of cluster discussed in this paper. In particular, explanations framed in the language of effective potentials for covalent clusters, using analogies with other, simpler systems, provide an alternative way of looking at the structures of these molecules. More quantitative analyses may also be possible in future work. Balanced structures provide a particularly helpful set of examples because we know in advance that they must be stationary points of some order, and the variation of the order with the type of potential may be most illuminating. Furthermore, balanced structures suggest possible structures for new stable species; for example, it would be interesting to investigate icosidodecahedral $B_{30}H_{30}^{2-}$, or even C_{30}, to see if this structure is ever a minimum or a transition state.

Acknowledgment. The computer time for this work was funded in part by a Grant from the National Science Foundation to Prof. R. S. Berry, whose hospitality throughout the author's stay in Chicago under a Lindemann Trust Fellowship is greatly appreciated. Dr. J. F. Stanton provided helpful advice on the use of the ACES optimiser. I am also grateful to the referees for some constructive criticisms of the original manuscript and to Dr. A. J. Stone for several discussions.

Electron-Tunneling Pathways in Ruthenated Proteins

David N. Beratan,*[†,‡] José Nelson Onuchic,[†,‡,§] Jonathan N. Betts,[‡] Bruce E. Bowler,[‡,∥] and Harry B. Gray[‡]

Contribution from the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, Beckman Institute,[⊥] California Institute of Technology, Pasadena, California 91125, Instituto de Física e Química de São Carlos, Universidade de São Paulo, SP, Brazil, and Department of Physics, University of California, San Diego, La Jolla, California 92093. *Received December 8, 1989*

Abstract: We implement a numerical algorithm to survey proteins for electron-tunneling pathways. Insight is gained into the nature of the mediation process in long-distance electron-transfer reactions. The dominance of covalent and hydrogen bond pathways is shown. The method predicts the relative electronic couplings in ruthenated myoglobin and cytochrome *c* consistent with measured electron-transfer rates. It also allows the design of long-range electron-transfer systems. Qualitative differences between pathways arise from the protein secondary structure. Effects of this sort are not predicted from simpler models that neglect various details of the protein electronic structure.

I. Introduction

Long-distance electron transfer in proteins involves electron tunneling through a polypeptide environment.[1] Although donor–acceptor electronic interactions in many proteins are relatively weak, they are not as weak as would be expected in the absence of the polypeptide bridge.[2] Theory suggests that the transfer rate should be sensitive to the molecular details of the tunneling bridge in a weakly coupled donor–acceptor molecule and that a molecular orbital approach is an appropriate one.[3]

Methods to calculate weak bridge-mediated donor–acceptor interactions have been of interest for some time in chemistry.[1,2] We recently developed a model for the dependence of the donor–acceptor coupling and transfer rate on bridge structure in small molecules[2e,4] and proteins.[5] While refinements[5b] are being added to the protein model (questions still remain concerning details of the electronic structure techniques[5c] and the density of important pathways, see next section), the model provides a framework for the interpretation and design of experimental systems.

In this paper we present a numerical implementation of our theoretical model for electron tunneling in proteins. The model divides the bridge that assists the electron-tunneling process into a number of blocks. The decay of the interaction across each block is sufficiently rapid that the overall coupling can be approximated as a product of decays per block where these decays depend only on the details of the particular block and the tunneling energy.[5] This approximation is an oversimplification based in perturbation theory, and strategies to generalize this treatment are discussed in the following sections. The description of the bridge as a combination of identifiable blocks is a useful one that applies even beyond the perturbation theory limit.[6] This is a central theme of our model, and future implementations that take more details of the bridge into account will be based on this description.

These blocks of orbitals between donor and acceptor define pathways that mediate electron transfer. Due to the approximately *exponential* decay of the coupling with the number of bridging groups in a tunneling pathway, one expects relatively few pathways to be important for coupling the donor and acceptor in a given protein or protein–protein complex. If gating of the electron-transfer reaction[7] becomes important, the calculation of the tunneling matrix element for the relevant pathway must be

[†] Jet Propulsion Laboratory, California Institute of Technology.
[‡] Beckman Institute, California Institute of Technology.
[§] University of California, San Diego, and Instituto de Física e Quimica de São Carlos.
[∥] Present address: Department of Chemistry, University of Denver, Denver, CO 80208.
[⊥] Contribution No. 8086.

(1) (a) Marcus, R. A.; Sutin, N. *Biochim. Biophys. Acta* **1985**, *811*, 265. (b) Newton, M. D.; Sutin, N. *Annu. Rev. Phys. Chem.* **1984**, *35*, 437. (c) *Photoinduced Electron Transfer*; Fox, M. A., Chanon, M., Eds.; Elsevier: Amsterdam, 1988; Vols. A–D.
(2) (a) Halpern, J.; Orgel, L. E. *Discuss. Faraday Soc.* **1960**, *29*, 32. (b) McConnell, H. M. *J. Chem. Phys.* **1961**, *35*, 508. (c) Hopfield, J. J. *Proc. Natl. Acad. Sci. U.S.A.* **1974**, *71*, 3640. (d) Larsson, S. *J. Am. Chem. Soc.* **1981**, *103*, 4034. (e) Beratan, D. N.; Hopfield, J. J. *J. Am. Chem. Soc.* **1984**, *106*, 1584. (f) Riemers, J. R.; Hush, N. S. *Chem. Phys.* **1989**, *134*, 323.
(3) Beratan, D. N.; Onuchic, J. N.; Hopfield, J. J. *J. Chem. Phys.* **1985**, *83*, 5325.
(4) (a) Onuchic, J. N.; Beratan, D. N. *J. Am. Chem. Soc.* **1987**, *109*, 6771. (b) Beratan, D. N. *J. Am. Chem. Soc.* **1986**, *108*, 4321.
(5) (a) Beratan, D. N.; Onuchic, J. N.; Hopfield, J. J. *J. Chem. Phys.* **1987**, *86*, 4488. (b) Onuchic, J. N.; Beratan, D. N. *J. Chem. Phys.* **1990**, *92*, 722. (c) Beratan, D. N.; Onuchic, J. N. *Photosynth. Res.* **1989**, *22*, 173.
(6) da Gama, A. A. S. *J. Theor. Biol.* **1990**, *142*, 251.
(7) Hoffman, B. M.; Ratner, M. A. *J. Am. Chem. Soc.* **1987**, *109*, 6237.

