Electron-Transfer Reactions in Proteins: An Artificial Intelligence Approach to Electronic Coupling

Prabha Siddarth and R. A. Marcus*

Arthur Amos Noyes Laboratory of Chemical Physics,† California Institute of Technology, Pasadena, California 91125

Received: August 13, 1992; In Final Form: October 26, 1992

The electronic interactions which are responsible for electron transfer in proteins are treated using an artificial intelligence (AI) approach and a quantum mechanical formulation of superexchange. An AI search procedure is devised to select the most important amino acid residues which mediate long-range electron transfer. All the valence orbitals of these amino acid residues are used in a diagonalization of the "bridge" orbitals. The electronic coupling matrix element is then calculated by using second-order perturbation theory to couple the bridge orbitals to the donor and acceptor orbitals. The relative values of the electronic coupling elements obtained with this model are shown to be in good agreement with experimental results for cytochrome *c* derivatives, without use of adjustable parameters. Further, an optimum path calculation in which the path consists of several amino acids is also presented and compared with the many amino acid calculation. The various results show that not merely the separation distance but also the nature of the protein medium separating the redox centers is an important factor in controlling the rate of these electron-transfer reactions.

I. Introduction

Recently, there has been considerable interest in long-range electron-transfer (ET) reactions in biological systems,[1-3] which have been shown to proceed rapidly over large (10–15-Å edge-to-edge) distances. Many experimental studies, using techniques such as site-directed mutagenesis and semisynthetic modification, have focused on electron-transfer reactions within proteins.[4-7] A key issue in this context is the dependence of the ET rate on the intervening medium between the electron donor and the electron acceptor. In the present paper, we investigate this question theoretically and explore how the protein structure controls the electronic coupling in electron-transfer reactions in modified proteins.

These long-range electron-transfer reactions in proteins are nonadiabatic. In electron-transfer theory, the rate constant for nonadiabatic transfer of an electron from a donor to an acceptor can be expressed in terms of a golden rule type expression, namely, as the product of the square of an electronic coupling matrix element (H_{DA}) and a nuclear Franck–Condon factor (FC):[8]

$$k_{ET} = \frac{2\pi}{\hbar}|H_{DA}|^2 FC \qquad (1)$$

The Franck–Condon factor has been treated quantum mechanically for the higher frequency motions, semiclassically, and in the classical limit for low-frequency motions. Various expressions for FC are given, for example, in a recent review.[9] A purely classical form of the FC, for example, is

$$FC = \frac{1}{(4\pi\lambda k_B T)^{1/2}} \exp\left(-\frac{(\Delta G^\circ + \lambda)^2}{4\lambda k_B T}\right) \qquad (2)$$

where ΔG° is the standard free energy of the reaction for the donor–acceptor pair DA for electron transfer at a fixed DA separation distance R and λ is a reorganizational term which contains both "solvent", which includes the protein and vibrational contributions.

H_{DA} is treated quantum mechanically and is influenced by the distance and structure of the medium separating the electron donor and acceptor. For small systems, H_{DA} has been calculated by using all-electron ab initio SCF methods[10,11] as well as by more approximate methods.[12,13] However, the application of the SCF methods to large biological systems such as proteins would be computationally very intensive and currently impractical. We have previously studied[14] electron transfer in modified myoglobin derivatives using a superexchange formalism. Other approaches include path integral methods,[15] pathway analyses,[16,17] and a perturbative form of superexchange theory.[18] The path integral method of Kuki and Wolynes[15] is instructive and makes use of a pseudopotential for the protein. In its present form, it includes electron tunneling and neglects hole transfer. In the pathway analysis of Beratan, Betts, Onuchic, and co-workers,[16] the electronic coupling is estimated by the number of links between the donor and the acceptor, each of which, in turn, is estimated by certain empirical decay factors for each type of link. Ulstrup and co-workers[18] have applied superexchange theory, treating perturbatively a chosen sequence of amino acid residues to calculate electronic coupling to various surface sites in plastocyanin. They used one molecular orbital per amino acid in their expression.

In our earlier treatment of electronic coupling in proteins,[14] we made use of a simple search procedure to select the "important" amino acids for electron transfer. Using these amino acids as the bridge, the bridge molecular orbitals and eigenvalues were then determined and H_{DA} was calculated from the interaction of the donor and acceptor orbitals with the fully diagonalized bridge orbitals. The molecular orbital eigenvectors and eigenvalues were obtained using extended Hückel theory. (The latter has been found to give reasonable *relative* values of H_{DA} in both synthetic model donor–acceptor systems[13,19] and in modified proteins,[14] as judged by comparison with "experimentally" determined values.)

In the present paper, we introduce an artificial intelligence (AI) search algorithm to choose the important amino acids. Depending on the cytochrome *c* derivative, some 10–20 amino acids are selected from the approximately 100 available. These amino acids are then used as the relevant subset of protein to calculate the electronic coupling. This procedure automatically incorporates all possible amino acid paths within this subset and all possible interactions between these paths. We also calculate H_{DA} using only one single path, the path being composed of amino acids. For this calculation, we choose the best path in our AI search and compare the results of this single-path calculation

† Contribution No. 8699.

