energy of the diprotein complex.[7−12] Yet another method, which involves simulation of Brownian dynamics, requires a reaction criterion and can be used in calculating and analyzing the rate of association.[13−18]

† Freie Universität Berlin.
‡ Iowa State University.
⊗ Abstract published in *Advance ACS Abstracts,* December 15, 1996.

(1) Shoichet, B. K.; Kuntz, I. D. *J. Mol. Biol.* **1991**, *221*, 327−346.
(2) Helmer-Citterich, M.; Tramontano, A. *J. Mol. Biol.* **1994**, *235*, 1021−1031.
(3) Cherfils, J.; Janin, J. *Curr. Opin. Struct. Biol.* **1993**, *3*, 265−269.
(4) Kollman, P. A. *Curr. Opin. Struct. Biol.* **1994**, *4*, 240−245.
(5) Jackson, R. M.; Sternberg, M. J. E. *J. Mol. Biol.* **1995**, *250*, 258−275.
(6) Zhang, T.; Koshland, D. E. *Protein Sci.* **1996**, *5*, 348−356.
(7) Poulos, T. L.; Kraut, J. *J. Biol. Chem.* **1980**, *255*, 10322−10330.
(8) Mauk, M. R.; Mauk, A. G.; Weber, P. C.; Matthew, J. B. *Biochemistry* **1983**, *25*, 7085−7091.
(9) Poulos, T. L.; Mauk, A. G. *J. Biol. Chem.* **1983**, *258*, 7369−7373.
(10) Matthew, J. J.; Weber, P. C.; Salemme, F. R.; Richards, F. M. *Nature* **1983**, *301*, 169−171.
(11) Roberts, V. A.; Freeman, H. C.; Olson, A. J.; Tainer, J. A.; Getzoff, E. D. *J. Biol. Chem.* **1991**, *266*, 13431−13441.
(12) Adir, N.; Axelrod, H. L.; Beroza, P.; Isaacson, R. A.; Rongey, S. H.; Okamura, M. Y. Feher, G. *Biochemistry* **1996**, *35*, 2535−2547.
(13) Ermak, D. L.; McCammon, J. A. *J. Chem. Phys.* **1978**, *69*, 1352−1360.
(14) Northrup, S. H.; Allison, S. A.; McCammon, J. A. *J. Chem. Phys.* **1984**, *80*, 1517−1524.
(15) Northrup, S. H.; Boles, J. O.; Reynolds, J. C. L. *J. Phys. Chem.* **1987**, *91*, 5991−5998.
(16) Northrup, S. H.; Boles, J. O.; Reynolds, J. C. L. *Science* **1988**, *241*, 67−70.
(17) (a) Eltis, L. D.; Herbert, R. G.; Barker, P. D.; Mauk, A. G.; Northrup, S. H. *Biochemistry* **1991**, *30*, 3663−3674. (b) Northrup, S. H.; Thomasson, K.; Miller, C. M.; Barker, P. D.; Eltis, L. D.; Guillemette, J. G.; Inglis, S. C.; Mauk, A. G. *Biochemistry* **1993**, *32*, 6613−6623. (c) Andrew, S. M.; Thomasson, K. A.; Northrup, S. H. *J. Am. Chem. Soc.* **1993**, *115*, 5516−5521.

S0002-7863(96)02237-8 CCC: $14.00

In most of the theoretical studies to date the aqueous environment and the protein flexibility have not been considered. Conformational flexibility has been included in calculations by molecular dynamics,[19] Monte Carlo methods,[20,21] and search for side chain rotamers.[22] Of the various methods for analyzing electron-transfer properties of proteins[23−31] the most popular is the one embodied in the program Pathways.[32−36]

This study deals with the blue copper protein plastocyanin, designated pc, and the heme protein cytochrome *f*, designated cytf, whose three-dimensional structures are known. These proteins are involved in photosynthetic electron transfer: cupriplastocyanin accepts an electron from ferrocytochrome *f*, and cuproplastocyanin donates an electron to the oxidized form of photosystem I. These proteins are well-suited to experimental and theoretical studies of association and electron-transfer reactions.[37−44] Plastocyanin is notable because it contains two distinct surface patches through which it can exchange electrons with redox partners. The broad, negatively-charged acidic patch around Tyr83 is remote from the copper atom, whereas the electroneutral, hydrophobic patch around His87, a ligand to the copper atom, is proximate to this atom. Both of these important amino acid residues are somewhat exposed on the surface. Despite the different distances, these two patches are approximately equally coupled to the copper site.[39,45−48]

The luminal part of cytochrome *f* can gently be cleaved from the short segment anchoring it, as a part of the cytochrome b_6f complex, to the membrane. Recent crystallographic analysis of this solubilized form of cytochrome *f* revealed a remarkable two-domain structure.[49] The larger domain contains heme with the amino group of the terminal residue Tyr1 as an axial ligand to the iron atom. The smaller domain contains a patch of positively-charged residues. Unexpectedly, the heme and this patch are relatively far apart.

When plastocyanin and cytochrome *f* are noninvasively cross-linked in a reaction mediated by a carbodiimide,[50] the resulting covalent complex cannot detectably undergo the internal electron-transfer reaction, which is fast within the electrostatic complex.[51,52] This unreactivity was taken as evidence that the two proteins dock and react with each other in different configurations.[52] This prediction and this analysis were nicely corroborated by subsequent publication of the structure of cytochrome *f*, which showed that the positively-charged patch and the heme are relatively far apart. Studies of the rearrangement in complexes that plastocyanin forms with native (iron-containing) cytochrome *c* and with its zinc derivative gave evidence for configurational fluctuation, in which the two associated protein molecules fluctuate around the initial docking configuration, without greatly deviating from it.[53−56]

The present study deals with the association shown in eq 1 and with the subsequent electron-transfer reaction shown in eq 2; the Roman numerals are the oxidation states of copper and iron, and the slant represents the diprotein complex.

$$\mathrm{pc(II)} + \mathrm{cytf(II)} \rightleftharpoons \mathrm{pc(II)/cytf(II)} \qquad (1)$$

$$\mathrm{pc(II)/cytf(II)} \rightarrow \mathrm{pc(I)/cytf(III)} \qquad (2)$$

We combine theoretical methods in a new way. First, we find numerous electrostatically-favorable predocking orientations of the rigid molecules by a Monte Carlo procedure, with implicit consideration of hydration. Next, we carry out molecular dynamics simulations for configurations representing local minima of Coulombic energy. At this stage we explicitly treat hydration, consider local hydrophobic effects, and allow for conformational flexibility. Next, we analyze energetics of each complex on the basis of a thermodynamic cycle including numerical solution of the Poisson−Boltzmann equation. Finally, we analyze electron-transfer paths in all the optimized configurations, and also in some others, by the Pathways method, taking into account differences in the covalency of the iron−ligand and copper−ligand bonds at the active sites. To our knowledge, this is the first quantitative simulation of the complex between