0002-7863/90/1512-7915$02.50/0

performed. Methods for including relatively small amplitude thermal fluctuations of groups on the pathways exist.[5a]

We seek to define the residues in specific proteins[8] that provide dominant contributions to the electronic coupling between donor and acceptor. The numerical calculation that we describe is the simplest implementation of our model for bridge-mediated tunneling in proteins. It includes details of the connectivity between atoms. The calculation distinguishes hydrogen bonds, covalent bonds, and through space interactions. However, it ignores differences in covalent interactions. In this paper we present numerical calculations that suggest specific protein bonds in ruthenium-labeled myoglobin and cytochrome *c* that dominantly contribute to the tunneling interaction. Comparisons with experimentally derived coupling matrix elements are made. Future work will probe the coupling between electron localization sites in native proteins[8a] as well as examine differences in the mediation properties of particular protein secondary structures.[9]

II. The Tunneling Pathway Model

Theoretical studies of protein-mediated electron transfer typically address the hamiltonian used for the system or the method of calculating the tunneling matrix element from a given hamiltonian. In earlier papers, we discussed strategies for choosing and improving the hamiltonian[3-5] (e.g., by improving the parameters for through bond and through space coupling). Throughout this paper, we will focus on the question of how to determine the bonds and the interactions that dominantly contribute to the tunneling matrix element for a given hamiltonian. The concept of tunneling pathways is central to this issue.

(a) Physical Pathways. There are two distinct mediation mechanisms that the intervening protein could provide to couple the donor and acceptor. One is to mediate the interaction by a few specific combinations of interacting bonds (fragments of amino acids) between the donor and acceptor. The bonds may couple donor and acceptor through covalent, hydrogen bonding, and noncovalent interactions. Each of these combinations is called a pathway and plays an important role in the donor–acceptor coupling. The actual role played by a specific pathway may be determined experimentally by using molecular biological techniques to modify the pathway.[8,10] The other distinct way that the protein might couple the donor and acceptor involves a sufficiently large number of pathways such that modifying a single pathway in this network will have a very small effect on the net coupling and the rate. In this case, the strategy of identifying specific pathways is nonproductive.

To focus the discussion of the donor–acceptor coupling mechanism, it is useful to introduce the concept of a *physical tunneling pathway*. A physical pathway is defined as a collection of interacting bonds in the medium around and between the donor and acceptor that makes some contribution to the donor–acceptor interaction. For a single physical pathway there are exact and perturbation theory methods for calculating the coupling arising from that physical pathway. The tunneling matrix element can be calculated from the propagation of the donor (acceptor)-localized state at the appropriate nuclear configuration.[11] Numerical strategies (for both exact and perturbation methods) usually write the decay of the wave function as a product of decays per bond (or delocalized group[5b]). Within a perturbation theory calculation, the per bond decay depends only on the tunneling energy and the nature of the particular bonds in the pathway. This method (applied to lowest order) neglects scattering corrections to the wave function propagation in the protein bridge. The scattering corrections for a given pathway arise from enumerations of bonds on the tunneling pathway longer than the shortest path from donor to acceptor. For example, a physical pathway consisting of bonds 1, 2, 3, 4, ... has the direct pathway 1–2–3–4... and the scattering pathways 1–2–3–2–3–4..., etc. The scattering pathways can be accounted for exactly in the electronic coupling calculation for a one-dimensional physical pathway by correcting the self energy of each orbital on the path.[6,9] Scattering pathways are the distinct combinations of bonds that give rise to a large number of additional paths within the same physical pathway. Exact methods, particularly Green's function approaches, often write the coupling as a product as well.[6,9] In this case, the terms in the product explicitly include these scattering corrections. (The simple description presented here is valid when the physical pathways are independent—not intersecting. This approach is adequate for the present survey purposes. The problem is more complicated if the pathways are not simple, such as pathways with loops. In this case, the distinction between physical and scattering pathways is more subtle.[9] Approaches that attempt numerical solution of the Schrödinger equation do not usually break the problem into products of decays across a collection of groups.

In more complex systems that, in addition to being aperiodic and finite, have real connections between pathways, the self-energy correction technique can be generalized.[9] However, our main concern here is to arrive at a qualitative description of the important paths. The goal is to identify important physical pathways, *not* to make firm quantitative estimates of the tunneling matrix elements or study interference between paths. References to pathways in this paper mean physical pathways.

In complex disordered three-dimensional systems, one might expect the number of physical pathways of a given effective length (i.e., coupling strength) to increase as the transfer distance increases. (We use the term effective length because different interactions in the medium contribute differently to the overall coupling.) As the number of pathways increases, the relative contribution of each one decreases with the effective length.[12] Further theoretical work is needed to analyze the growth in the number of physical pathways in a protein vs the relative contribution of each to the tunneling matrix element. The approach here is dictated by the strong suspicion that few pathways dominate. Further theoretical and experimental work in this area is needed. In general, the answer to this question will likely depend on the structure of the intervening protein and the electron tunneling energy. The relatively rapid decay of the tunneling interaction with distance (particularly for through space decay), the relatively low density of residues, and the anisotropic packing of bonds in proteins suggest relatively few significant pathways.[5]

(b) Contributions from a Single Physical Pathway. The rate of electron transfer between weakly coupled localization sites in proteins can be written[11]

(8) (a) Beratan, D. N.; Onuchic, J. N.; Gray, H. B. In *Metal Ions in Biological Systems*; Sigel, H., Sigel, A. Eds.; Marcel Dekker: New York, Vol. 27, in press. (b) Cowan, J. A.; Upmacis, R. K.; Beratan, D. N.; Onuchic, J. N.; Gray, H. B. *Annu. New York Acad. Sci.* **1988**, *550*, 68. (c) Moore, J. M.; Case, D. A.; Chazin, W. J.; Gippert, G. P.; Havel, T. F.; Powls, R.; Wright, P. E. *Science* **1988**, *240*, 314. (d) McLendon, G. *Acc. Chem. Res.* **1988**, *21*, 160. (e) Bowler, B. E.; Raphael, A. L.; Gray, H. B. *Prog. Inorg. Chem.* In press. (f) Mayo, S. L.; Ellis, W. R., Jr.; Crutchley, R. J.; Gray, H. B. *Science* **1986**, *233*, 948. (g) Gray H. B. *Chem. Soc. Rev.* **1986**, *15*, 17. (h) Gray, H. B.; Malmström, B. G. *Biochemistry* **1989**, *28*, 7499. (i) Liang, N.; Pielak, G. J.; Mauk, A. G.; Smith, M.; Hoffman, B. M. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 1249. (j) Liang, N.; Mauk, A. G.; Pielak, G. J.; Johnson, J. A.; Smith, M.; Hoffman, B. M. *Science (Washington, D.C.)* **1988**, *240*, 311. (k) Elias, H.; Chou, M. H.; Winkler, J. R. *J. Am. Chem. Soc.* **1988**, *110*, 429. (l) Conrad, D. W.; Scott, R. A. *J. Am. Chem. Soc.* **1989**, *111*, 3461. (m) Pan, L. P.; Durham, B.; Wolinska, J.; Millett, F. *Biochemistry* **1988**, *27*, 7180. (n) Durham, B.; Pan, L. P.; Long, J. E.; Millett, F. *Biochemistry* **1989**, *28*, 8659. (o) Farver, O.; Pecht, I. *FEBS Lett.* **1989**, *244*, 379. (p) Jackman, M. P.; McGinnis, J.; Powls, R.; Salmon, G. A.; Sykes, A. G. *J. Am. Chem. Soc.* **1988**, *110*, 5880. (q) Osvath, P.; Salmon, G. A.; Sykes, A. G. *J. Am. Chem. Soc.* **1988**, *110*, 7114. (r) Faraggi, M.; Klapper, M. H. *J. Am. Chem. Soc.* **1988**, *110*, 5753. (s) Hazzard, J. T.; McLendon, G.; Cusanovich, M. A.; Das, G.; Sherman, F.; Tollin, G. *Biochemistry* **1988**, *27*, 4445. (t) Cusanovich, M. A.; Meyer, T. E.; Tollin, G. *Adv. Inorg. Biochem.* **1987**, *7*, 37.