0022-3654/93/2097-2400$04.00/0

with those obtained from the more complete (containing the many more amino acids) calculation.

II. Theory: An Artificial Intelligence Search

In the present paper, we treat electron transfer in ruthenium-modified cytochrome *c* proteins. Even a small protein such as cytochrome *c* has about 100 amino acids and over 800 heavy atoms. Not only would a full calculation involving all these atoms not be feasible, but it would possibly be even unnecessary. It is intuitively clear that amino acid residues far away from both the donor and the acceptor, as well as from the line of centers of the donor and the acceptor, should have relatively little effect on the electronic coupling matrix element. In that case, it is sufficient to consider a suitably chosen subset of the approximately 100 amino acid residues.

The task then is to select a small number of "important" amino acid residues. This selection process can be phrased in terms of a search problem: Given the starting node (the electron donor D in the protein) and the goal (the electron acceptor A), what are the most important paths connecting these two sites? There are two standard search techniques, namely, the depth-first search and the breadth-first search,[20] which can be applied to find paths from D to A. In a depth-first search, the search is conducted from the starting node as far as it can be performed until either a "dead end" or the goal is encountered. If a dead end is encountered, the search is backed up one level and continued. Breadth-first is a form of exhaustive search in which each of the nodes is systematically explored at each level of the search until the goal is attained. It is usual to use a heuristic function to guide the search along each step. For example, costs or distances between each intermediate node is a common heuristic function. For our particular application, the intermediate nodes are chosen to be the various protein atoms, and instead of costs, we can assign electronic couplings H_{MN} between these protein atoms. Presently we use, but only for the purpose of the AI search, the following simplified expression for H_{MN} between two atoms M and N along a path:

$$H_{MN} = \max\left|KS_{mn}\left(\frac{\epsilon_m + \epsilon_n}{2}\right)\right| \tag{3}$$

where m denotes an atomic orbital on atom M, n denotes an atomic orbital on atom N, S_{mn} is the overlap integral between m and n, ϵ_m and ϵ_n are the orbital energies, and K is a constant. The expression within the absolute value signs in eq 3 is the well-known Wolfsberg–Helmholtz formula[21] for the resonance integral between two atomic orbitals. (K is usually taken to be 1.75.[22]) H_{MN}, as calculated above, depends directly on the overlap between atomic orbitals and thus takes into account the effect of mutual orientations of amino acids.

Once the electronic couplings between all pairs of atoms in the protein are calculated, we can formulate a search from the donor to the acceptor using a best-first strategy: At each step of the process, the most promising of the nodes is selected by applying the heuristic function (in our case, the value of the electronic coupling, as given by H_{MN}). This method combines the advantages of both a depth-first and a breadth-first search into a single method.[23]

A heuristic function based on H_{MN} alone, defined as above, results in a *blind* or an *uninformed* search, since the intermediate nodes that are selected are unaffected by information concerning the goal state or the unexplored region between the explored states and the goal state. For a true AI search, also known an an *informed search*, knowledge, of the unexplored region is made part of the heuristic or the evaluation function. We presently use an informed search which exploits knowledge about the unexplored domain to find paths from D to A.

In the electron-transfer problem, for each intermediate node I, the true measure of electronic coupling from the starting point D to the node I is estimated from the product of the appropriate H_{MN}'s divided by an energy denominator. Let this quantity be denoted as V_{DI}:

$$V_{DI} = H_{D1}\frac{H_{12}}{\Delta E}\frac{H_{23}}{\Delta E}\cdots\frac{H_{I-1I}}{\Delta E} \tag{4}$$

In the above expression, we have used identical energy denominators. Since eq 4 is only used as part of the evaluation function in the AI search and not in the final calculation, approximations such as replacing the exact ΔE's by some mean value are minor, except when the actual ΔE_{IJ} is of the order of an H_{IJ}, and in that case replacing that $H_{IJ}/\Delta E_{IJ}$ by unity would be adequate.

If we can now estimate the coupling from node J to the goal, then this knowledge can be incorporated into the evaluation function. It has been well documented in the literature that the rate of falloff of electronic coupling with distance is approximately exponential,[8] $H_{MN} = H^\circ_{MN}\exp(-\beta R/2)$, where $\beta \approx 1$ Å^{-1} and H°_{MN} is the electronic coupling between atoms M and N at their contact distance. Therefore, we choose to estimate the coupling from any given node to the goal as

$$T_{IA} = C\exp(-0.5R) \tag{5}$$

where R is the distance between node I and the acceptor A and C is a constant whose value is estimated in footnote 24. Thus, the composite evaluation function used to evaluate the promise of a successor node is

$$\mathrm{EF} = \frac{V_{DI}T_{IA}}{\Delta E} \tag{6}$$

where ΔE is a suitably chosen energy difference,[25] V_{DI} is the actual coupling from the donor to node I, and T_{IA} is an estimate of the coupling from node I to the acceptor. It is in the factor T_{IA} that knowledge about the problem domain is exploited.[26]