(18) Northrup, S. H. *Curr. Opin. Struct. Biol.* **1994**, *4*, 269−274.
(19) Wendoloski, J. J.; Matthew, J. B.; Weber, P. C.; Salemme, F. R. *Science* **1987**, *238*, 794−797.
(20) Zacharias, M.; Luty, B. A.; Davis, M. E.; McCammon, J. A. *J. Mol. Biol.* **1994**, *238*, 455−465.
(21) Totrov, M.; Abagyan, R. *Struct. Biol.* **1994**, *1*, 259−263.
(22) Leach, A. R. *J. Mol. Biol.* **1994**, *235*, 345−356.
(23) (a) Siddarth, P.; Marcus, R. A. *J. Phys. Chem.* **1990**, *94*, 8430−8434. (b) Siddarth, P.; Marcus, R. A. *J. Phys. Chem.* **1993**, *97*, 2400−2405. (c) Siddarth, P.; Marcus, R. A. *J. Phys. Chem.* **1993**, *97*, 6111−6114. (d) Siddarth, P.; Marcus, R. A. *J. Phys. Chem.* **1993**, *97*, 13078−13082.
(24) Stuchebrukhov, A. A. *Chem. Phys. Lett.* **1994**, *225*, 55−61.
(25) (a) Moser, C. C.; Keske, J. M.; Warncke, K.; Farid, R. S.; Dutton, P. L. *Nature* **1992**, *355*, 796−802. (b) Farid, R. S.; Moser, C. C.; Dutton, P. L. *Curr. Opin. Struct. Biol.* **1993**, *3*, 225−233.
(26) Christensen, H. E. M.; Conrad, L. S.; Mikkelsen, K. V.; Nielsen, M. K.; Ulstrup, J. *Inorg. Chem.* **1990**, *29*, 2808−2816.
(27) (a) Skourtis, S. S.; Onuchic, J. N. *Chem. Phys. Lett.* **1993**, *209*, 171−177. (b) Skourtis, S. S.; Beratan, D. N.; Onuchic, J. N. *Chem. Phys.* **1993**, *176*, 501−520.
(28) Skourtis, S. S.; Mukamel, S. *Chem. Phys.* **1995**, *197*, 367−388.
(29) (a) Larson, S. *J. Am. Chem. Soc.* **1981**, *103*, 4034−4040. (b) Larson, S. *J. Chem. Soc., Faraday Trans. 2* **1983**, *79*, 1375−1388.
(30) Gruschus, J. M.; Kuki, A. *J. Phys. Chem.* **1993**, *97*, 5581−5593.
(31) (a) Evenson, J. W.; Karplus, M. *J. Chem. Phys.* **1992**, *96*, 5272−5278. (b) Evenson, J. W.; Karplus, M. *Science* **1993**, *262*, 1247−1249.
(32) (a) Beratan, D. N.; Onuchic, J. N. *Photosynth. Res.* **1989**, *22*, 173−186. (b) Onuchic, J. N.; Beratan, D. N. *J. Chem. Phys.* **1990**, *92*, 722−733.
(33) Beratan, D. N.; Onuchic, J. N.; Betts, J. N.; Bowler, B. E.; Gray, H. B. *J. Am. Chem. Soc.* **1990**, *112*, 7915−7921.
(34) Beratan, D. N.; Betts, J. N.; Onuchic, J. N. *Science* **1991**, *252*, 1285−1288.
(35) Beratan, D. N.; Onuchic, J. N.; Winkler, J. R.; Gray, H. B. *Science* **1992**, *258*, 1740−1741.
(36) Onuchic, J. N.; Beratan, D. N.; Winkler, J. R.; Gray, H. B. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 349−377.
(37) Redinbo, M. R.; Yeates, T. O.; Merchant, S. *J. Bioeng. Biomembr.* **1994**, *26*, 49−66.
(38) Gross, E. L. *Photosynth. Res.* **1993**, *37*, 103−116.
(39) Sykes, A. G. *Adv. Inorg. Chem.* **1991**, *36*, 377−408.
(40) Sykes, A. G. *Struct. Bonding (Berlin)* **1991**, *75*, 177−244.
(41) Drepper, F.; Hippler, M.; Nitschke, W.; Haehnel, W. *Biochemistry* **1996**, *35*, 1282−1295.
(42) Haehnel, W.; Jansen, T.; Gause, K.; Klösgen, R. B.; Stahl, B.; Michl, D.; Huvermann, B.; Karas, M.; Herrmann, R. G. *EMBO J.* **1994**, *13*, 1028−1038.
(43) Hervas, M.; Navarro, J. A.; Diaz, A.; Bottin, H.; De la Rosa, M. A. *Biochemistry* **1995**, *34*, 11321−11326.
(44) Sigfridsson, K.; Young, S.; Hanson, Ö. *Biochemistry* **1996**, *35*, 1249−1257.
(45) Kyritsis, P.; Lundberg, L. G.; Nordling, M.; Vängard, T.; Young, S.; Tomkinson, N. P.; Sykes, A. G. *J. Chem. Soc., Chem. Commun.* **1991**, *8*, 1441−1442.
(46) Solomon, E. I.; Lowery, M. D. *Science* **1993**, *259*, 1575−1581.
(47) Ullmann, G. M.; Kostić, N. M. *J. Am. Chem. Soc.* **1995**, *117*, 4766−4774.
(48) Qin, L.; Kostić, N. M. *Biochemistry* **1996**, *35*, 3379−3386.
(49) Martinez, S. E.; Huang, D.; Szczepaniak, A.; Cramer, W. A.; Smith, J. L. *Structure* **1994**, *2*, 95−105.
(50) (a) Davis, D. J.; Hough, K. *Biochem. Biophys. Res. Commun.* **1983**, *116*, 1000−1006. (b) Morand, L. Z.; Frame, M. K.; Colvert, K. K.; Johnson, D. A.; Krogmann, D. W.; Davis, D. J. *Biochemistry* **1989**, *28*, 8039−8047. (c) Morand, L. Z.; Frame, M. K.; Colvert, K. K.; Johnson, D. A.; Krogmann, D. W.; Davis, D. J. *Biochemistry* **1989**, *28*, 10093.
(51) Qin, L.; Kostić, N. M. *Biochemistry* **1992**, *31*, 5145−5150.
(52) Qin, L.; Kostić, N. M. *Biochemistry* **1993**, *32*, 6073−6080.
(53) Zhou, J. S.; Kostić, N. M. *J. Am. Chem. Soc.* **1992**, *114*, 3562−3563.
(54) Zhou, J. S.; Kostić, N. M. *Spectrum* **1992**, *5*, 1−6.
(55) Zhou, J. S.; Kostić, N. M. *J. Am. Chem. Soc.* **1993**, *115*, 10796−10804.
(56) Peerey, L. M.; Kostić, N. M. *Biochemistry* **1989**, *28*, 1861−1868.

Table 1. Atomic Charges for the Blue Copper Site of Cupriplastocyanin[a,b]

ligand	atom	charge	ligand	atom	charge
	Cu	0.35	His87	$N_{\delta1}$	−0.42
Cys84	S_γ	−0.26		C_γ	0.16
	C_β	0.26		$C_{\varepsilon1}$	0.41
His37	$N_{\delta1}$	−0.47		$C_{\delta2}$	0.12
	C_γ	0.16		$N_{\varepsilon2}$	−0.41
	$C_{\varepsilon1}$	0.41		$H_{\varepsilon2}$	0.39
	$C_{\delta2}$	0.12	Met92	S_δ	−0.24
	$N_{\varepsilon2}$	−0.41		C_γ	0.19
	$H_{\varepsilon2}$	0.39		C_ε	0.25

[a] Supplied by E. I. Solomon. [b] Hydrogen atoms are united with carbon atoms to which they are bound.

plastocyanin and cytochrome *f* and the first combination of the aforementioned theoretical and computational methods.

Methods

Structures and Charges. The crystal structure of the luminal domain of turnip cytochrome *f*[49] is the entry 1ctm in the Protein Data Bank. The two C-terminal residues absent from this structure were modeled by a simulated annealing procedure based on molecular dynamics, while the rest of the protein was kept rigid. Sixteen similar structures of bean plastocyanin in solution, determined by ^{1}H NMR spectroscopy,[57] comprise the entry 9pcy; each was used in the calculations.

Charges of most atoms, including those in the heme[58] and hydrogen atoms in polar bonds, were taken from the parameters of CHARMM.[59] Carbon and hydrogen atoms bonded to each other were united. Charges of the copper atom and of its ligands, calculated by a density functional method and consistent with spectroscopic evidence, were kindly supplied by Professor E. I. Solomon; see Table 1. Charges of titratable groups are those at pH 7.0, assuming standard pK_a values. Electrostatic calculations by a published method[60,61] showed His142 in cytochrome *f* to be electroneutral at pH 7.0.

Monte Carlo Calculations. In the search for the predocking orientation the protein molecules were treated as rigid bodies. The center of mass of cytochrome *f* was placed at the center of two concentric cubes having edges of 200 and 400 Å and of two concentric spheres having radii of 100 and 160 Å. The dielectric constant was $\varepsilon = 80.0$. The grid in the small box encompassing cytochrome *f*, which was used in testing whether plastocyanin entered the excluded volume, had the spacing of 0.5 Å. The Coulombic potential of cytochrome *f* was mapped on two cubic grids, with spacings of 1.0 Å in the inner cube and of 2.0 Å in the outer, each having 200 × 200 × 200 points. See Figure 1. The energy of plastocyanin in this electrostatic field was evaluated by multiplying the atomic charges with the field values, obtained by linear interpolations at the atomic positions.

Calculating the interaction energies for thousands of orientations with the Poisson−Boltzmann equation would demand excessive memory and time. If the potential of cytochrome *f* were calculated with this equation, its dielectric inhomogeneity would have been treated exactly. Then calculating the energy of plastocyanin as was done here would have neglected the dielectric inhomogeneity of this protein. In this case, the solvation contribution would have been considered for the former protein but neglected for the latter. In our method this contribution is consistently neglected for the sake of internal consistency, so that the same interaction energy is obtained regardless of which protein molecule is moving in the field of which.

The first configurations of the diprotein complex were simulated by an annealing procedure. Initial positions and orientations of plastocyanin were chosen randomly, by placing its center of mass on the inner sphere. Each step consisted of a random displacement and a random rotation. The components of these two vectors were taken from normal distributions with means of zero and with respective deviations of 2.0 Å and 5.0°. Each step was accepted or rejected according to the Metropolis algorithm.[62] If the energy of the new situation was lower than that of the preceding one, the step was always accepted. If the new energy was higher by ΔE than the preceding one, the probability of the step was taken to be proportional to $\exp(-\Delta E/RT)$. This step was accepted or rejected depending on whether this probability was greater or less than the random probability of the given ΔE. If plastocyanin entered the excluded volume of cytochrome *f*, the step was rejected. A total of 32 000 trajectories were created. After 3000 accepted steps at 300 K the temperature was lowered by 10 K. After each 300 accepted steps thereafter the temperature was again lowered by 10 K. An unproductive trajectory was ended if the distance between the centers of mass of the two proteins exceeded 160 Å (that is, if plastocyanin left the outer sphere). A productive trajectory was ended if 2000 consecutive steps were rejected; this was the criterion for a local minimum. At 0 K the steps were continued until either of these criteria was met. The procedure just described yielded approximately 2000 different diprotein configurations. Those with the highest and the lowest Coulombic energies were found. One hundred fourty configurations had electrostatic energies lower than the average value of these two and were analyzed further. Those whose geometrical structures were similar, i.e., those with root-mean-square (rms) deviations of 8.0 Å or less, were grouped together. Six families of configurations emerged. The configuration having the lowest Coulombic energy in each family was further refined, as described below.