(9) (a) Goldman, C. In preparation. (b) Onuchic, J. N.; de Andrade, P. C. P.; Beratan, D. N. In preparation.

(10) (a) Bowler, B. E.; Meade, T. J.; Mayo, S. L.; Richards, J. H.; Gray, H. B. *J. Am. Chem. Soc.* **1989**, *111*, 8757. (b) Therien, M. J.; Selman, M. A.; Chang, I.-J.; Winkler, J. R.; Gray, H. B. *J. Am. Chem. Soc.* **1990**, *112*, 2420.

(11) (a) Onuchic, J. N.; Wolynes, P. G. *J. Phys. Chem.* **1988**, *92*, 6495. (b) Bialek, W.; Bruno, W. J.; Joseph, J.; Onuchic, J. N. *Photosynth. Res.* **1989**, *22*, 15. (c) Onuchic, J. N.; Beratan, D. N.; Hopfield, J. J. *J. Phys. Chem.* **1986**, *90*, 3707.

(12) One might refer to the coupling between donor and acceptor as a free tunneling matrix element. The influence of many pathways, rather than one, on the coupling is analogous to an entropic component in this quantity.

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

T_{DA} is the electron tunneling matrix element arising from bridge-mediated coupling of the donor and acceptor by all physical pathways. It is sensitive to the energetics of the donor, acceptor, and bridge. FC is the Franck–Condon factor that reflects details of the coupling between the electronic and nuclear degrees of freedom in the protein. For a single physical pathway, the tunneling matrix element can be written[2b,e,4-6,13]

$$t_{DA} = \text{prefactor} \prod_{i=1}^{N} \epsilon_i \quad (2)$$

Neglecting interactions between pathways within the protein bridge, T_{DA} is a sum over t_{DA}'s for all physical pathways. For a pathway, ϵ_i for each block in the path[5b] may be calculated approximately or exactly as discussed above. The prefactor depends on details of the interaction between the donor (acceptor) with the first (last) bond of the tunneling pathway. When experimental systems with similar (or properly scaled) prefactors and FC factors are compared, differences in electron-transfer rates are expected to be due to differences in the coupling via the dominant physical pathways of the systems. The challenge in proteins, then, is to identify chains of orbitals that define dominant pathways. The dominant tunneling pathways correspond to the combinations of bonds in the protein that maximize the products in eq 2.

It is convenient to associate ϵ_i with the decay from one bond to another for typical σ bonds (π systems may require a somewhat different treatment[5]). In this representation, ϵ_i is categorized as occurring either through bond (between two covalent bonds sharing a common atom, or between two hydrogen-bonded groups) or through space (between nonbonded groups).[5] The decay factors associated with these interactions are ϵ_B, ϵ_H, and ϵ_S, respectively. The perturbation theory value of ϵ_B in a one-electron tight-binding limit is

$$\epsilon_i = \frac{\beta_i\gamma_i}{(E - \alpha_L^i)(E - \alpha_R^i) - \beta_i^2} \quad (3a)$$

β_i is the resonance integral between hybrid orbitals in bond i, and γ_i is the resonance integral between bonds connected to the same atom. α_L^i and α_R^i are the site energies of the two orbitals in the ith bond, and E is the tunneling energy determined by the donor and acceptor energetics and vibronic coupling. To emphasize the role played by hole and/or electron tunneling, ϵ_i can be written

$$\epsilon_i = \epsilon_e + \epsilon_h = -\frac{\gamma_{12}^{eff}}{E - E_a^{(2)}} + \frac{\gamma_{12}^{eff}}{E - E_b^{(2)}} \quad (3b)$$

E_a (E_b) is the antibonding (bonding) orbital energy of bond i. If either mediation mechanism dominates for a particular bond,[5]

$$\epsilon_i \simeq \frac{\gamma_i}{E - \alpha_i} \quad (3c)$$

α_i is the energy of the bonding or antibonding orbital that mediates the coupling across the particular bond.

In this paper, we implement a numerical search procedure to identify the combinations of bonds that maximize $\prod_i\epsilon_i$ between specified sites in a protein (based on the crystallographic coordinates). The assumptions described in section III that are used to calculate the pathway decay parameters, ϵ_i, for through bond, hydrogen bond, and noncovalent interactions (and the neglect of interactions between paths) simplify the calculation. The kinds of pathways that are found are not very sensitive to changes in these decay parameters. In section IV we describe the implementation of the search for pathways that maximize $\prod_i\epsilon_i$. Correlations between effective pathway lengths and experimentally derived electronic matrix elements are discussed in section V.

(13) Davydov, A. S. *Phys. Stat. Solidi B* **1978**, *90*, 457.

III. Numerical Implementation of the Model

To probe the nature of the physical pathways that mediate electron transfer in proteins and to estimate relative coupling strengths requires a set of rules to calculate ϵ_i's and a strategy to search for pathways. Estimates of decay parameters from prior theoretical and experimental work are adequate to put us into the correct range to carry out this study. The search procedure begins with a protein crystal structure, identifies covalent and hydrogen bonds, calculates bonded and nonbonded interactions, and searches for combinations of bonds on a path connecting donor and acceptor that maximize $\prod_i\epsilon_i$ according to the prescription below.