In the present implementation of the above algorithm, we have chosen to employ a tier model as was done previously.[14] From the donor atom, a set of successor atoms are selected on the basis of their evaluation function. These atoms constitute the first tier of the search. In the next tier, the most promising atoms from each of these first-tier atoms are selected. Then from each of these tier-2 atoms, the optimal path to the acceptor was found by an iterative procedure: Once a path to the acceptor is found, a net measure of electronic coupling, calculated as the product of H_{MN}'s along the path (divided by energy differences), is assigned to it. Then, other paths are found from the same tier-2 atom to the acceptor, and finally, only that path that, for the particular pair of tier-1 and tier-2 atoms, has the greatest net electronic coupling (evaluation function) is retained.

To capture the amino acids in the vicinity of A democratically with those of D, it is desirable to perform searches in both directions, i.e., starting from the donor and reaching the acceptor, as well as starting from the acceptor and reaching the donor. A pre-set threshold value for the net evaluation function for any path is used to avoid particularly unimportant paths. Since the number of atoms chosen for expansion in the first and second tiers of the search can be varied, there is a greater probability of selecting all the possibly important amino acid residues. Also, it is to be noted that a number of paths from the donor to the acceptor are found, since there are a number of tier-2 atoms. (There are, in the procedure given later and barring duplicates, 15 best D-to-A paths selected and the same number of best A-to-D paths, when the number of tier-1 atoms is chosen to be 5, and for each of these, the number of tier-2 ones is chosen to be 3.)

Thus, the above AI search selects a number of atoms that lie on paths from D to A. The amino acids to which these atoms belong are then considered as the important amino acids for mediating electron transfer in the protein. Since some of the selected amino acids will be bonded to amino acids which are not

selected, there will be "dangling" bonds, which are completed by addition of H atoms.

III. Theory: Calculation of H_{DA}

Once the important amino acids are selected by the AI search, they collectively are considered to be the bridge for mediating ET between D and A. We then diagonalize the bridge orbitals and use second-order perturbation theory to calculate the electronic coupling matrix element by what is known in the literature as the superexchange mechanism.[28,29] Quantitative comparisons and justification for using second-order perturbation theory to treat ET occurring via off-resonant bridge orbitals were described earlier.[13,19] (A comparison of this application of second-order perturbation theory was made with a calculation in which the totality of relevant D, A, and bridge orbitals was used in the diagonalization.)

The transition state for electron transfer occurs at the "intersection" of the energy surface for the reactants with that for the products, i.e., where there is zero change in electronic energy upon electron transfer. (These surfaces are those of electronic energy vs nuclear coordinates.) On this intersection, the energy of the donor orbital E_χ equals that of the acceptor orbital E_μ, and this result is used in obtaining eqs 7 and 8 below and in later remarks in the paper. Equations 7 and 8 are valid only at this intersection.

We recall next the salient features of the superexchange formalism in a one-electron description of the system containing a donor D, a bridge B, and an acceptor A. For the case where B is connected to the donor (acceptor) by only one atomic orbital, H_{DA} is given by[28,30]

$$H_{DA} = T_D T_A \sum_\alpha \frac{C_{D\alpha} C_{A\alpha}}{E_\chi - E_\alpha} \tag{7}$$

where T_D (T_A) is the matrix element for interaction between D(A) and the adjacent atomic orbital of B, $C_{D\alpha}$ ($C_{A\alpha}$) is the coefficient of the αth bridge orbital at the point of contact of B with D(A), E_α is the energy of the αth molecular orbital of B, and E_χ is the energy of the localized molecular orbital of D, which is equal to that of A, E_μ, in the transition state for the electron transfer.

In a more general case, where D and A are connected to B by a number of atomic orbitals, the electron-transfer matrix element is given by[13]

$$H_{DA} = \sum_\alpha \sum_{m_D, m_B, m_A} \frac{(C^*_{\chi m_D} H_{m_D m_B} C_{m_B \alpha})(C^*_{\alpha m_B} H_{m_B m_A} C_{m_A \mu})}{E_\chi - E_\alpha} \tag{8}$$

where χ and μ denote the molecular orbitals on D and A that are involved in the electron transfer, the α's denote the MO's of the bridge B, and m_D, m_B, and m_A are the atomic orbitals of D, B, and A, respectively. $H_{m_D m_B}$ are the interaction matrix elements of the donor and bridge orbitals, and $H_{m_B m_A}$ are the interaction matrix elements of the bridge and acceptor orbitals. We have used the term electron transfer, but it is intended to include, in eq 8, transfer by any occupied orbitals χ of the bridge as well, i.e., hole transfer.