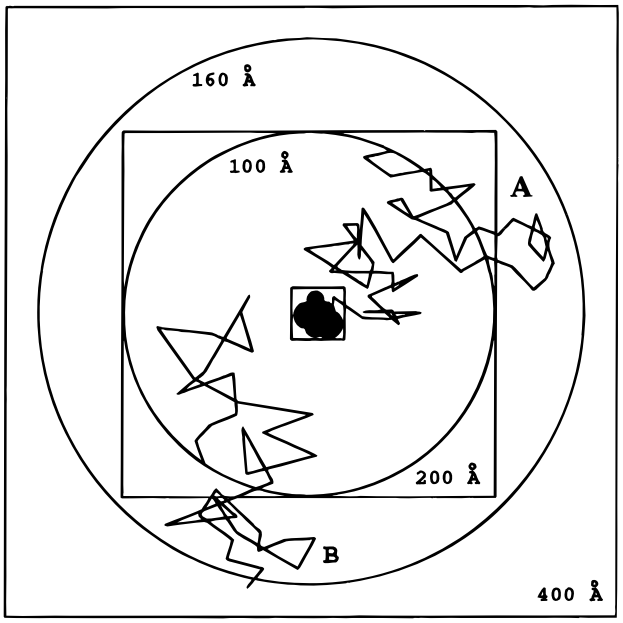


Figure 1. Monte Carlo procedure for finding predocking configurations. The energy of plastocyanin in a Coulombic field of cytochrome *f* (at the center) was calculated on a finer grid within the inner cube (having an edge of 200 Å) and on a coarser grid elsewhere in the outer cube (having an edge of 400 Å). Trajectories started randomly on the inner sphere (having a radius of 100 Å). For example, trajectory A ended unproductively when plastocyanin exited the outer sphere (having a radius of 160 Å), and trajectory B ended when plastocyanin found a local minimum of energy near cytochrome *f* without entering the excluded volume.

Molecular Dynamics Simulation. At this stage conformational flexibility was imparted to each of the six electrostatically-favorable configurations, and hydration was treated explicitly. The simulations were done with the program CHARMM.[59] A sphere with a diameter of 30 Å was placed at the geometrical center of plastocyanin and filled with 3587 water molecules; those that overlapped with either protein were removed. Water was treated according to the model TIP3P.[63]

(57) Moore, J. M.; Lepre, C. A.; Gippert, G. P.; Chazin, W. J.; Case, D. A.; Wright, P. E. *J. Mol. Biol.* **1991**, *221*, 533−555.
(58) Mishra, K. C.; Mishra, S. K.; Das, T. P. *J. Am. Chem. Soc.* **1983**, *105*, 7729−7735.
(59) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(60) Bashford, D.; Karplus, M. *Biochemistry* **1991**, *29*, 10219−10225.
(61) Beroza, P.; Fredkin, D. R.; Okamura, M. Y.; Feher, G. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 5804−5808.
(62) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H. *J. Chem. Phys.* **1953**, *21*, 1087−1092.

Since the structure of plastocyanin had been determined in solution, by NMR spectroscopy, no water molecules were included in it. A total of 114 water molecules found in the crystal structure of cytochrome *f* were included; those that overlapped with protein atoms after energy minimization were removed. Water was contained in the sphere by a suitable potential.[64] The whole plastocyanin and the part of cytochrome *f* inside the sphere were completely flexible, subject only to the potential defined by the force field. Each atom of cytochrome *f* outside the sphere was gently fixed with a harmonic potential defined by the force constant of 0.05 kcal·mol^{-1}·Å^{-1}. The Coulombic and Lennard-Jones interactions of each atom were calculated with so-called group switch boundary conditions. At distances greater than 11.0 Å no potential existed, from 11.0 to 9.0 Å the potential was gradually applied, and at distances less than 9.0 Å the full potential was applied. The water molecules within a 2.5-Å layer below the sphere surface were coupled to a heat bath at 300 K. In Langevin dynamics the friction coefficient was $\beta = 8$ ps^{-1}; it accounts for the removal of energy from the system. The lengths of the covalent bonds between heavy atoms and hydrogen atoms were constrained with the SHAKE algorithm,[65] so that the intergration could be done in 2-fs steps. Each of the six configurations was simulated for 200 ps at 300 K and then cooled by coupling to heat baths at 200, 100, and 0 K. The simulations continued for 20 ps at each of these lower temperatures. Finally, the energy of each configuration was minimized. Each of them was then quantitatively analyzed, as described below.

Energetics of Association. At least four factors contribute to the change in free energy of the protein pair upon association, ΔG_D, according to eq 3. Changes in solvation energy come from electrostatic,

$$\Delta G_D = \Delta\Delta G_R + \Delta G_{NE} + \Delta G_C + \Delta G_{TRC} \quad (3)$$

ΔG_R, and nonelectrostatic, ΔG_{NE}, interactions; the latter depends only on the shapes of the protein molecules. There are also changes in Coulombic electrostatic energy, ΔG_C. The term ΔG_C is simply energy calculated with the Coulomb law, using a single dielectric constant (for a uniform medium). The term G_R, which is calculated with the Poisson−Boltzmann equation, includes G_C, but contains an additional term if more than one dielectric medium is present. In calculations of solvation energy, with a thermodynamic cycle, the Coulombic terms cancel out in the difference. Loss of translational, rotational, and conformational mobility is represented by ΔG_{TRC}. Each of these four terms is the energy difference between the docked complex pc/cytf on the one side and the noninteracting plastocyanin and cytochrome *f* on the other. The last term in eq 3 is likely to have similar values for all the configurations of the docked complex; this is assumed in studies of this kind. Since we are interested in differences among the configurations, i.e., in their relative stabilities, neglect of this term is justifiable. The remaining three terms can be estimated with the help of various approximations.

The electrostatic contribution to the solvation energy arises from polarization of the medium by the charges in the protein. This contribution can be calculated by solving numerically the Poisson−Boltzmann equation.[66−71] The energy required to bring a molecule with the dielectric constant ε_s from a medium with the constant ε_m to a medium with the constant ε_s is given in eq 4. The energy of a molecule within its reaction field is defined in eq 5, in which q_i are the atomic charges in the molecule and ϕ_i is the electrostatic potential at the position of the charge *i* inside a dielectric cavity with a dielectric constant ε_s is a medium with a constant ε_i such that ε_i is ε_m or ε_s. The cavity is defined by the water-accessible surface of the molecule. The potentials ϕ_i can be obtained by solving numerically the Poisson−Boltzmann equation or its linearized form.

$$\Delta G_R = G_R(\varepsilon_s,\varepsilon_m) - G_R(\varepsilon_s,\varepsilon_s) \quad (4)$$

$$G_R(\varepsilon_s,\varepsilon_i) = (1/2)\sum_i q_i\phi_i(\varepsilon_s,\varepsilon_i) \quad (5)$$

The nonelectrostatic contribution to the solvation energy is caused by van der Waals interactions between the solute and the solvent and also by creation or enlargement of the cavity for the solute against the solvent pressure. This, the second term in eq 3, can be estimated by the empirical eq 6, in which ΔA is the change in the water-accessible

$$\Delta G_{NE} = a + b\Delta A \quad (6)$$

surface upon protein docking. The water-accessible surface (A) is defined by rolling a sphere with a diameter of 1.4 Å over the protein molecule. The quantity ΔG_{NE} is negative because this surface decreases in the docking. The parameters *a* and *b* are empirically fitted.[72] Since we are interested in relative, not absolute, energies of various docking configurations, we justifiably neglected the parameter *a*, which is common to all of them. The parameter *b* was set at 5.0, 6.8, or 20.0.[6,72,73]

Coulombic energy arises from pairwise interactions among all the atoms, within each protein and between the two proteins. It can be calculated according to eq 7, in which *i* and *j* run over all the atoms.

$$G_C(\varepsilon_s) = (1/2)\sum_{i\neq j}\frac{q_i q_j}{\varepsilon_s \tau_{ij}} \quad (7)$$

We also define electrostatic energy (ΔG_E) as the sum of the so-called reaction field and Coulombic energy (eq 8) and the total energy (ΔG_T) as the sum of all three terms, i.e., of the electrostatic and nonelectrostatic contributions (eq 9). In eqs 4, 5, and 7, ε_m was set at 80.0 for water

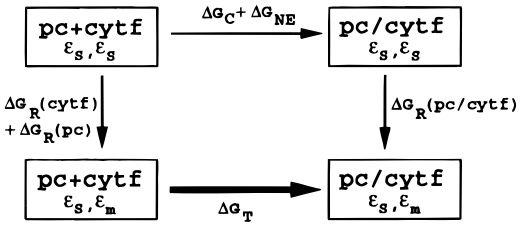
$$\Delta G_E = \Delta\Delta G_R + \Delta G_C \quad (8)$$

$$\Delta G_T = \Delta\Delta G_R + \Delta G_C + \Delta G_{NE} = \Delta G_E + \Delta G_{NE} \quad (9)$$

and ε_s was set at 1.0, 2.0, or 4.0 for the protein matter. Since the nuclear rearrangement was treated explicitly, the value 2.0 probably is the most realistic one. In eqs 4 and 5, the ionic strength was set at 0.10 M and a Stern layer of 2.0 Å was used.