(a) Description of the Computer Model. The strategy for pathway searches is as follows. From the crystal structure, covalent and hydrogen-bonded connections between atoms are established. This allows calculation of the ϵ_i's for pairs of bonds. Within a search radius around each atom, noncovalent connections are established if the through space coupling between the atoms is stronger than the through bond coupling. In these searches, all through bond decays (ϵ_B) are taken to be equivalent, although flexibility to vary the coupling for specific bond types was included in the program. Through space, hydrogen bond, and covalent couplings are each calculated differently as described in section IIIb. Limitations of the model are presented in section IV. An improved calculation would include orbital energy and symmetry effects as outlined in ref 5b. Pathways that are found are stored in a circular list. Once the list is full, the poorest pathway in the list is replaced by the next path found, providing that the coupling in the new path is larger than that of the largest coupling already on the list (within a preset factor). The list size can be varied so that there is no danger of overwriting any significant paths.

Once the bond–bond connections and decay factors are calculated, a tree-like search algorithm is performed to find optimum pathways. The program begins at a specified atom (donor or acceptor) and proceeds along a pathway allowed by the connection list until $\prod_i\epsilon_i$ is smaller than a preset cutoff value or the other specified atom (acceptor or donor) is found. If the product of decays to reach a particular atom is smaller than the cutoff, another pathway is sought from the prior atom. If this fails, the pathway backs down the tree by one more bond and seeks another route. Markers are left at the visited atoms, along with the largest $\prod_i\epsilon_i$ value found to that point. Future pathways that reach the same point with a smaller coupling (within a variable factor) are rejected (i.e., the search backs down by one atom and continues).

The process of looking for alternate routes, backing down, branching, and moving forward continues until the acceptor is found. When an acceptable pathway is found, $\prod_i\epsilon_i$ is stored as the new cutoff for acceptable pathways if it is larger than the existing cutoff. Searches then work backward toward the donor to find additional pathways. Only pathways within a set factor of the largest $\prod_i\epsilon_i$ found so far in the search are accepted. As the largest found $\prod_i\epsilon_i$ increases (i.e., as better and better paths are found), poor pathways are rejected progressively earlier in their analysis. The best accumulated pathways are retained.

(b) Choice of Parameters. In this model, which suppresses differences between covalent bond types (differences may be introduced by varying the factor const in eq 4a), there are three key decay factors characterized by six parameters. The decay factors are

$$\epsilon_B = \text{const} \exp[-\beta_0(R - R_{eq}^B)] \quad (4a)$$

$$\epsilon_S = \sigma_S\bar{\epsilon}_B \exp[-\beta_1(R - R_{eq}^B)] \quad (4b)$$

$$\epsilon_H = \sigma_H\bar{\epsilon}_B^2 \exp[\beta_2(R - R_{eq}^H)] \quad (4c)$$

$\bar{\epsilon}_B$ is the value of ϵ_B for an equilibrium length bond (const). The β factors specify the distance dependence of the interactions and the σ factors specify their orientation dependence. In practice, we choose $\sigma_H = 1.0$, $\sigma_S = 0.0$–1.0, $\beta_0 = \beta_1 = \beta_2 = 1.0$–$1.7$ Å^{-1}, and const = 0.4–0.6 for pathway surveys (three parameters). Detailed discussion of these parameters is found in ref 5b. In a standard extended-Hückel calculation, the β's in eq 4 depend on the orbital binding energies. References 3 and 5b show that in the electron-transfer problem it is more appropriate to use the electron tunneling energy to calculate β. In any case, the pathways are not strongly dependent on the particular way in which the energy is chosen.

The reason all β's and σ's are the same is that we intentionally chose to neglect differences between bond types and orientations to obtain qualitative mappings of protein tunneling pathways. The strategy for including bond differences and angular effects is described in ref 5b. Although angular effects require greater attention, the differences between decay factors for different bond types will not cause gross changes in the pathways. The strategy presented here would be meaningless if

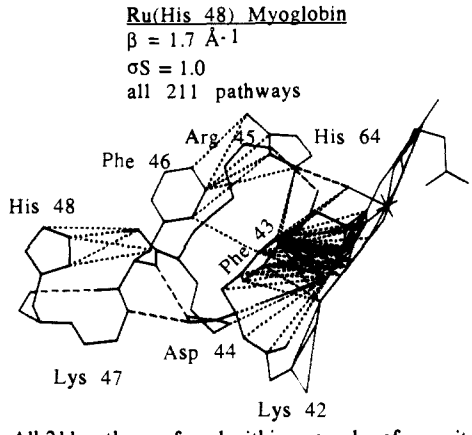

Figure 1. All 211 pathways found within one order of magnitude of the best path for Ru (His 48) myoglobin. $\beta = 1.7$ Å^{-1}, $\sigma_S = 1.0$. In all figures solid lines represent covalent bonds, dashed lines represent hydrogen bonds, and dotted lines represent through space contacts along pathways.

the qualitative aspects of the predictions were dependent on fine details of the decay parameters.

The σ values can be purposely varied to find pathways that exclude through space or hydrogen bonded segments. The decay factors include the qualitative aspects of the coupling, such as the similarity between covalent and hydrogen-bonded coupling[5b] as opposed to through space coupling. The rough choice of parameters suffices in a first attempt to obtain a qualitative understanding of dominant pathways.

Realistic values of ϵ_B are defined by the resonance integral for the bond and tunneling electron energy relative to the bond energy. As discussed in refs 3–5, typical values of ϵ_B are 0.4–0.6 for the bonds of interest. The value of β, the decay length of the through space interactions, is determined by the binding energy of the tunneling electron.[5] A 10-eV binding energy corresponds to $\beta \sim 1.7$ Å^{-1}. Changes in the pathways caused by varying β and σ are discussed in section IV. Since we only intend to classify pathways and estimate matrix element ratios, the conclusions are insensitive to the precise values of these three parameters. The prefactors σ_H and σ_S allow the inclusion of orientational factors in an average sense for hydrogen bond and through space contacts.

IV. Results

(a) Sensitivity to Parameters: Example of Ru-His 48 Myoglobin. As an example of an implementation of the program, four known ruthenated myoglobin derivatives[8b] were studied with ϵ_B fixed at 0.6, $\sigma_H = 1.0$, and σ_S varied between 0.01 and 1.0. β was varied from 1.0 (~3 eV binding energy) to 1.7 Å^{-1} (~10 eV binding energy), and minor changes were found in the physical pathways. Decreasing β does introduce more physical pathways. However, the pathways generally differ in minor ways from those found with a larger value.