IV. Results and Discussion

The proteins considered here are the ruthenium-modified cytochrome *c* derivatives studied experimentally by Gray and coworkers.[31-33,6] Four singly ruthenated cytochromes ($Ru(bpy)_2$(im)(His*X*), where im = imidazole and *X* = 33, 39, 62, and 72) have been prepared.[31] The $Ru(bpy)_2$(im)(His) complex was first photoexcited, and in the presence of a quencher, the transient Ru^{3+}–Cyt *c*–Fe^{2+} was formed. Intramolecular ET rates from Fe^{2+} to Ru^{3+} were then measured. A distinct advantage of working with these systems is that these ET reactions are nearly activationless, since the driving force of the reaction ($-\Delta G^\circ$ = 0.74 eV) nearly equals the reorganization energy, λ. (The reorganization energy has been estimated[32] to be 0.8 eV for [Re-$(bpy)_2$(im)(His–Cyt *c*)].) Under these conditions, the ET rates are mainly controlled by the electronic matrix element that couples the donor and the acceptor at the transition state.

TABLE I: Amino Acid Residues Selected by the AI Search

His 33	His 39	His 72	His 62
Tyr 67	Met 80	Met 80	Met 64
Met 80	Asn 52	Phe 82	Asp 60
Leu 32	Val 57	Cys 17	Pro 30
Pro 30	Leu 58	Lys 79	Asn 31
Asn 31	Tyr 67	Ile 81	Leu 32
His 26	Trp 59	Pro 71	Ile 35
His 33	Phe 82	Asn 70	Phe 36
Tyr 59	Gly 41	His 72	Trp 59
His 18	Asn 56	Thr 78	Asn 63
Thr 19	Pro 71		His 62
Cys 14	Ser 40		Tyr 74
Ala 15	His 39		Val 57
Lys 22	Ile 35		
Gly 34	Thr 78		
Leu 35	Lys 55		
Arg 38			
Gly 24			
Gly 23			
Glu 21			
Lys 25			

The protein coordinates were obtained from the crystal structure of cytochrome *c*.[34] Using BIOGRAF,[35] the protein structure was then "ruthenated" by adding $Ru(bpy)_2$(im) to the histidines. This modification does not significantly change the cytochrome *c* structure.[31] The resulting structure files are used in the AI search.

In the first tier of the search, five atoms were retained, as noted earlier, and for each of these, three atoms were retained in the second tier. Then from each of these 15 atoms in the second tier, a path to the acceptor was found, as described earlier. The search was performed in the reverse direction also, namely, starting from the acceptor and seeking the donor. In Table I, we give the amino acid residues that were selected by the AI search for each of the four modified cytochrome *c* derivatives. Cytochrome *c* itself has about 100 amino acids and the AI search yielded about 10–20 important amino acids that need to be considered for electron transfer.

We then made a quantum mechanical calculation of the electronic coupling matrix element using, as the bridge, the amino acid residues selected by the search. The molecular orbital coefficients *C*'s and eigenvalues E_α of the bridge as well as the interaction matrix elements *H*'s are obtained using extended Hückel theory.[36] A common problem in most calculations is determining the metal donor (acceptor) orbital energy relative to the bridge orbitals at the transition state. Presently, as discussed in a recent article,[19] we obtain E_χ, the energy of D(A) at the transition state, using the absorption spectrum maximum associated with charge transfer from porphyrin to Fe(III) in ferricytochrome *c*[37,38] (cf. the present Appendix).

The results obtained for the four cytochrome *c* systems are given in Table II. As in earlier papers, it should be stressed that the results from extended Hückel calculations have been more reliable for relative values of H_{DA} than for absolute values. (For example, the relative values of H_{DA} are unaffected by approximations in T_D and T_A in eq 7.) Accordingly, we have given in Table II ratios of the electronic matrix elements, as well as their absolute values. The values of H_{DA} extracted from experimental data, as well as their ratios, are also given in Table II. (The absolute values of the "experimental" H_{DA} may also be model dependent and might change somewhat if a different analysis of the experimental data were used.) Hence, for both reasons, it is

TABLE II: Calculated and Experimental Electronic Coupling Matrix Elements and Their Ratios, for "Full" and Single-Path Treatments

		H_{DA}, cm^{-1}		$H_{DA,i}/H_{DA,1}$	
i (deriv)	R, Å	calc[a]	expt[b]	calc[a]	expt[b]
1 (His 33)	11.1	0.01(0.007)	0.10	1(1)	1
2 (His 39)	12.3	0.01(0.007)	0.11	1(1)	1.1
3 (His 72)	8.4	0.007(0.003)	0.06	0.7(0.4)	0.6
4 (His 62)	14.8	0.002(0.001)	0.006	0.2(0.14)	0.06

[a] The values in parentheses are the results obtained for the single-path calculation. [b] The experimental values are from ref 31.