Thermodynamic Cycle. Numerical solution of the Poisson−Boltzmann equation requires moving charge distribution discretely on the grid. The resulting artifact is termed the grid energy. It precludes direct calculation of the binding energy by comparing separate and associated proteins, along the thick horizontal line in Scheme 1. This

Scheme 1. Thermodynamic Cycle



artifact can be avoided by calculating the other three steps in the thermodynamic cycle. We calculated the energies required to transfer the separate proteins from a medium with the dielectric constant of water to a medium with the dielectric constant of the protein (the vertical on the left) and the energy required to transfer each of the six optimized configurations of the diprotein complex in the reverse direction (the vertical on the right). The grid energies canceled out, and we obtained reliable quantities for the two "vertical" processes in Scheme 1. Because in the calculation of G_C the dielectric constants of the solute

(63) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impley, R.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926−935.
(64) Brooks, C. L.; Karplus, M. *J. Chem. Phys.* **1983**, *79*, 6312−6325.
(65) Ryckaert, J.; Ciccotti, G.; Berendsen, H. *J. Comput. Phys.* **1977**, *23*, 327−341.
(66) Warwicker, J.; Watson, H. C. *J. Mol. Biol.* **1982**, *186*, 671−679.
(67) Klapper, I.; Fine, R.; Sharp, K. A.; Honig, B. H. *Proteins* **1986**, *1*, 47−59.
(68) Gilson, M. K.; Sharp, K. A.; Honig, B. H. *J. Comput. Chem.* **1987**, *9*, 327−335.
(69) Gilson, M. K.; Honig, B. H. *Proteins* **1988**, *4*, 7−18.
(70) Honig, B.; Nicholls, A. *Science* **1995**, *268*, 1144−1149.
(71) Sharp, K. A.; Honig, B. *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 301−332.
(72) Ben-Naim, A. *Curr. Opin. Struct. Biol.* **1994**, *4*, 264−268.
(73) BIOSYM. *DelPhi User Guide*; 1993.

and the medium are equal, this term was calculated with the Coulomb law and a single dielectric constant. Estimation of ΔG_{NE} completed the cycle and allowed an estimation of ΔG_T. This procedure was repeated for each of the six configurations of the diprotein complex; the structures of the separate proteins were kept the same.

The conformations of the individual proteins and configuration of the diprotein complex were flexible in this study. Other thermodynamic cycles, which are applicable to rigid structures,[5] are inapplicable here because the grid energy would not be exactly canceled out.

Our calculations were done with the program package MEAD 1.1.3.[60,74−76]

Iron−Ligand Bonding in Cytochrome *f*. The coordinates of the heme were taken directly from the protein structure. All the peripheral substituents in the porphyrin were retained; the two cysteine side chains that form covalent bonds to the porphyrin were represented with thio ether groups; the imidazole group of histidine and the terminal amino group axially coordinated to iron were represented with imidazole and methylamine, respectively, and both propionate groups were reasonably deprotonated. The electronic structure of the ferrous complex was calculated by the Extended Hückel method.[77,78]

Electronic Coupling between the Redox Sites. Both the theoretical basis and the algorithm for the Pathways method are described in detail elsewhere.[32−36,79,80] We kept in mind its assumptions and approximations but also its utility and soundness, established in many of the previous applications.[81−84] Here we briefly explain only the salient features of the method.

The rate constant (k) for the electron-transfer reaction in eq 2 is proportional to the square of the tunneling matrix element (T_{DA}) between ferroheme as the donor (D) and the cupric site as the acceptor (A) and to the density of vibrational states weighed by their Franck−Condon factors (the exponent) (eq 10). The symbols λ and $\Delta G°$, respectively,

$$k = \frac{2\pi}{\hbar}|T_{DA}|^2 \frac{1}{(4\pi\lambda RT)^{1/2}} \exp\left(-\frac{(\Delta G° + \lambda)^2}{4\pi\lambda RT}\right) \qquad (10)$$

represent the reorganizational energy and Gibbs free energy of the reaction.[85,86] An electron tunneling path is a combination of interacting covalent bonds, hydrogen bonds, and van der Waals contacts (interactions through space) that link the donor and the acceptor. The respective decay parameters for attenuation of electronic coupling via these bonds and contacts are the unitless quantities ϵ_C, ϵ_H, and ϵ_S, defined in eq 11

$$\epsilon = \alpha \exp[-\beta(r - r_{eq})] \qquad (11)$$

and calculated with the standard parameters α, β, and r_{eq}; r is the distance between the interacting atoms. (Note the difference between the symbol ε in eqs 4, 5, and 7 and the symbol ϵ in eq 11.) Coupling within aromatic rings of heme, histidine, phenylalanine, tyrosine, and tryptophane and within the guanidinium group of arginine was defined in two ways. In one, the bonds were treated as usual covalent bonds ($\epsilon = 0.6$). In the other, the enhanced coupling was recognized by neglecting the attenuation ($\epsilon = 1.0$).

The tunneling matrix element for a single path (t_{DA}) is proportional to the relative coupling, according to eq 12. The total tunneling matrix

$$t_{DA} \propto \prod_i \epsilon_C(i) \prod_j \epsilon_H(j) \prod_k \epsilon_S(k) \qquad (12)$$

element (T_{DA}) in eq 10 includes the elements t_{DA} for various possible paths. The difficult problem of interactions among paths, i.e., of the relation between T_{DA} and t_{DA}, has not yet been solved. Since the ϵ parameters in eq 12 are ultimately based on experimental results, these parameters implicitly account for the multiplicity of paths. Usually, however, one or a few paths dominate the overall coupling.

We recognize that electronic coupling via a given path depends on the degree of covalency of the iron−ligand and copper−ligand bonds included in that path. As in our previous study, in which this treatment was introduced into the Pathways algorithm,[47] we scale the relative coupling coefficients (γ) for the contributions of the relevant ligands (L) to the redox molecular orbitals of the electron donor (D) and acceptor (A);[87,88] see eq 13. This scaling is based on the reasonable

$$t_{DA} \propto \gamma_{DL}^2 \gamma_{AL}^2 \prod_i \epsilon(i) \qquad (13)$$

assumption that the expansion coefficients are independent of the relative coupling. The values for γ were 0.68 for Cys84, 0.11 for His37, 0.11 for His87, and 0 for Met92 in plastocyanin[47,89] and 1.00 for the porphyrin ligand and 0.6 for each axial ligand in cytochrome *f*. Because γ values are essentially the expansion coefficients in linear combinations of orbitals, their squares are proportional to orbital overlap and electron density.

The couplings were calculated with the program Greenpath 0.973.[90] Figures 2−6 were drawn with the program MOLSCRIPT.[91]

Results and Discussion

Experimental Studies of the Electron-Transfer Reaction between Plastocyanin and Cytochrome *f*. Kinetic effects of chemical modification[92−94] and of site-directed mutagenesis[95−97] in plastocyanin indicate that this protein uses its acidic patch, and Tyr 83 in particular, for binding (eq 1) and the electron-transfer reaction (eq 2) with cytochrome *f*. These processes, however, are quite intricate. Replacement of Leu12 by various amino acids seems to affect the association constant, whereas neutralization of a negative charge in the mutant Asp42Asn seems not to, even though position no. 12 lies in the hydrophobic patch and position no. 42 lies in the acidic patch. Morever, mutation Phe35Tyr in the hydrophobic patch, not far from Leu12, appears not to affect the association constant.[95] Mutations at position no. 12 may affect the reaction indirectly, by perturbing the redox potential of the nearby copper site.[44] Conclusive analysis of kinetic effects of mutation requires direct observation of the intracomplex electron-transfer reaction in eq 2; this was achieved at low ionic strength.[51,52] Effects of

(74) Bashford, D.; Gerwert, K. *J. Mol. Biol.* **1992**, *224*, 473−486.
(75) Bashford, D.; Case, D. A.; Dalvit, C.; Tennant, L.; Wright, P. E. *Biochemistry* **1993**, *32*, 8045−8056.
(76) You, T.; Bashford, D. B. *Biophys. J.* **1995**, *69*, 1721−1733.
(77) Hoffmann, R. *J. Chem. Phys.* **1963**, *39*, 1397−1412.
(78) Zerner, M.; Gouterman, M.; Kobayashi, H. *Theor. Chim. Acta* **1966**, *6*, 363−400.
(79) Betts, J. N.; Beratan, D. N.; Onuchic, J. N. *J. Am. Chem. Soc.* **1992**, *114*, 4043−4046.
(80) Regan, J. J.; Risser, S. M.; Beratan, D. N.; Onuchic, J. N. *J. Phys. Chem.* **1993**, *97*, 13083−13088.
(81) Casimiro, D. R.; Wong, L.-L.; Colon, J. L.; Zewert, T. E.; Richards, J. H.; Chang, I.-J.; Winkler, J. R.; Gray, H. B. *J. Am. Chem. Soc.* **1993**, *115*, 1485−1489.
(82) Casimiro, D. R.; Richards, J. H.; Winkler, J. R.; Gray, H. B. *J. Phys. Chem.* **1993**, *97*, 13073−13077.
(83) Wuttke, D. S.; Bjerrum, M. J.; Winkler, J. R.; Gray, H. B. *Science* **1992**, *256*, 1007−1009.
(84) Jacobs, B. A.; Mauk, M. R.; Funk, W. D.; Gillivray, R. T. A. M.; Mauk, A. G.; Gray, H. B. *J. Am. Chem. Soc.* **1991**, *113*, 4390−4394.
(85) Marcus, R. A.; Sutin, N. *Biochim. Biophys. Acta* **1985**, *811*, 265−322.
(86) Canters, G. W.; van de Kamp, M. *Curr. Opin. Struct. Biol.* **1992**, *2*, 859−869.
(87) Newton, M. D. *J. Phys. Chem.* **1988**, *92*, 3049−3056.
(88) Newton, M. D. *Chem. Rev.* **1991**, *91*, 767−792.
(89) Lowery, M. D.; Guckert, J. A.; Gebhard, M. S.; Solomon, E. I. *J. Am. Chem. Soc.* **1993**, *115*, 3012−3013.
(90) Regan, J. J. *Greenpath Software V0.973*; San Diego, 1995.
(91) Kraulis, P. J. *J. Appl. Crystllogr.* **1991**, *24*, 946−950.
(92) Anderson, G. P.; Sanderson, D. G.; Lee, C. H.; Durell, S.; Anderson, L. B.; Gross, E. L. *Biochim. Biophys. Acta* **1987**, *894*, 386−398.
(93) Gross, E. L.; Curtiss, A. *Biochim. Biophys. Acta* **1991**, *1056*, 166−172.
(94) Christensen, H. E. M.; Conrad, L. S.; Ulstrup, J. *Biochim. Biophys. Acta* **1992**, *1099*, 35−44.
(95) Modi, S.; Nordling, M.; Lundberg, L. G.; Hansson, Ö.; Bendall, D. S. *Biochim. Biophys. Acta* **1992**, *1102*, 85−90.
(96) Modi, S.; He, S.; Gray, J. C.; Bendall, D. S. *Biochim. Biophys. Acta* **1992**, *1101*, 64−68.
(97) He, S.; Modi, S.; Bendall, D. S.; Gray, J. C. *EMBO* **1991**, *10*, 4011−4016.