The number of pathways found with a donor–acceptor coupling within an order of magnitude of the best coupling is ~200 for the ruthenated His 48 derivative of myoglobin (~13 Å transfer distance). Figure 1 shows that most pathways differ in minor ways, such as in the details of the through space connections between the Phe 43 ring and the porphyrin. By eliminating trivially different pathways, a compact network of relatively few bonds is found (corresponding to relatively few physical pathways). Figure 2a displays only one path per group of calculated pathways consisting of identical amino acids in a unique order for ruthenated His 48 myoglobin.

Figure 2a–c shows the sorted His 48 pathways with $\beta = 1.7$ Å^{-1} and σ_S, the through space orientation factor, equal to 1.0, 0.1, and 0.01. For $\sigma_S = 0.01$, all paths with through space legs are eliminated. However, for $\sigma_S = 0.1$, about one-half of the through space contacts present in the $\sigma_S = 1.0$ calculation remain important. These pathways have been sorted as described above. The factor $\exp[-\beta(R - R_{eq})]$ for a 3.5 Å through space contact decreases by a factor of 7 as β changes from 1.0 to 2.0 Å^{-1}. For $\beta = 1.7$ Å^{-1}, an additional 1 Å in through space distance changes the coupling factor for the interaction by a factor of about 5. Through space contacts make significant contributions to pathways

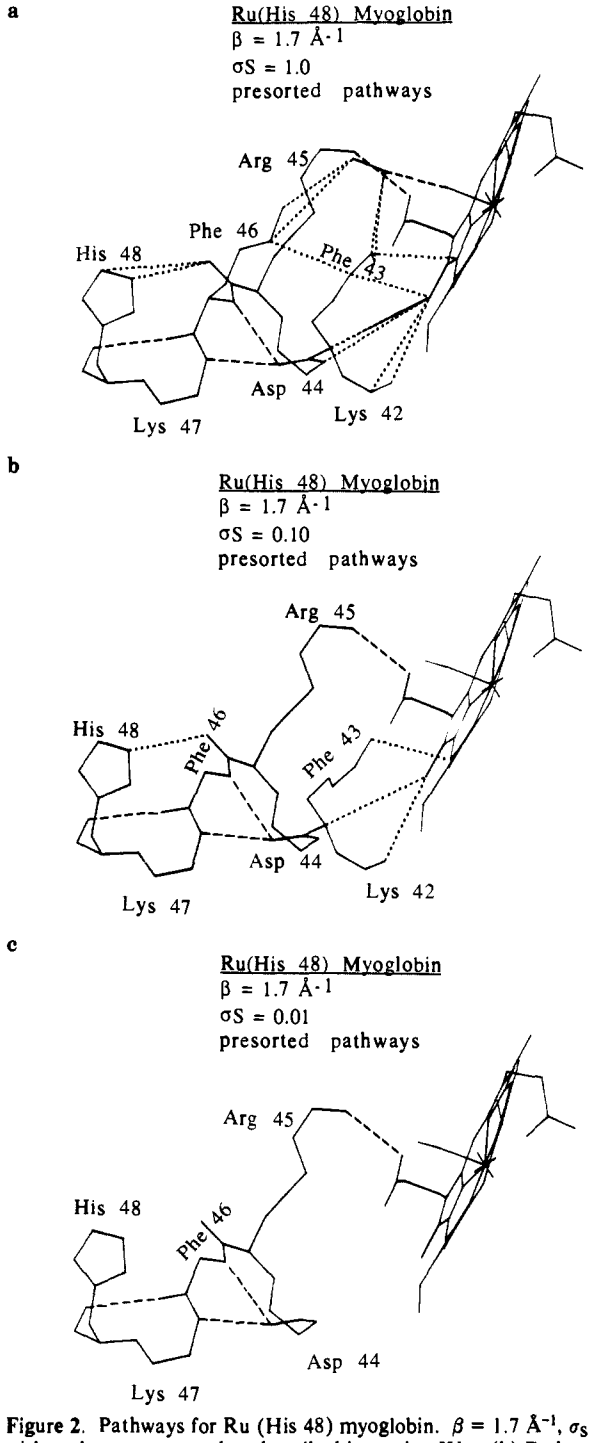

Figure 2. Pathways for Ru (His 48) myoglobin. $\beta = 1.7$ Å^{-1}, $\sigma_S = 1.0$ with pathways presorted as described in section IVa. (b) Pathways for Ru (His 48) myoglobin. $\beta = 1.7$ Å^{-1}, $\sigma_S = 0.1$ with pathways presorted. (c) Pathways for Ru (His 48) myoglobin. $\beta = 1.7$ Å^{-1}, $\sigma_S = 0.01$ with pathways presorted.

when they shortcut a large number of through bond links. The precise value of β, therefore, defines the point at which through space contacts replace longer through bond paths. In isomers with an essential through space contact, the number of useful interconnections between through bond legs will depend on the value of β (e.g., details of interhelix contacts are β dependent, for example, in the Ru–myoglobin His 12 isomer). Figure 3 shows the effect of varying β on the paths for $\sigma_S = 0.1$ (see Figure 2b for comparison).

σ_S is expected to be quite sensitive to the relative orientation of the groups interacting through space[5] and may be much smaller than 1. Naturally, if the distance saved by making a through space connection vs propagating through covalent or hydrogen bonds is offset by this penalty, the through space contributions will be

Table I. Dominant Pathways[a]

isomer	through bond links[b]	through space links	through space distance, Å	effective through bond links	$\Pi_i\epsilon_i$[c]
myoglobin					
His 48	15 (1 H bond)	1	3.26	24.0	4.65×10^{-6}
His 81	22 (4 H bonds)	1	3.19	31.6	9.76×10^{-8}
His 12	21 (no H bonds)	2	3.16, 3.19	37.5	4.73×10^{-9}
His 116	17 (1 H bond)	1	4.17	29.2	3.36×10^{-7}
cytochrome *c*					
His 33	18 (1 H bond)	0		17.6	1.26×10^{-4}
His 39	15 (1 H bond)	0		16.3	2.47×10^{-4}
His 62	22 (3 H bonds)	0		22.6	9.74×10^{-6}

[a] Calculated with $\sigma_S = 0.5$, $\beta = 1.7$ Å^{-1}, and counting hydrogen bonds as two through bond connections from heteroatom to heteroatom. [b] Bonds counted from Ru to the porphyrin ring edge or to the porphyrin metal atom for paths involving a ligand of the porphyrin metal (His 33 cytochrome *c* only); hydrogen bonds counted as two bonds. [c] Relative couplings, squares give relative transfer rates assuming equal activation parameters, donor/bridge, and acceptor/bridge couplings.