TABLE III: Amino Acid Residues Used in the Single-Path Calculation

His 33	His 39	His 72	His 62
His 18	Asn 52	Met 80	Met 64
Pro 30	Gly 41	Thr 78	Asp 60
Asn 31	Asn 56	His 72	His 62
Leu 32	His 39		
His 33			

preferable to use the present results for comparisons among the four derivatives than for the absolute values of the electronic coupling elements.

The coupling elements for the His 33 and His 39 derivatives are rather similar, both for theory and experiment. The His 62 derivative with a donor–acceptor separation of 14.8 Å has the smallest coupling element. The striking result is that the His 72 derivative, which even though it has a smaller donor–acceptor separation than the His 33 or His 39 derivatives, has a smaller coupling element than the latter two. This behavior is primarily due to the nature of the protein medium which separates Fe^{2+} and Ru^{3+} in the His 72 derivative, as pointed out earlier by Wuttke et al.[31] There is almost a 2.7 Å cavity in the medium, which substantially reduces the electronic coupling.

In the present study, we also investigated the coupling resulting from considering only one path of amino acids as opposed to a number of such paths. We considered the "best" amino acid path determined from the AI search (i.e., the amino acid path with greatest net electronic interaction as estimated from the evaluation function). The amino acid residues of this single path calculation are given in Table III and are seen to be only 3–5 in number, instead of 10–20. With only these amino acids being considered as the bridge, H_{DA} was calculated, as before, using eq 8. Both the absolute values and the relative values of H_{DA} obtained with this amino acid "path" calculation are given in Table II. It is seen that these values correlate well with the values from the full calculation, indicating that, in the present instance at least, the relatively few amino acids used in the path calculation are sufficient to obtain relative values of electronic coupling. It is again significant that, for the His 72 derivative, the coupling is smaller than that for the His 33 or His 39 derivatives.

Considering the amino acid residues used in the single-path calculation more closely, the His 33, His 39, and His 62 paths are relatively directly linked (Figures 1–3, respectively), while the His 72 path (Figure 4) has a considerably long through-space connection, making the electronic coupling smaller. This result clearly indicates that the nature of the protein medium between the donor and the acceptor can control the electronic coupling and hence the rate of electron-transfer reactions.

It is perhaps worthwhile at this point to compare the present approach with the atom pathway analyses of Beratan and co-workers.[16,17] Our approach to estimating electronic coupling in protein electron-transfer reactions consists of using an AI search to determine the important amino acids and, with these amino acids as the bridge, to use second-order perturbation theory to calculate H_{DA}. Beratan and co-workers also use a search to evaluate pathways, where each chemical bond (all chemical bonds between any two atoms are treated as equivalent), H-bond, or through-space link contributes its own empirical decay factor in the calculation of the electronic matrix elements. The pathway (comprised of atoms rather than amino acids as in our method) with the largest electronic coupling is selected. For the derivatives of cytochrome c discussed in the present paper, this method has given a useful qualitative mapping of electron-transfer pathways and ratios of electronic coupling elements in good agreement with experimental results.[16,31]

We have developed earlier a molecular fragment approach[19] for calculating electronic coupling in large systems by dividing the bridge into smaller molecular fragments. We plan next to

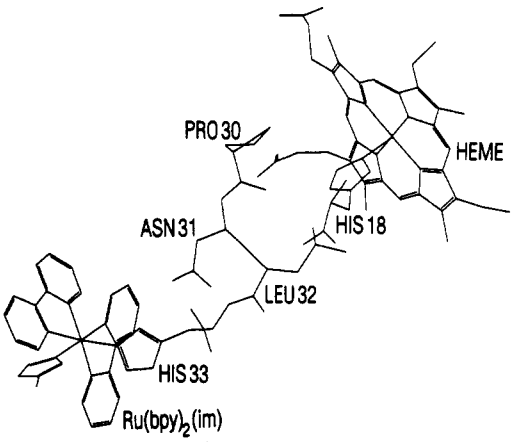

Figure 1. Amino acid residues in the best path for the His 33 derivative of cytochrome c.

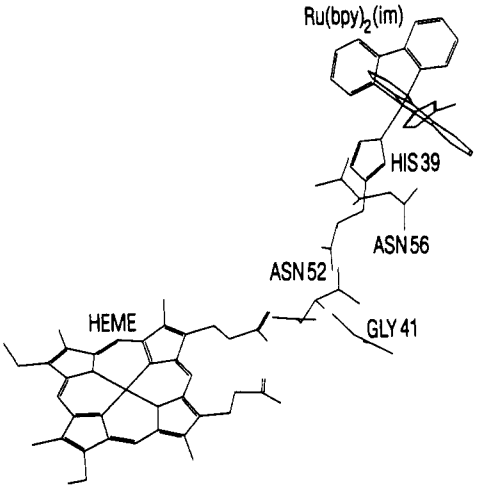

Figure 2. Amino acid residues in the best path for the His 39 derivative of cytochrome c.