mutation on bimolecular rate constants, determined at intermediate ionic strength, can be partitioned into contributions from the two steps in eqs 1 and 2, but this partitioning may be uncertain. A small but intriguing dependence of the electron-transfer rate constant on ionic strength may be due to a mismatch between thermodynamic stability and redox activity of a diprotein complex formed at low ionic strength and a rearrangement at higher ionic strength.[98] Alternative explanations are conceivable. The small dependence may be due to a reaction of the diprotein complex with free plastocyanin or with cytochrome *f*. An explanation in terms of different contributions by the net charges, local charges, and dipoles to the electrostatic interaction energy at different ionic strengths has been offered.[99] A similar explanation can probably be attempted by an extension of the van Leeuwen theory[100] including higher multipoles. A very recent study of plastocyanin mutants claimed that the upper cluster of anionic residues (nos. 59−61) is not involved in the electron-transfer reaction, but that the lower cluster (nos. 42−45) is.[101]

The aforementioned studies show that intricacy of investigating the problem of association and reaction of plastocyanin and cytochrome *f*. The acidic patch in the former and the basic patch in the latter are important for the reaction. It is not clear, however, whether the prominent residue Tyr83 is involved in the docking, in the reaction, or in both. We will discuss this question below.

Covalency of Copper−Ligand and Iron−Ligand Bonds. The electronic structure of the cupric site in plastocyanin is relatively well understood.[46,102−105] The short and highly-covalent bond between the copper(II) atom and the thiolate anion of Cys84 provides strong electronic coupling to the adjacent Tyr83, a residue between the two anionic clusters at the acidic patch. The ligand His87, partially exposed at the hydrophobic surface, also provides a good path for electronic tunneling to the copper(II) atom. Indeed, both surface sites are well coupled electronically with the copper atom. We will discuss later their roles in electron tunneling.

Our simple calculation, by the Extended Hückel method, seems to be the first quantum-chemical study of the unusual heme complex found in cytochrome *f*. The several highest filled molecular orbitals have similar energies. The HOMO is delocalized over the porphyrin ring; the three molecular orbitals just below it are mostly composed of the iron 3d orbitals. These high-lying molecular orbitals have very small contributions from the two axial ligands. See the Supporting Information, Table 1 and Figures 1−5. Our finding that electron transfer in cytochrome *f* involves mainly the π-electron system agrees with similar findings by others concerning other cytochromes.[106,107] Indeed, a porphyrin-to-iron charge-transfer transition is observed spectroscopically.[108]

Table 2. Six Configurations of the Diprotein Complex That Emerged from Monte Carlo Calculations, Molecular Dynamics Simulations with Allowance for Flexibility and with Inclusion of Water, and Energy Minimization

configuration	Cu−Fe distance (Å)	interacting side chains	
		anions in pc	cations in cytf
A	34	Glu59, Glu60, Glu61	Lys187, Lys185
		Asp44, Glu43, Glu45	Lys65, Lys66
B	31	Glu59, Glu60	Lys185, Lys187
		Glu43	Arg209, Lys45
C	37	Glu59	Lys187
		Asp45, Glu43	Lys65, Lys66, Lys58
D	14	Glu43, Asp42, Asp44, Glu45	Lys187, Arg209
		Glu59, Glu60	Lys65, Lys58
E	20	Asp44, Glu45, Asp42, Glu43	Lys187, Arg209
		Glu59, Glu60, Glu61	Lys58, Lys65, Lys66
F	35	Asp42, Asp44, Glu43	Lys58, Lys65
		Glu59, Glu60, Glu61	Arg209, Lys187

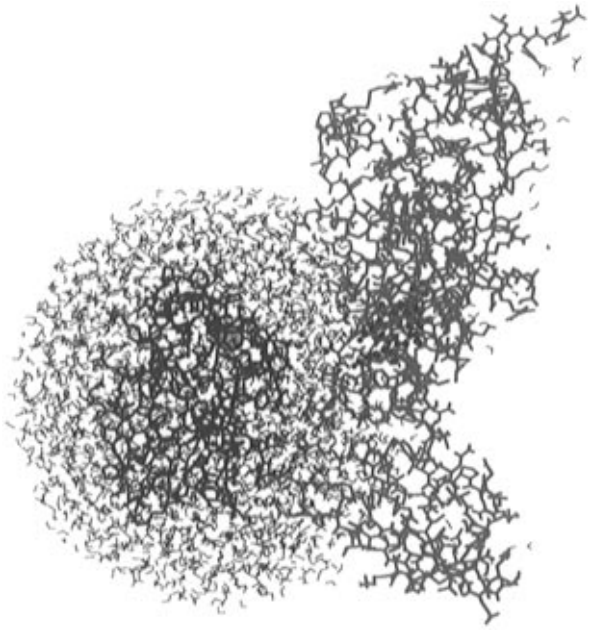


Figure 2. One of the six configurations of the diprotein complex between plastocyanin (blue) and cytochrome *f* (red) that emerged from the Monte Carlo procedure. A sphere with a radius of 15 Å was filled with water (black), and molecular dynamics simulation was performed.

Six Configurations of the Diprotein Complex. The series of calculations culminating in Table 2 began with 32 000 Monte Carlo trajectories obtained with rigid proteins. Approximately 2000 of them ended at local minima of energy. Of these, 140 configurations were relatively stable and were clustered into six families on the basis of structural similarity. The most stable member of each family was subjected to a thorough molecular dynamics simulation, in which the protein molecules were hydrated and given conformational flexibility. Finally, the energy of each configuration was minimized. They are designated A−F.

The criterion for a salt bridge between a carboxylate anion in plastocyanin and an ammonium cation or a guanidinium in cytochrome *f* was the O···N distance of 3.2−3.6 Å. Such interactions are found only in configuration F, between Asp42 and Lys65 and between Asp44 and Lys65. This scarcity of salt bridges is understandable because they are energetically less favorable than hydration of both ions.[109−111] In simulations in

(98) Meyer, T. E.; Zhao, Z. G.; Cusanovich, M. A.; Tollin, G. *Biochemistry* **1993**, *32*, 4552−4559.
(99) Watkins, J. A.; Cusanovich, M. A.; Meyer, T. E.; Tollin, G. *Protein Sci.* **1994**, *3*, 2104−2114.
(100) van Leeuwen, J. W. *Biochim. Biophys. Acta* **1983**, *743*, 408−421.
(101) Lee, B. H.; Hibino, T.; Takabe, T.; Weisbeek, P. J.; Takabe, T. *J. Biochem. (Tokyo)* **1995**, *117*, 1209−1217.
(102) Penfield, K. W.; Gay, R. R.; Himmelwright, R. S.; Eickman, N. C.; Norris, V. A.; Freeman, H. C.; Solomon, E. I. *J. Am. Chem. Soc.* **1981**, *103*, 4382−4388.
(103) Solomon, E. I.; Baldwin, M.; Lowery, M. D. *Chem. Rev.* **1992**, *92*, 521−542.
(104) Penfield, K. W.; Gerwirth, A. A.; Solomon, E. I. *J. Am. Chem. Soc.* **1985**, *107*, 4519−4529.
(105) Larson, S.; Broo, A.; Sjölin, L. *J. Phys. Chem.* **1995**, *99*, 4860−4865.
(106) Nakagawa, H.; Koyama, Y.; Okada, T. *J. Biochem. (Tokyo)* **1994**, *115*, 891−897.
(107) Stuchebrukhov, A. A.; Marcus, R. A. *J. Phys. Chem.* **1995**, *99*, 7581−7590.
(108) Gadsby, P. M. A.; Thomson, A. J. *J. Am. Chem. Soc.* **1990**, *112*, 5003−5011.
(109) Hendsch, Z. S.; Tidor, B. *Protein Sci.* **1994**, *3*, 211−226.

