Electron Transport in Disordered Polymeric and Biological Systems

Vijay S. Pande[1] and José Nelson Onuchic[2]
[1]*Department of Physics, University of California at Berkeley, Berkeley, California 94720-7300*
[2]*Department of Physics, University of California at San Diego, La Jolla, California 92093-0319*
(Received 3 May 1996)

In current models of electron transport through proteins and other polymers, calculations are performed by finding a single path which dominates. This model may not be applicable for certain proteins, in which disorder caused by differences in couplings between sites or the length of paths lead to a wide distribution of Green's functions. As in other disordered systems, the quantity which is averaged over disorder has a fundamental effect on the physical picture. We examine two different experimental regimes and comment on the role of disorder. [S0031-9007(96)02039-X]

PACS numbers: 87.15.–v

Electron transfer (ET) reactions play a key role in several biological processes such as photosynthesis and throughout cell metabolism [1]. Starting with the "pathways" approach by Beratan and Onuchic [2], several new theoretical advances have tried to predict how details of the protein environment control the coupling between the donor and acceptor sites [3,4]. Going beyond the isotropic description, the pathways approach used the fact that electron tunneling mediated through bonds are longer range than interactions through space. Thus, it is convenient to break this coupling into through space and through bond steps. The pathways strategy estimates the coupling as the product of decay factors along the strongest coupling route (one-dimensional virtual connection) between donor and acceptor. This approach is successful as long as a single pathway or tube of orbitals dominates the coupling.

The reasons for pathways success are also the cause of its main weaknesses. For example, the decay per bond in covalent networks are taken as a constant value, independently of the bond type and of the local environment. However, for certain systems, one would expect that disorder is important and cannot be ignored. Also, pathways assumes that a single pathway tube dominates the coupling without a quantitative justification of why contributions of the remaining of the protein can be neglected. No further understanding of how the protein environment mediates electron tunneling will be possible without addressing these questions. Even though some initial work has started to deal with the issues discussed above, such as the multiple pathways question [3–5], no theoretical approach to deal with both of the questions above, different decay parameters and multiple paths, is presently available. Moreover, in addition to the relevance to proteins, a many path theory can be compared with considerably simpler systems such as dendrimers [5], which are now under study [6].

Generalizing techniques that have been used to understand complex landscapes and disorder in physical systems [7], we develop a framework that may be used to quantitatively address the questions raised above. In Sections I and II, we explore the simplest possible problem where donor and acceptor sites are connected by a linear (single pathway tube) bridge. We include, however, the possibility of disorder for the electronic coupling between bridge orbitals, given rise to the possibility of a breakdown of the constant per bond decay assumption. Depending on the level of disorder, a transition is observed from the "pathways" limit with an exponential decay in time with a rate proportional to the square of the average tunneling matrix element to a nonexponential decay that at short times is controlled by the strong coupling molecules and at long times by the weak coupling ones. These effects are due to the fact that an ensemble of molecules are measured experimentally. Moreover, the nonexponential decay described by Miller *et al.* [8] is perhaps due to these effects. In Section III, a more relevant situation for proteins, with donor and acceptor sites interaction via several pathway tubes, is analyzed. In this case, the transition from a single to many pathways in a single molecule is analyzed.

(I) Disorder in electron transport.—Most biological electron transfer rates (k_{ET}) are described successfully with a nonadiabatic formulation due to a large separation between donor (D) and acceptor (A) sites (and, therefore the tunneling matrix element T_{DA} is small), so

$$k_{ET} = \frac{2\pi}{\hbar} T_{\mathrm{DA}}^2 (FC), \tag{1}$$

where (FC) is the Franck-Condon nuclear factor associated with the nuclear modes activation barrier.

T_{DA} reflects how the protein environment couples the electronic states between the donor and acceptor sites. A detailed understanding of this coupling proved elusive until recent theoretical and experimental advances [1,9]. The early simple models completely neglected the details of the protein medium and assumed that this coupling would decay exponentially with distance.