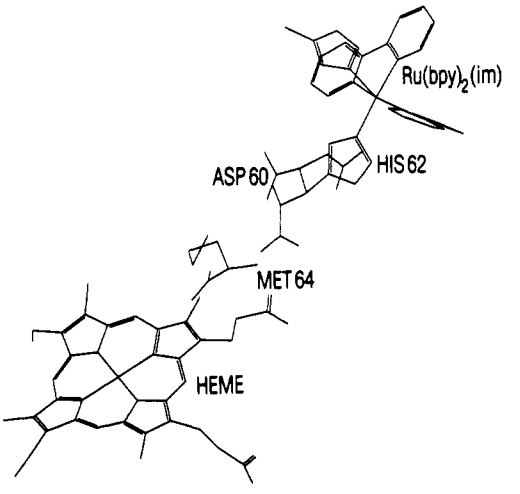

Figure 3. Amino acid residues in the best path for the His 62 derivative of cytochrome c.

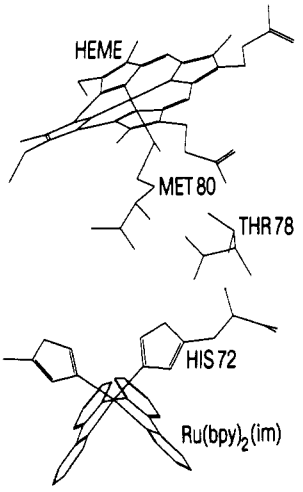

Figure 4. Amino acid residues in the best path for the His 72 derivative of cytochrome *c*.

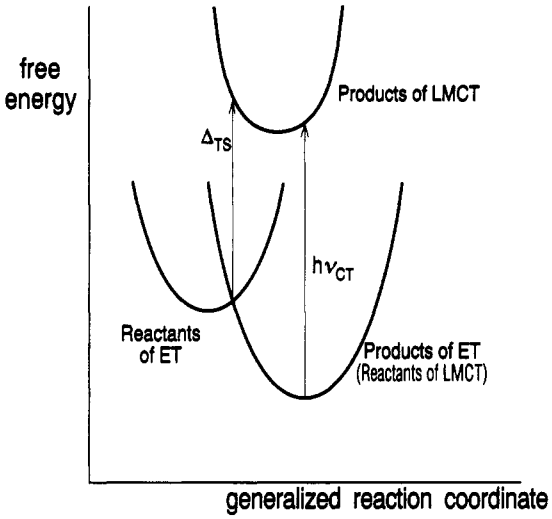

Figure 5. Free-energy curves of reactants and surrounding medium and that of products and surrounding medium.

study electron transfer in proteins using this approach, particularly in these cytochrome *c* derivatives and in the modified myoglobins. That study would enable us to determine the important amino acid paths, in contrast to the atom pathways most commonly used.[16,17] Further, this approach would also permit us to determine the number of orbitals of the amino acids which contribute importantly to the electronic coupling. In the polypeptide systems that we investigated earlier using the molecular fragment approach,[19] typically five molecular orbitals per amino acid were found to contribute significantly to H_{DA}.

V. Conclusions

In the present paper, we have developed an AI search to determine the most important amino acid residues for mediating electron transfer in proteins. This method can be implemented for any protein molecule, and presently, we have considered the ruthenium-modified cytochrome *c* derivatives studied experimentally by Gray and co-workers. Once the search is accomplished, a quantum mechanical method which makes use of all the valence orbitals of the selected amino acid residues is used to calculate values of electronic coupling matrix elements. The results are shown to be consistent with experimental data.

We have also considered an amino acid path calculation and have found that, for these particular proteins, a calculation involving only the limited number of residues along a single path is sufficient to estimate relative values of electronic coupling. However, the full calculation, which considers all potential paths and all possible interactions between these paths, is expected to be the method of choice when more than one path dominates.

Acknowledgment. It is a pleasure to dedicate this article to our friend and colleague, Dudley Herschbach, on the occasion of his 60th birthday. His style, enthusiasm, and choice of problems have been an inspiration to us over these many years. It is a pleasure to acknowledge the support of this research by the National Science Foundation and the Office of Naval Research.

Appendix. Estimation of the Donor (Acceptor) Orbital Energy, E_χ, at the Transition State

In the expression for the electronic matrix element H_{DA}, eq 8, a key quantity is $E_\chi - E_\alpha$, where E_χ is the donor energy level (equal to that of the acceptor E_μ at the transition state for electron transfer) and E_α the energy of a molecular orbital of the bridge. Extended Hückel theory and most other treatments of electronic structure are too approximate to obtain this quantity accurately. Instead, we estimate it using a cytochrome *c* charge-transfer spectrum. We have previously employed[19] a similar method of estimating $E_\chi - E_\alpha$ for a series of polypeptide-bridged systems.