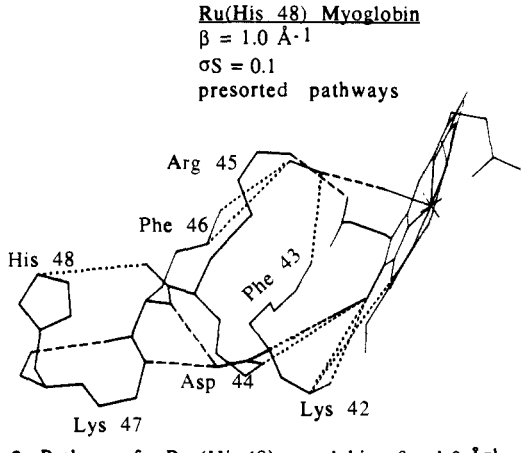

Figure 3. Pathways for Ru (His 48) myoglobin. $\beta = 1.0$ Å^{-1}, $\sigma_S = 0.1$ with pathways presorted.

much less important. For example, in the Ru–myoglobin system, decreasing σ_S to 0.1 in the His 48 derivative shuts off many of the through space pathways within an order of magnitude of the best pathway (Figure 2b).

(b) Analysis of Other Ruthenated Myoglobin Derivatives. Pathways with limited through space or hydrogen bond connections can be found by changing the prefactors, σ_S and σ_H. Using $\sigma_S = 0.5$ and $\beta = 1.7$ Å^{-1}, we have analyzed the three other ruthenated myoglobin isomers. As mentioned in other communications,[5c,8a] His 116 and His 12 derivatives lack fully covalent/hydrogen-bonded paths. The requirement for through space coupling may make the net interaction quite small. This result is of particular interest from an experimental point of view (see section V).

Figure 4 shows the sorted pathways in His 12, His 81, and His 116 ruthenated myoglobins. The His 81 isomer has two parallel sets of pathways that utilize α-helix and contain significant covalent/hydrogen-bonded pathways. The His 12 isomer involves two segments bridged by Trp 14 or Leu 76. No significant purely covalent/hydrogen-bonded pathways exist for this isomer. Pathways in the His 116 isomer are along a single segment of α-helix, but all pathways into the porphyrin involve through space connections.

(c) Analysis of Ruthenated Cytochrome *c*. Ruthenated derivatives of cytochrome *c* labeled at His 33 (horse heart), His 39 (*C. krusei*), and His 62 (*S. cerevisiae*) provide an interesting case study. In these isomers,[10,14] there are roughly one order of magnitude fewer paths than in the myoglobin derivatives because of differences in protein secondary structure. All of the dominant cytochrome *c* pathways involve covalent and hydrogen bonds. The smaller amount of α-helix in cytochrome *c* produces few nearly equivalent hydrogen bond mediated pathways compared to myoglobin. None of the cytochrome *c* derivatives has as large a number of pathways as myoglobin for a comparable donor–acceptor distance. The density of pathways to other surface residues in cytochrome *c* is also low.

The through space (ruthenium to porphyrin ring edge) distances in the His 33 and 39 derivatives are about the same (~13.2 and 13.0 Å, respectively).[10b] The pathways in 39 are predicted to be about a factor of 2 better coupled (the paths differ in effective length by one bond). Typical constant potential barrier models (simple exponential decay models) would assign couplings differing by no more than 10% (see Figure 5) (see refs 10 and 14). In contrast, the His 62 derivative, with a distance 2.5 Å longer (i.e., 15.6 Å), is predicted to have a path with an effective length 4–6 bonds longer than the His 39 and His 33 derivatives.[8a,14] This difference should result in a coupling difference of about a factor of 10. This is also a larger coupling difference than typical exponential decay models would predict.

Table I summarizes the pathway lengths, their characteristics, and the relative coupling strengths for the best paths in the ruthenated derivatives of myoglobin and cytochrome *c*.

V. Discussion

This paper presents a simplified numerical implementation of the electron tunneling pathway model.[5] We predict qualitative differences in the nature of tunneling pathways in different proteins as well as in different ruthenated isomers of the same protein. The calculated size of the decay factor product (eqs 2 and 4) predicts an effective tunneling pathway length for each isomer. This effective length is directly related to the relative electron transfer rates in the isomers, provided that changes in FC factors are accounted for. In some instances, experimental reaction rate vs free energy studies have led to determinations of the coupling matrix elements and direct comparisons with theory.[10,14]

The ordering of the predicted coupling is the following: His 39 cytochrome *c* ~ His 33 cytochrome *c* > His 62 cytochrome *c* ~ His 48 myoglobin > His 81 myoglobin ~ His 116 myoglobin ~ His 12 myoglobin. It is encouraging that this ordering is in accord with the available experimental information.[8b,10,14] Notably, the pathway calculations predict the small observed differences between the His 33 and His 39 couplings as well as their relatively large size compared to His 62 cytochrome *c*, which has a transfer distance ~4 Å longer but an effective pathway length ~5 bonds longer.[14] Both the experimental values and theoretical estimates of the couplings for cytochrome *c* are more reliable than those for myoglobin.

The ruthenated His 116 and His 12 myoglobin isomers lack fully covalent/hydrogen-bonded pathways.[5c,8b] The through space linkages in the His 12 and 116 isomers may lead to smaller net couplings than in the His 81 isomer, despite the quite similar transfer distances in the three derivatives. Recent work suggests that the His 116 myoglobin intramolecular $ZnP^* \rightarrow Ru^{3+}$ electron-transfer rate may be smaller than previously estimated.[15]

(14) (a) Therien, M. J.; Bowler, B. E.; Selman, M. A.; Gray, H. B. In *Electron Transfer in Inorganic, Organic, and Biological Systems*; Bolton, J., McLendon, G. L., Mataga, N., Eds.; ACS Symposium Series; American Chemical Society: Washington, DC, in press. (b) Beratan, D. N.; Onuchic, J. N., *Ibid.*

(15) Upmacis, R. K. Unpublished results.

(16) Kuki, A.; Wolynes, P. G. *Science* **1987**, *236*, 1647.