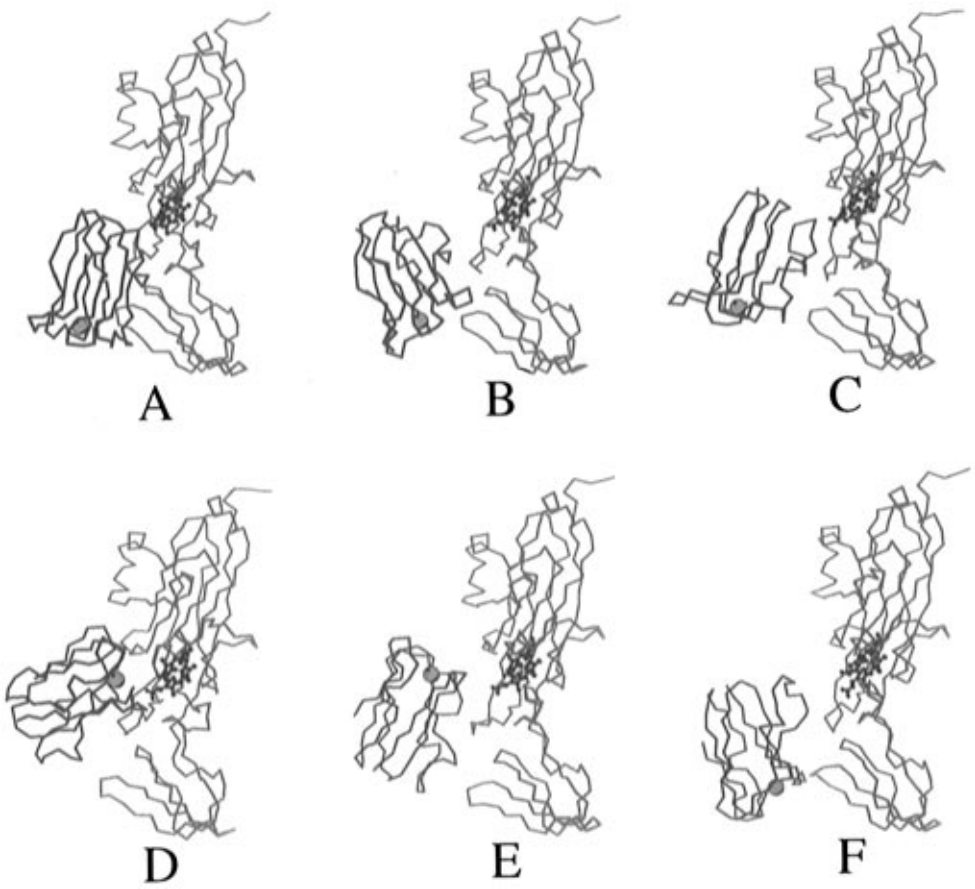


Figure 3. Optimized six configurations of the diprotein complex between plastocyanin (blue) and cytochrome *f* (red) that emerged from molecular dynamics simulations for 260 ps and energy minimization. In the simulations, hydration was treated explicitly and conformational flexibility was allowed. The copper atom and the heme are highlighted.

which the solvent water is neglected salt bridges are commonly found, and their contribution to the stability of the complex is usually overestimated. In our simulations, in which water is explicitly treated, salt bridges occur only in buried regions of the protein interface. The ionic side chains located in the exposed regions prefer to be hydrated, when they are allowed to.

Despite the scarcity of salt bridges there are numerous electrostatic attractions, listed in Table 2. Two oppositely-charged ions were considered interacting if the carbon atom of the carboxylate ion approached the nitrogen atom of the ammonium or guanidinium cation at less than 8.0 Å. This distance includes the radii of the ions and the thickness of the water layer.

Only in configurations D and E is the copper−iron distance shorter than 30 Å. As Figure 3 shows, in the other four configurations the copper atom points away from the heme.

In configuration B, the prominent residue Tyr83 lies outside of the protein interface and is not involved in docking. In configuration A, this residue lies at the edge of the interface. In the remaining four configurations, C−F, this residue is buried in the interface. In three of them Tyr83 appears to form a hydrogen bond with the following residues of cytochrome *f*: Arg209 in configurations D and F and Lys65 in configuration E. Because Arg209 is not conserved in amino acid sequences of cytochromes *f*, only the interaction between Tyr83 and Lys65 will be discussed in detail below.

Cytochrome *f* contains three loops in the region that abuts plastocyanin: Lys185−Gly189, Pro208−Glu212, and Leu61−Lys65. These sections of the protein chain show large temperature factors in the crystal structure, an indication of mobility. Molecular dynamics simulations revealed that association with plastocyanin causes slight reorientation of these loops in nearly all of the configurations. The antiparallel β-sheet Gly157−Asn167 in configuration D swings by as much as 8 Å. Molecular dynamics simulations showed no significant structural changes in plastocyanin in any of the six configurations.

Table 3. Energies (kJ/mol) of the Complex between Plastocyanin and Cytochrome *f* Calculated with $\varepsilon_s = 2.0$ cal/Å^2 and $b = 20.0$ cal/Å^2

configuration	$\Delta\Delta G_R$ (reaction field)	ΔG_C (Coulombic)	ΔG_{NE} (nonelectrostatic)	$\Delta G_E{}^a$ (electrostatic)	$\Delta G_T{}^b$ (total)
A	1483	−2851	−120	−1368	−1488
B	2112	−3975	−94	−1863	−1957
C	2572	−4289	−116	−1717	−1833
D	2676	−4394	−157	−1718	−1875
E	2713	−4585	−107	−1872	−1979
F	3139	−4752	−99	−1613	−1712

^{a} $\Delta G_E = \Delta\Delta G_R + \Delta G_C$. ^{b} $\Delta G_T = \Delta\Delta G_R + \Delta G_C + \Delta G_{NE} = \Delta G_E + \Delta G_{NE}$.

Energetics of the Docking and Analysis of the Stability of the Six Configurations. Different contributions to the energy of the diprotein complex are defined in eqs 4 and 6−9 and shown in Table 3. These energy terms at ionic strengths of 0.10 M (commonly used in experiments in vitro) and 0.30 M (prevalent in the cell) differ only slightly, generally by much less than 1.0%. Only the former results are shown. None of the trends and relative stabilities in Table 3 depend on ionic strength. No single component of the docking interaction correlates with the calculated total energies, ΔG_T. The stabilities

(110) Dougherty, D. A. *Science* **1996**, *271*, 163−168.
(111) Waldburger, C. D.; Schildbach, J. F.; Sauer, R. T. *Struct. Biol.* **1995**, *2*, 122−128.

Table 4. Best Electron-Tunneling Paths from Fe(II) to Cu(II) in Six Configurations, Shown in Figure 3, of the Plastocyanin/Cytochrome *f* Complex Calculated for Two Levels of Hydration with Two Parametrizations of the Coupling within the Aromatic Rings (ϵ Values) and Considering the Metal−Ligand Covalency as Isotropic or Anisotropic

pc/cytf configuration	10^{20}(relative coupling)^2 for the best path					
	no H_2O, isotropic		H_2O, isotropic		H_2O, anisotropic	
	$\epsilon = 0.6$	$\epsilon = 1.0$	$\epsilon = 0.6$	$\epsilon = 1.0$	$\epsilon = 0.6$	$\epsilon = 1.0$
A	310	3.9×10^3	310	3.9×10^3	65	825
B	3.6×10^{-3}	0.078	8.1	62	1.2×10^{-3}	8.9×10^{-3}
C	7.6	79	7.6	79	0.2	2.2
D	1.8×10^{11}	1.4×10^{12}	2.4×10^{12}	2.4×10^{12}	2.2×10^9	1.7×10^{10}
E	7.7×10^6	5.9×10^7	2.8×10^7	1.5×10^8	4.4×10^5	3.4×10^6
F	44	950	44	950	1.2	26.5

of different configurations can be properly analyzed only by recognizing the interplay of the different components.

If only the Coulombic energy were considered, configuration F would be the most stable one. When, however, the change in the reaction field is taken into account, this configuration becomes distinctly unfavorable. This destabilizing contribution of the reaction field may be due to the presence of two salt bridges, discussed above. This example clearly shows the peril of analyzing protein complexes solely, or mostly, in terms of Coulombic interactions even when the proteins are highly charged. Although this approach to molecular modeling remains popular,[11,12] non-Coulombic contributions to electrostatic energy should be considered as well.

As Table 3 shows, the nonelectrostatic term is less than 10% of the electrostatic term, but it too should be taken into account. It makes a significant contribution to the total energy of configuration D.

Configuration B has a relatively small interface, a sign of loose packing together of the two proteins. Consequently, its nonelectrostatic term is the smallest of all in Table 3. Since Tyr83 is not involved in docking and since the metal atoms are far apart, this configuration probably is unimportant for the electron-transfer reaction. Configurations A and C, which have the longest copper−iron distances, are likewise of less interest.