We will discuss both the single and many path regimes. The analyses of both of these cases begins with the matrix elements of a particular path T_{DA}, which are directly related to the D-A bridge Green's function,

$$T_{\mathrm{DA}} = \sum_{i,j} \nu_{\mathrm{D}i} G_{ij}(E) \nu_{j\mathrm{A}}, \tag{2}$$

 0031-9007/96/78(1)/146(4)$10.00

where ν_D and ν_A are the donor and acceptor states, respectively. For the simplified case of a linear bridge of l orbitals and neglecting backscattering, (see Fig. 1)

$$G(E) \sim \frac{1}{E - \alpha_{B_1}} \prod_{i=1-1}^{l} \frac{v_{i,i+1}}{E - \alpha_{B_{i+1}}}, \qquad (3)$$

where E is the energy of the tunneling electron and α_B is the energy of the orbital [10], and for notational simplicity, we will assume that the constant $E - \alpha_B$ is included in the couplings v. Note that backscattering over a single bond can be trivially included in this model by a redefinition of $v \to v/1 - v^2$; in fact, this approximation works well for $v < 0.38$.

As for the nature of disorder, in this work, we employ a log normal distribution $\mathcal{P}(G)$ for the Green's function of a given path,

$$\mathcal{P}(G)\, dG = \frac{dG}{G\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\ln G/G_m)^2}{2\sigma^2} \right], \qquad (4)$$

or in other words, the probability distribution for $x \equiv \ln G/G_m$ is normal, where G_m is the mean of G; for notational simplicity, we take $G_m = 1$ and thus G and σ are unitless.

There are two types of disorder for which this approximation holds. (1) *Disorder in couplings:* for paths with fixed length l and random couplings v (and therefore random $\ln v$), $\ln G = \sum_i^l \ln v_i$, and (for long enough chains, $l \gg 1$) the probability distribution for $\ln G$ is normal, due to the central limit theorem. (2) *Disorder in lengths:* while the couplings are constant, there is a Gaussian distribution of path lengths l and thus a log normal distribution of Green's functions v^l. We will also consider another form of disorder which involves the overall sign of the paths' Green's functions (which depends on whether l is odd or even) in the many path regime.

The first question in any theoretical analysis of disordered systems is what quantity do we average over disorder. Of course, averaging different quantities yields different physical interpretations. In the following sections, we examine two cases which have clear experimental meanings; in analogy with other disordered systems, we examine the annealed and quenched cases in Sections II and III, respectively.

(II) Many realizations of disorder.—When one performs time-dependent ET experiments, one measures the decay with time of the density of occupation of the donor state. For a distribution of "identical" molecules, a single rate will exist and gives rise to the following decay probability vs time: $P(t) \sim \exp[-2\pi T_{DA}^2 (FC) t/\hbar]$. (Assuming that the back rate is small compared to the forward rate.) Initially, we consider the measurement of $P(t)$ in a sample which contains many molecules, each with different realizations of disorder. We first examine the case when there is only a single path between donor and acceptor, although the Green's functions for this path varies over the molecules in the sample.

(A) Single path between donor and acceptor.—We start with the simplest case: a linear set of orbitals connecting donor and acceptor. In this case, we must average $P(t)$ over disorder, which means calculating

$$\langle P(t) \rangle = \int_0^\infty dG\, \mathcal{P}(G) \exp(-G^2 t). \qquad (5)$$

Examining the form of (5), it is easy to make an analogy to a nondisordered statistical system, in which $\langle P(t) \rangle$ is a partition function, the realizations of disorder are the states, G^2 is the energy, and t is the inverse temperature. Thus, $\mathcal{P}(G)$ is the density of states, and one can view $\exp(-G^2 t)$ as a Boltzmann weight. This analogy also allows some simple qualitative arguments. For example, at high temperatures (small t), we expect entropy to dominate and partition function $\langle P(t) \rangle$ is dominated by the most common realizations of disorder. For small temperature (large t), we expect only the lowest energy states (realizations of disorder with the smallest rate G^2) to contribute. Thus, the measurement of $P(t)$ yields essentially the partition function of this system at a range of temperatures. As one cannot calculate the integral (5) exactly, we approximate it in several regimes.

First, while it is tempting to perform a small t expansion of $P(t)$, calculate the integral for each term (which are just Gaussian integrals), and then re-sum all terms, the resulting series converges only asymptotically, for small t. Another approach is to calculate the integral (5) by mean field and then add fluctuations to Gaussian order. This assumes that $\mathcal{P}(x) \exp(-e^{2x} t)$ looks Gaussian in $x \equiv \ln G$. This is reasonable up to the very large disorder limit, as we will discuss below.