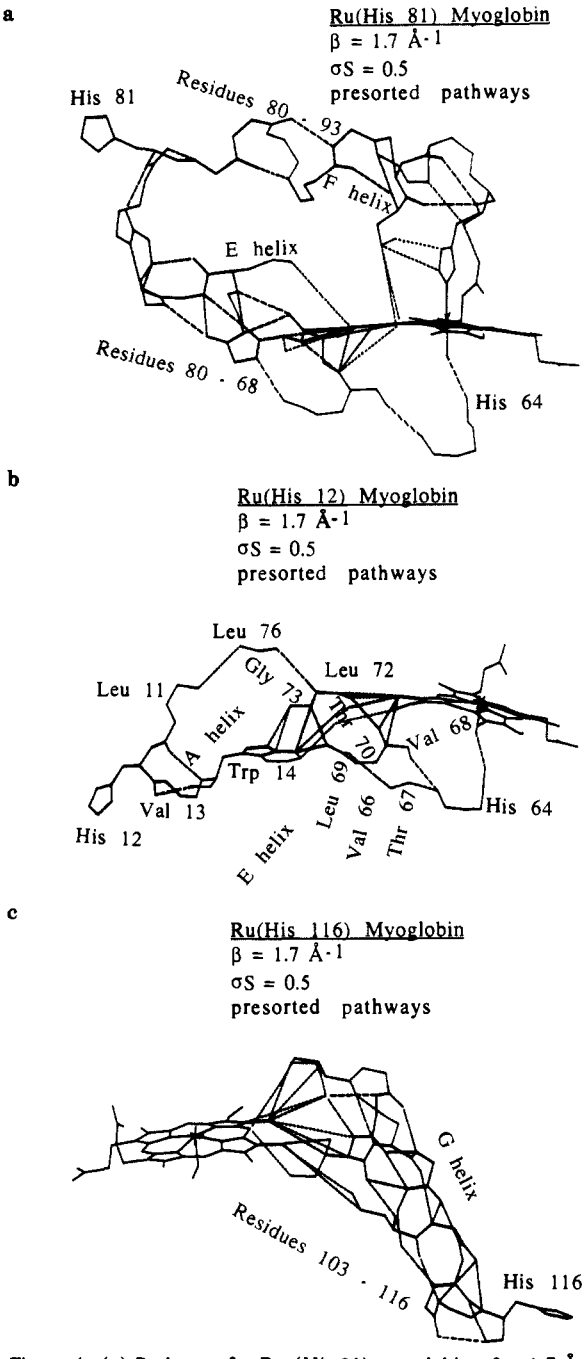

Figure 4. (a) Pathways for Ru (His 81) myoglobin. $\beta = 1.7$ Å^{-1}, $\sigma_S = 1.0$ with pathways presorted. (b) Pathways for Ru (His 12) myoglobin. $\beta = 1.7$ Å^{-1}, $\sigma_S = 0.5$ with pathways presorted. (c) Pathways for Ru (His 116) myoglobin. $\beta = 1.7$ Å^{-1}, $\sigma_S = 0.5$ with pathways presorted.

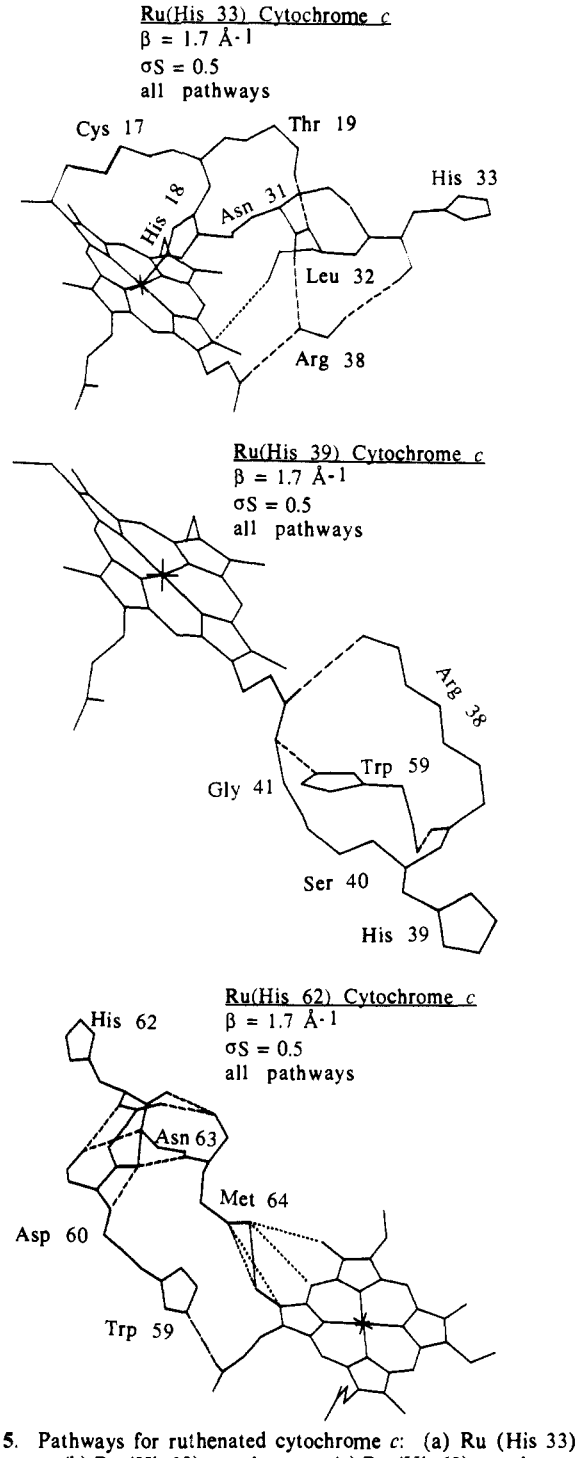

Figure 5. Pathways for ruthenated cytochrome *c*: (a) Ru (His 33) cytochrome *c*; (b) Ru (His 39) cytochrome *c*; (c) Ru (His 62) cytochrome *c*. $\beta = 1.7$ Å^{-1}, $\sigma_S = 0.5$. Note the limited number of paths and the rather direct ones into the π-cloud of the porphyrin for the His 39 isomer. The coordinates were generated with use of the tuna X-ray structure, substituting the appropriate amino acids, and performing an energy minimization.[10b] The low density of pathways eliminated the need for sorting the output paths.

Additional experiments are needed on both the His 12 and His 116 isomers to resolve the intramolecular contributions to the observed $ZnP^* \rightarrow Ru^{3+}$ rates. It is also possible that distance fluctuations in through space contacts could be "frozen out" at low temperatures. The intramolecular rates in His 12 and His 116 isomers should be examined carefully to see if they exhibit unusual temperature dependences related to temperature-dependent pathways.