The most stable configuration, E, owes its stability largely to the very favorable electrostatic energy. This finding is consistent with kinetic experiments, which showed a marked dependence of the rate constants for the reaction in eq 2 on ionic strength.[51,98]

Interactions of Tyr83 with Cytochrome *f* Residues. The hydroxyl group of Tyr83 in plastocyanin emerges from several simulations as an acceptor in hydrogen bonds. The donor, from cytochrome *f*, is Lys65 in the most stable configuration, designated E. Because this residue is positively charged, we were intrigued by the possibility that the putative hydrogen bond is in fact an interaction between a cation and the aromatic ring, a so-called cation−π interaction.[110,112]

Pyramidal complexes between aromatic molecules and cations such as Ag^+ have long been known. These surprisingly strong, noncovalent interactions are being emphasized in the latest studies of enzyme−substrate interactions and of molecular recognition in synthetic host−guest systems.[110,113] A pair of hydrated ions is more stable than a salt bridge, but a single cation is more stable in a complex with an aromatic molecule than hydrated.[110,112] Unfortunately, the state of the art in molecular mechanics calculations is inadequate for a correct description of cation−π interactions; their energies are greatly underestimated.[110−112] Satisfactory force fields must include contributions from polarization, induced dipoles, dispersion forces, charge transfer, and possibly other interactions and processes.[115,116] Such calculations are only beginning to be done, and only with small molecules.[114] Applications to proteins, let alone structural optimization of protein complexes, are challenges for the future.

The probability of this strong interaction between plastocyanin and cytochrome *f* impelled us to a broader analysis of our findings about association (here) and about electron transfer (below). The observed 40-fold decrease of the bimolecular rate constant upon the mutation Tyr83Leu[95] is consistent with a decrease in the binding affinity. Attribution of this decrease, wholly or in part, to a changed electron-transfer ability is a matter of kinetic analysis, which is further complicated by the possibility of the rearrangement of the diprotein complex.

Electron-Tunneling Paths. The method Pathways is especially applicable to electron-transfer systems in which the main consideration is the nature of the matter between the donor and the acceptor, not solvation and other effects. We applied it, in the first detailed study of this kind, to the complex between plastocyanin and cytochrome *c*, in which both redox sites are enclosed in the protein matter.[47] As in this previous study, we now consider trends, not absolute values, in the quantities t_{DA}^2 for various paths.

Table 4 shows that configuration D, which has the shortest copper−iron distance, also has by far the strongest electronic coupling between the metal sites, i.e., the most efficient path. Next comes configuration E, with a longer distance and a less strong coupling. The other four configurations seem to be unfavorable for electron transfer. Inclusion of water slightly enhances the coupling in configurations D and E and greatly enhances it in configuration B. The small interface in this configuration (see below) benefits from hydration; the best, but still inefficient, path includes three water molecules. Generally speaking, paths via multiple water molecules are unlikely because the positions of these molecules must be simultaneously favorable for electron transfer to occur.

Because of the approximations in the Pathways method, even the relative magnitudes of the couplings in Table 4 must be taken cautiously. More efficient paths may be discovered by more rigorous calculations. We sought additional paths by widening the search to include relative couplings lower than those along the best path in each of configurations D and E. The results are given in the Supporting Information, Table 2. The best of all paths, which occurs in configuration D, is shown in Figure 4. There is a van der Waals contact between a propionate group of the heme and Pro86 in plastocyanin. Depending on the treatment of copper(II)−ligand bonding, this

(112) Kumpf, R. A.; Dougherty, D. *Science* **1993**, *261*, 1708−1710.
(113) Sussman, J.; Silman, I. *Curr. Opin. Struct. Biol.* **1992**, *2*, 721−729.
(114) Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4177−4178.
(115) Kim, K. S.; Lee, J. Y.; Lee, S. J.; Ha, T.-K.; Kim, D. H. *J. Am. Chem. Soc.* **1994**, *116*, 7399−7400.
(116) Lee, J. Y.; Lee, S. J.; Choi, H. S.; Cho, S. J.; Kim, K. S.; Ha, T.-K. *Chem. Phys. Lett.* **1995**, *232*, 67−71.

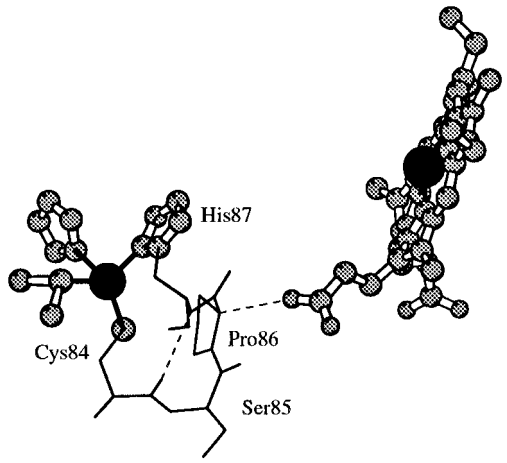


Figure 4. The most efficient electron-tunneling path between the cupric site in plastocyanin (left) and the ferroporphyrin group in cytochrome *f* (right) calculated by the Pathways method. This path was found in configuration D of the diprotein complex.

path can take somewhat different approaches to the copper(II) atom, always within the short segment 84−87. If anisotropy of this bonding is ignored, the path goes via covalent bonds, through His87. If the strong coupling of Cys84 to copper(II) is recognized, the path goes via Ser85 and Cys84. This example shows the intricacies of analyzing electron-tunneling paths at their beginnings and ends, near the donor and acceptor sites.

Comparison of Configurations D and E. Configuration E emerged as the most stable one (Table 3), whereas configuration D turned out to be the most reactive one with respect to internal electron transfer (Table 4). Their relative stabilities probably are not due to interactions involving Arg209, which are present in both configurations. Because of the importance of this difference for the analysis of the reaction in eq 2, we checked whether the relative stabilities change when parameters in the energy calculations are changed.

The dielectric properties of proteins depend on reorientation of permanent and induced dipoles. Much has been written about the dielectric constant.[117−119] The value $\varepsilon_s = 4.0$ is appropriate if ionic residues are treated as point charges and if nuclear rearrangement is neglected. Because this value is used most often,[70,117,120] we used it also even though we assigned partial charges to all atoms and explicitly treated nuclear rearrangement. In a calculation of pK_a values of proteins the best results were obtained with a dielectric constant of 20.0.[121] This high value may have been needed because the side chains were rigid, so that the states with different charges had the same structure; a high dielectric constant corrected this problem. Because this problem was absent in our calculations, in which the proteins were conformationally flexible, we did not test this high value. Atomic charges in the CHARMM program are adjusted so that the dielectric contstant of 1.0 is used in molecular dynamics simulations. In cases such as the so-called special pair in the photosynthetic reaction center, when the protein environment is rigid, the value $\varepsilon_s = 1.0$ best reproduces certain experimental results.[122,123] Therefore, we also used this value. With the aforementioned parametrization of charges, CHARMM implicitly recognizes reorganization of induced, but not of permanent, dipoles. Since both of these effects contribute nearly equally to the value of the dielectric constant, the value $\varepsilon_s = 2.0$ seemed the more realistic. The results in Table 3 were obtained with this value. Three different values of the parameter *b* in eq 6 were tested.

Table 5. Differences in Electrostatic, Nonelectrostatic, and Total Energies (kJ/mol) between Configuration D, Having the Best Heme−Copper Coupling, and Configuration E, Having the Lowest Total Energy

protein dielectric constant, ε_s	*b* in eq 6	$\Delta\Delta G_E$	$\Delta\Delta G_{NE}$	$\Delta\Delta G_T$
1.0	5	323	−12	311
	6.8	323	−17	306
	20	323	−50	273
2.0	5	154	−12	142
	6.8	154	−17	137
	20	154	−50	104
4.0	5	72	−12	60
	6.8	72	−17	55
	20	72	−50	22

Results of these exploratory calculations are shown in Table 5. The variation of the parameters ε and *b* did not change the main finding: The configuration having the best electronic coupling between the copper and heme sites (D) is different from the configuration with the greatest affinity for protein association (E).

Possible Involvement of Tyrosine 83 and of Cationic Side Chains in Electron Transfer. It is important to keep in mind the approximations embodied in the Pathways method. In it and in the other methods for estimating electronic coupling an effective two-state Hamiltonian is used to describe the interactions between the donor and the acceptor. This description becomes invalid if the electronic states of the "bridging" groups (those interposed between the donor and the acceptor) strongly interact with the donor state or with the acceptor state. The description in terms of the superexchange mechanism fails also when the electron-transfer reaction involves an intermediate.[28,29,85]

If, as discussed above, a cationic side chain in cytochrome *f* and the aromatic ring of Tyr83 in plastocyanin form a special bond, then the LUMO of this so-called cation−π complex may act as an electron acceptor, so that a radical intermediate is formed in the course of electron transfer from the ferroheme to the cupric site. Indeed, recent quantum-chemical calculations showed that interaction of a σ^* orbital of an ammonium cation and a π orbital of benzene plays an important role in stabilizing this pair.[115,116] The LUMO is delocalized over the whole complex and is well suited to accept the electron in the putative intermediate.

To our knowledge, a radical of unmodified Tyr83 has not been detected in studies of electron-transfer reactions. Most of these studies, however, were done with reducing agents that cannot form the special bond to the aromatic ring of Tyr83. The physiological partner, cytochrome *f*, may be able to. In the few studies of the reaction in eq 2, radical intermediates were not considered, let alone sought. A short-lived intermediate may be possible, and this question is worthy of experimental study in the future.