The electron-transfer reaction of interest is

$$\text{ET: Fe(II)–Cyt } c\text{–Ru(III)} \rightarrow \text{Fe(III)–Cyt } c\text{–Ru(II)} \quad \text{(A1)}$$

It has been reported in the literature[37,38] that ferricytochrome *c* exhibits a porphyrin to metal-charge-transfer absorption at 1750 nm. This process can be depicted as

$$\text{LMCT:} \quad \text{Fe(III)–Cyt } c \rightarrow \text{Fe(II)–Cyt } c^+ \quad \text{(A2)}$$

The free-energy surfaces of these two processes can be schematically plotted as in Figure 5. Δ_{TS} denotes the vertical energy difference from the transition state of the electron-transfer reaction to the products of the charge-transfer excitation process and $h\nu_{CT}$ is the absorption maximum associated with this charge transfer.

Using the same procedure as employed in ref 19, the difference $\Delta_{TS} - h\nu_{CT}$ is given by

$$\Delta_{TS} - h\nu_{CT} = -\frac{\lambda_i}{\lambda}(\Delta G^\circ + \lambda) \quad \text{(A3)}$$

where λ_i is the reorganization energy of $Fe^{3+/2+}$–Cyt *c*, λ is the reorganization energy for the entire electron-transfer process (i.e., including the $Ru^{3+/2+}$ contribution), and ΔG° is the free energy of the ET reaction. For the present system, λ_i and λ have been estimated to be[32] about 0.5 and 0.8 eV, respectively, and ΔG° is measured to be[31] –0.74 eV. From the charge-transfer absorption spectrum,[38] $h\nu_{CT}$ is known to be 0.7 eV. Hence, Δ_{TS} is given approximately by 0.67 eV.

Δ_{TS} represents the energy difference $E_\chi - E_\alpha$, where E_α is the energy of the HOMO of the porphyrin from which the charge is transferred in the LMCT. Presently, E_α is calculated (by extended Hückel theory) to be –11.35 eV. Hence, E_χ is estimated to be –10.68 eV (–11.35 eV + Δ_{TS}). The extended Hückel value for the energy of the donor orbital, without correcting for solvational and reorganizational effects of the medium, was –16.26 eV. In the calculations, therefore, the energy of the donor orbital was adjusted upward by this difference, 5.58 eV, so that its new value is now –10.68 eV. Correspondingly, the energy of the acceptor orbital was also adjusted to be –10.68 eV.

References and Notes

(1) *Metal Ions in Biological Systems*; Sigel, H., Sigel, A., Eds.; Marcel Dekker: New York, 1991; Vol. 27.

(2) Long Range Electron Transfer in Biology. *Structure and Bonding*; Palmer, G., Ed.; Springer-Verlag: Berlin, 1991; Vol. 75.

(3) *Electron Transfer in Inorganic, Organic and Biological Systems*; ACS Advances in Chemistry Series No. 228; Bolton, J. R., Mataga, N., McLendon, G., Eds.; American Chemical Society: Washington, DC, 1991.