The systematic pathway analysis presented here serves several purposes: (1) to provide design guidelines for engineered electron-transfer proteins, (2) to suggest significant physical pathways for further detailed theoretical analysis, (3) to provide interpretations of experimental electron transfer experiments in native and mutant proteins,[8,10,14] (4) to provide a basis for interpretation of theoretical models for electron transfer in proteins that explicitly include contributions from all paths,[9,16] and (5) to test the dependence of the pathways on specific aspects of the electronic structure of proteins. The method has succeeded in providing guidance in the ruthenated protein research and in providing some understanding of the origin of electronic coupling differences in several myoglobin and cytochrome *c* isomers. The pathway calculations are now being used to design proteins with or without hydrogen bonds along putative pathways.[10]

We reiterate our prediction that relatively few pathways contribute to the donor–acceptor coupling and recall that this observation is consistent with our understanding that the coupling decays rather rapidly with distance and is very sensitive to details of the interactions along the pathway. This concluison is ex-

perimentally testable with modern synthetic peptide techniques.[17] If many pathways are important, the rate should be weakly dependent on the details of the intervening medium.

Finally, we emphasize that this theoretical work is of a survey nature aimed at providing design and interpretation guidelines. Interpretation of transfer rates requires corrections due to differences in reaction free energies, reorganization energies, and tunneling matrix element prefactors (donor-first bridge orbital and last bridge orbital-acceptor coupling factors) prior to attempting a correlation with pathway length. From a theoretical point of view, interference between multiple pathways, high-order (multiple scattering) corrections, bond orientation, and energetic effects have been neglected. The importance of these effects in causing qualitative changes in the pathways and quantitative changes in the couplings is being investigated. In spite of the limitations of the calculations, this work describes the physical mechanism for electron tunneling through the protein environment and it provides a tool to guide the design and interpretation of new protein-electron-transfer experiments. Preliminary results from this study, especially on cytochrome *c*, are encouraging since effective tunneling pathway lengths correlate very well with experimental estimates of the tunneling matrix elements.[14]

VI. Software

Software was developed with the Digital Equipment implementation of FORTRAN 77 to run on Digital Vax and Microvax machines. The source code and users manual are available from the authors.[18] The program inputs parameters from four data files and one BIOGRAF[19] structure file derived from the X-ray crystal structure. The software outputs an ascii file listing the pathways and a BIOGRAF vector file with their coordinates.

Acknowledgment. This work was performed in part at the Jet Propulsion Laboratory, California Institute of Technology and was sponsored in part by the Department of Energy's Energy Conversion and Utilization Technologies Division—ECUT, through an agreement with the National Aeronautics and Space Administration. Particular credit is due Minoo Dastoor, whose support and encouragement made this work possible. J.B. thanks Caltech for a Summer Undergraduate Research Fellowship (SURF). B.B. acknowledges the Medical Research Council (Canada) for a postdoctoral fellowship. J.N.O. and D.N.B. thank the National Science Foundation and CNPq (Brazil) for a Binational Research Grant that allowed international visits during which much of this work was performed and the Brazilian agencies FINEP and CNPq for additional support. J.N.O. also thanks the University of California, San Diego for start-up funding support. Work on protein electron transfer at Caltech (H.B.G.) is supported by grants from the National Science Foundation and the National Institutes of Health.

(17) Raphael, A. L.; Gray, H. B. *Proteins* **1989**, *6*, 338.

(18) A users guide to the pathway program, written by B. E. Bowler, J. Betts, and D. N. Beratan, 1989. The software has recently been modified to run on Silicon graphics workstations.

(19) BIOGRAF is a product of Biodesign, Inc., Pasadena, CA.

Analysis of Intramolecular Motions by Filtering Molecular Dynamics Trajectories

Pnina Dauber-Osguthorpe and David J. Osguthorpe*

Contribution from the Molecular Graphics Unit, University of Bath, Claverton Down, Bath BA2 7AY, UK. *Received December 20, 1989*

Abstract: A novel method for analyzing molecular dynamics trajectories has been developed, which uses digital signal processing techniques to eliminate unwanted motion and retains only motions of interest. In particular, it is possible to filter out the high-frequency motions and focus on the low-frequency collective motions of molecules. The trajectories of each of the atoms in the system (or any subset of interest) are Fourier transformed to the frequency domain, a filtering function is applied, and then an inverse transformation back to the time domain yields the filtered trajectory. The validity and merits of the filtering method were studied in detail for acetamide and *N*-acetylalanine-*N'*-methylamide as models for peptides and proteins. Initially the technique was tested for fluctuations around one local minimum. The normal modes obtained by diagonalizing the mass-weighted second-derivative matrix were combined to generate a well-characterized "normal-mode trajectory". The frequency distribution of this trajectory approximates a δ function with a peak for each of the frequencies in the normal-mode analysis. By use of a filtering function that retains only one peak of the spectral distribution, a single mode was extracted. This filtered mode had all the characteristics of the original normal mode. Technical aspects such as the effects of simulation length and sampling frequency were also examined by using normal-mode trajectories. The method was then applied to "real" molecular dynamics trajectories. We have shown that information about the structural mobility at the vicinity of a minimum, traditionally obtained from normal-mode analysis, can also be extracted from molecular dynamics simulations by using the filtering technique. In addition, molecular dynamics simulations that include conformational transitions, such as $\alpha_R \rightarrow C_7^{eq}$, were used for evaluating the merits of the filtering technique in anharmonic regions. We concluded that with this technique it is possible to characterize the important motions of a molecule in a way analogous to normal-mode analysis, without confining the study to harmonic oscillations and one local minimum energy conformation. The filtering technique is very flexible and can easily be applied to sections of a molecule or whole molecular systems, and various types of motion can be selected by designing the appropriate filtering function. Since it is also not very demanding computationally, it can serve as a powerful tool for the characterization of the dynamic behavior of small and large molecular systems.

Introduction

Molecular mechanics calculations have been used since the late 1960s[1,2] as an important tool for understanding the structural and energetic preferences of molecules.[3,4] In recent years the theoretical (and experimental) study of the dynamic behavior of molecules expanded as well. In particular, the flexibility and internal motion of biomolecules such as proteins (enzymes, receptors, etc.) or nucleic acids (DNA and RNA) are of importance

(1) Lifson, S.; Warshel, A. *J. Chem. Phys.* **1968**, *49*, 5116.
(2) Allinger, N. L.; Sprague, J. T. *J. Am. Chem. Soc.* **1972**, *94*, 5734.
(3) Ermer, O. *Struct. Bonding (Berlin)* **1976**, *27*, 161.
(4) Burkert, U.; Allinger, N. L. In *Molecular Mechanics*; ACS Monograph Series 177; American Chemical Society, Washington, Dc, 1982.

0002-7863/90/1512-7921$02.50/0