If a radical intermediate is involved, then analysis by the Pathways method must be done in two parts—from ferroheme to Tyr83 and from Tyr83 to the cupric site. Because the state of the art in molecular mechanics is inadequate for description

(117) Harvey, S. C. *Proteins* **1989**, *5*, 78−92.
(118) Warshel, A.; Russel, S. T. *Q. Rev. Biophys.* **1984**, *17*, 283−422.
(119) Warshel, A.; Aquist, J. *Annu. Rev Biophys. Biophys. Chem.* **1991**, *20*, 267−298.
(120) Gilson, M. K. *Curr. Opin. Struct. Biol.* **1995**, *5*, 216−223.
(121) Antosiewicz, J.; McCammon, A. J.; Gilson, M. K. *J. Mol. Biol.* **1994**, *238*, 415−436.
(122) Ullmann, G. M.; Muegge, I.; Knapp, E.-W. In *The Reaction Centers of Photosynthetic Bacteria*; Michel-Beyerle, E. M., Ed.; Springer-Verlag: Berlin, 1996; pp 143−155.
(123) Muegge, I.; Apostolakis, J.; Ermler, U.; Fritsch, G.; Lubitz, W.; Knapp, E.-W. *Biochemistry* **1995**, *35*, 8359−8370.

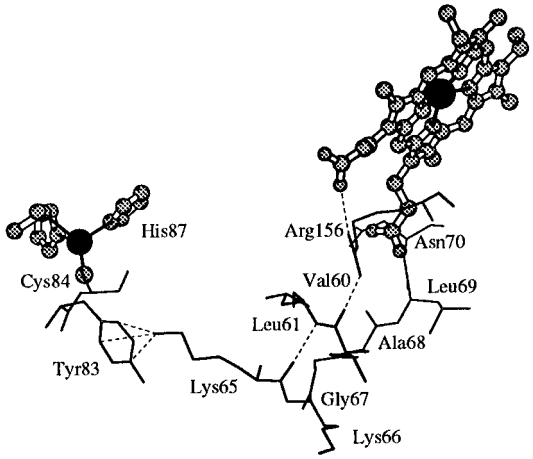


Figure 5. Two special electron-tunneling paths between the cupric site in plastocyanin (left) and the ferroporphyrin group in cytochrome *f* (right) found in configuration E of the diprotein complex. The ammonium cation of Lys65 is shown above the aromatic ring of Tyr83, in so-called cation$-\pi$ interaction. One path (through residues 156, 60, 61, and 65) involves the hydrogen bonds (dashed lines), whereas the other path (residues 70−65) leads through the protein backbone.

of interactions between cations and aromatic rings (see above), we did not limit our analysis to optimized configurations, results of complete molecular dynamics simulations for 260 ps. We also considered the evolving structures of the diprotein complex at the earlier stages of simulation and the actual structures of plastocyanin and cytochrome *f*. The main result of this analysis is the interesting pattern shown in Figure 5 and in the Supporting Information, Table 2.

A tunneling path starts at the iron atom and goes through the porphyrin ring, via a salt bridge involving the propionate group in the pyrrole ring D and the guanidinium group of Arg156, via another hydrogen bond to the oxygen atom of Val60, via the peptide bond to Leu61, then to the oxygen atom of Lys65, and then to the ammonium cation in the side chain that presumably interacts with the aromatic ring of Tyr83. This pattern is present in the early stages of simulation of configuration E but disappears after approximately 80 ps. Since, however, the aforementioned hydrogen bonds are evident in the crystal structure of cytochrome *f*,[49] we believe that their disappearance from the simulation is caused by the inability of the force field to reorganize the special interaction of Lys65 and Tyr83. Instead of simulating this interaction, the force field simulates others that are more tractable, such as the attraction of Lys65 to the acidic patch in plastocyanin; see Table 2. From this point of view, "diversion" of Lys65 somewhat strains this residue and residues bound to it; consequently the aforementioned hydrogen bonds and the path requiring them are disrupted. Then a path along the backbone of cytochrome *f*, from Asn70 to Lys65, becomes relatively favorable, with a coupling of approximately 10% of the previous one; see Figure 5.

This diprotein system, and likely others in which interactions between cationic and aromatic side chains may occur, should be thoroughly investigated in the future. Quantum-mechanical calculations should be combined with classical mechanical simulations based on much-improved force fields for analysis of these newly-recognized interactions.

Unreactivity of the Covalent Diprotein Complex. In the presence of carbodiimide, direct amide bonds form between lysine side chains in cytochrome *f* and carboxylate groups in glutamate or aspartate side chains in plastocyanin.[50] Structures and redox properties of the active sites of these proteins are not significantly perturbed,[50b,c] but the intracomplex electron-transfer reaction (eq 2), which is fast in the noncovalent complex, becomes undetectably slow in the covalent complex.[52]

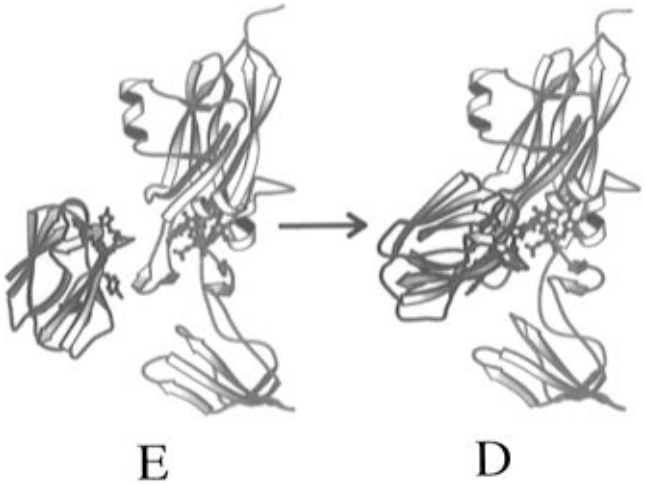


Figure 6. Rearrangement of the diprotein complex between cupriplastocyanin (blue) and ferrocytochrome *f* (red) that may be involved in the intracomplex electron-transfer reaction. Configuration E has the highest binding affinity, whereas configuration D provides the best electronic coupling between the redox sites, which are highlighted.

This study offers two possible explanations of this interesting finding. One is that direct cross-links (without any tethers) make the covalent complex rigid and preclude the rearrangement, shown in Figure 6, from the most stable configuration (E) into the most reactive one (D). This rearrangement is possible in the case of the noncovalent complex, which is flexible. The other explanation is that electron transfer in the noncovalent diprotein complex occurs in the configuration with the highest binding affinity (E), but via the proposed special bond between the ammonium cation of Lys65 and the aromatic ring of Tyr83, shown in Figure 5. In the presence of carbodiimide this noncovalent bond may be broken and Lys65 forced into covalent cross-linking with a proximate acidic group in plastocyanin. Indeed, the residues Glu59 and Glu60, which have been implicated in covalent cross-linking between the two proteins,[50b,c] lie close to Lys65, see Table 2. This diversion of Lys65 would disrupt the electron-transfer path and neutralize the cation required for the stabilization of the anion radical of Tyr 83.

We plan an experimental study to answer the important questions that arose from this theoretical study.

Conclusion

Protein docking can be modeled qualitatively, as in a very recent analysis of cytochrome *f* and plastocyanin.[124] In our study the simulation was done in stages, by Monte Carlo calculations followed by molecular dynamics calculations, for a relatively long period of time; in this study this period was 260 ps, at different temperatures. The protein molecules can be given much conformational freedom, and hydration can be treated explicitly. The structures obtained by thorough simulations can be analyzed in terms of various interactions that contribute to the binding energy. Possible artifacts can be sought, and ruled out, by critically comparing the results obtained with different reasonable values of the adjustable parameters. Only then can simulated structures be analyzed in functional terms and the theoretical results used to explain experimental ones.

The fact[52] that structurally noninvasive covalent cross-links in the diprotein complex pc/cytf abolish the electron-transfer reaction in eq 2 can be explained in two ways. First, the rigid cross-links may prevent the rearrangement shown in Figure 6.[52]

(124) Pearson, D. C., II; Gross, E. L.; David, E. *Biophys. J.* **1996**, *71*, 64−76.

Subsequent determination of the molecular structure of cytochrome *f*[49] gave credence to this hypothesis. Similar dynamics has been proposed for cytochrome *c* and cytochrome *c* peroxidase.[125] Second, Lys65 may be captured in a cross-link and thus diverted from its putative interaction with Tyr83, proposed in this study.

Our theoretical analysis of electron-tunneling paths agrees with a recent biochemical study[126] that Tyr1, an axial ligand to the iron atom, is not involved in electron-transfer paths.

Acknowledgment. This work was supported by the Deutsche Forschungsgemeinschaft SFB 312, Project D7, and by the U.S. National Science Foundation, through Grant MCB-9222741. G.M.U. thanks Boehringer Ingelheim Fonds for a fellowship. We thank Edward I. Solomon and Louis B. LaCroix for charge distribution in the active site of cupriplastocyanin, Jeffrey J. Regan for the program Greenpath 0.973, Donald Bashford for the program MEAD 1.1.3, Daniel Hoffmann for help and consultations, Brian M. Hoffman for a preprint of ref 125, and Mandy Colditz for assistance with the paper.

Supporting Information Available: Tables of HOMO and several filled molecular orbitals below it in ferrocytochrome *f* and results of "removal" from best paths certain amino acid residues or just their side chains and figures of ferrocytochrome *f* (MO numbers 144, 143, 142, 141, and 140) (16 pages). See any current masthead page for ordering and Internet access instructions.

JA962237U

(125) Nocek, J. M.; Zhou, J. S.; Forest, S. D.; Priyadarshy, S.; Beratan, D. N.; Onuchic, J. N.; Hoffman, B. M. *Chem. Rev.* **1996**, *96*, 2459−2489.
(126) Zhou, J.; Fernández-Velasco, J. G.; Malkin, R. *J. Biol. Chem.* **1996**, *271*, 6225−6232.