(4) Bowler, B. E.; Raphael, A. L.; Gray, H. B. *Prog. Inorg. Chem.* **1990**, *38*, 259.
(5) Axup, A. W.; Albin, M.; Mayo, S. L.; Crutchley, R. J.; Gray, H. B. *J. Am. Chem. Soc.* **1988**, *110*, 435.
(6) Bowler, B. E.; Meade, T. J.; Mayo, S. L.; Richards, J. H.; Gray, H. B. *J. Am. Chem. Soc.* **1989**, *111*, 8757.
(7) Michel-Beyerle, M. E.; Plato, M.; Deisenhofer, J.; Michel, H.; Bixon, J.; Jortner, J. *Biochim. Biophys. Acta* **1988**, *932*, 52.
(8) Marcus, R. A.; Sutin, N. *Biochim. Biophys. Acta* **1985**, *811*, 265.
(9) Marcus, R. A.; Siddarth, P. In *Photoprocesses in Transition Metal Complexes, Biosystems and Other Molecules: Experiment and Theory*; Kochanski, E., Ed.; Kluwer: Norwall, MA, 1992, p 49.
(10) Newton, M. D. *Int. J. Quantum Chem., Quantum Chem. Symp.* **1980**, *14*, 363. Logan, J.; Newton, M. D. *J. Chem. Phys.* **1983**, *78*, 4086. Newton, M. D. *J. Phys. Chem.* **1988**, *92*, 3049.
(11) Ohta, K.; Closs, G. L.; Morokuma, K.; Green, N. J. *J. Am. Chem. Soc.* **1986**, *108*, 1319.
(12) Larsson, S.; Volosov, A. *J. Chem. Phys.* **1986**, *85*, 2548. Larsson, S.; Volosov, A. *J. Chem. Phys.* **1986**, *85*, 6623.
(13) Siddarth, P.; Marcus, R. A. *J. Phys. Chem.* **1990**, *94*, 2985.
(14) Siddarth, P.; Marcus, R. A. *J. Phys. Chem.* **1990**, *94*, 8430.
(15) Kuki, A.; Wolynes, P. G. *Science* **1987**, *236*, 1647.
(16) Beratan, D. N.; Onuchic, J. N.; Betts, J. N.; Bowler, B. E.; Gray, H. B. *J. Am. Chem. Soc.* **1990**, *112*, 7915.
(17) Onuchic, J. N.; Beratan, D. N. *J. Chem. Phys.* **1990**, *92*, 722.
(18) Christensen, H. E. M.; Conrad, L. S.; Mikkelsen, K. V.; Nielsen, M. K.; Ulstrup, J. *Inorg. Chem.* **1990**, *29*, 2808.
(19) Siddarth, P.; Marcus, R. A. *J. Phys. Chem.* **1992**, *96*, 3213.
(20) Rich, E. *Artificial Intelligence*; McGraw-Hill: New York, 1983. Barr, A.; Feigenbaum, E. A. *The Handbook of Artificial Intelligence*; HeurisTech: Stanford, 1981; Vol. I. Nilsson, N. J. *Problem-Solving Methods in Artificial Intelligence*; McGraw-Hill: New York, 1971.
(21) Wolfsberg, M.; Helmholtz, L. *J. Chem. Phys.* **1952**, *20*, 837.
(22) Hoffmann, R. *J. Chem. Phys.* **1963**, *39*, 1397; **1964**, *40*, 2474, 7745.
(23) The chief problem of a breadth-first search, if employed at every tier of the search, is a combinatorial explosion. The number of nodes to be explored grows as c^n, where c is the number of branches at each node and n is the number of levels. Even for a 10-fold search with 5 branches at each level, the number of possible nodes is of the order of 10^7. On the other hand, while the depth-first strategy avoids a large number of unproductive intermediate nodes, it can miss the goal when the protein has a complex branch structure. The best-fit technique, by expanding the most promising successor evaluated to date, combines the advantages of these two strategies.
(24) An approximate value for C can be obtained by estimating it for two H atoms as 0.525 au.
(25) Since in this formulation of the search ΔE does not depend on the particular node I, the absolute value chosen for ΔE does not affect the promise of a node but is incorporated here to make the usual connection to second-order perturbation theory clear. The value of ΔE does affect, however, the relative values for two paths which differ in the number of nodes.
(26) This procedure is analogous to the A* algorithm proposed by Hart et al. (see ref 27) where the actual cost of reaching the present node from the starting point is added to an estimate of the additional cost of getting from the current node to the goal state. In the present application, we cannot add these two quantities but combine them, as in eq 6, according to perturbation theory. We note further that in the case when the estimate T_{IA} never underestimates the true (but unknown) coupling from I to A, this algorithm is guaranteed to find the optimal path from D to A, if one exists.
(27) Hart, P. E.; Nilsson, N. J.; Raphael, B. *IEEE Trans. SSC* **1968**, *4*, 100.
(28) Halpern, J.; Orgel, L. E. *Discuss. Faraday Soc.* **1960**, *29*, 32. McConnell, H. M. *J. Chem. Phys.* **1961**, *35*, 508.
(29) Newton, M. D. *Chem. Rev.* **1991**, *91*, 767.
(30) Larsson, S. *J. Am. Chem. Soc.* **1981**, *103*, 4034.
(31) Wuttke, D. S.; Bjerrum, M. J.; Winkler, J. R.; Gray, H. B. *Science* **1992**, *256*, 1007.
(32) Chang, I.-J.; Winkler, J. R.; Gray, H. B. *J. Am. Chem. Soc.* **1991**, *113*, 7056.
(33) Therien, M. J.; Selman, M. A.; Gray, H. B.; Chang, I.-J.; Winkler, J. R. *J. Am. Chem. Soc.* **1990**, *112*, 2420.
(34) The coordinates of horse hearth cytochrome *c* (His 33 and His 72 calculations) (Bushnell, G. W.; Louie, G. V.; Brayer, G. D. *J. Mol. Biol.* **1990**, *214*, 585) and yeast-iso-1 cytochrome *c* (His 39 and His 62 calculations) (Louie G. V.; Brayer, G. D. *J. Mol. Biol.* **1990**, *214*, 527) were kindly provided by G. D. Brayer.
(35) BIOGRAF/III: BIOGRAF was designed and written by S. L. Mayo, B. D. Olafson, and W. A. Goddard, III. It is a product of Biodesign Inc., Pasadena, CA.
(36) QCPE Program No. 517, Indiana University, Bloomington, IN. The basis set and valence-state ionization energies for the metal atoms were obtained from QCPE Program No. 387.
(37) Moore, G. R.; Pettigrew, G. W. *Cytochrome c: Evolutionary, Structural and Physicochemical Aspects*; Springer-Verlag: Berlin, 1990.
(38) Gadsby, P. M. A.; Thomson, A. J. *J. Am. Chem. Soc.* **1990**, *112*, 5003.