We calculate $\langle P(t) \rangle = \int dx \exp[f(x)]$, where

$$f(x) = -\frac{x^2}{2\sigma^2} - \exp(2x)t \approx -\frac{x^2}{2\sigma^2} - t\left(1 + 2x + \frac{1}{2}4x^2\right). \qquad (6)$$

Performing the Gaussian integral over x yields

$$\langle P(t) \rangle = \exp\left[-t + t\,\frac{2\sigma^2 t}{1 + 4\sigma^2 t} \right] [1 + 4\sigma^2 t]^{-1/2}. \qquad (7)$$

At $t\sigma^2 \ll 1$, we get modifications to exponential decay. For t greater than $1/\sigma^2$, we expect this approximation to break down, as $\exp(-G^2 t)$ begins to deform $\mathcal{P}(x)$ away from Gaussian. To a reasonable approximation, we can assume that this deformation is essentially a truncation of the Gaussian. In terms of our physical analogy, we are examining the partition function at a very low temperature, and we are running out of states. We need only take the low energy (small G^2) states up to some cutoff, as in the long time limit, all states with fast rate (large G) have already contributed. As time enters into the formulas by the weight $\exp[-\exp(2x + \ln t)]$, integration over a truncated Gaussian yields

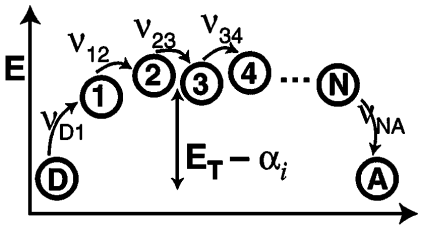

FIG. 1. Schematic diagram of a donor and an acceptor bound to a linear bridge of orbitals with disordered couplings.

$$\langle P(t)\rangle \approx \frac{1}{2}\left[1 - \mathrm{erf}\left(\frac{\ln t}{2\sigma\sqrt{2}}\right)\right]. \tag{8}$$

However, there is yet another possible regime. As any real system is finite, eventually all molecules will have contributed to $P(t)$, except $\mathcal{O}(1)$ molecules at the very tail of the spectrum. Thus, $P(t)$ eventually "freezes," i.e. the decay time becomes sample dependent, analogous to glassy systems below the glass temperature. This freeze occurs always for $t \sim 1/\sigma^2$ (however, the exact position of this freeze depends on the number of molecules). Notice that $1/t$ is equivalent to temperature in a conventional thermodynamic system. Therefore, for times longer than $1/\sigma^2$ nonexponential decay takes place, and, as the time increases, smaller values of G^2 (equivalent to small energies) dominate $P(t)$. Freezing can occur in either the modified exponential or the error function regime, depending on the number of molecules. If most of the decay takes place before freezing, then the process will appear to be controlled by a single rate.

(B) Many paths between donor and acceptor.—For many (p) paths between a single donor and acceptor, we sum over the paths in the Green's function. Upon squaring the Green's function, we get a correlator between Green's function of the paths,

$$\langle P(t)\rangle = \int_0^\infty \prod_{k=1}^{p} dG_k\, \mathcal{P}(G_k) \exp\left[-\sum_{i,j}^{p} G_i G_j t\right]. \tag{9}$$

Moreover, p depends on many aspects, including the geometry of the molecule and, in general, p should increase with the separation of donor and acceptor. Also, one must consider that the Green's functions for these paths may have different signs. Thus, while the square of the total Green's function will, of course, be positive, the correlator in (9) will have varying signs based upon the sign of the product of the Green's functions. We now incorporate this into our model. This means that the $\mathcal{P}(G)$ we have been using (log normal distribution) is really a $\mathcal{P}(|G|)$, i.e., a probability distribution for the magnitude. In the simplest case discussed next, $\mathcal{P}(|G_k|)$ is the same for all paths. We also assume that the sign of the Green's function is random.

When we move to $x \equiv \ln|G|$ space, the sign on G just means that each x_i has a parity attached to it $\chi_i \in \pm 1$. Thus, we get for $pt\sigma^2 \ll 1$

$$\begin{aligned}\langle P(t)\rangle &= \int_{-\infty}^{\infty} \prod_{k=1}^{p} dx_k\, \mathcal{P}(x_k) \exp\left[-t\sum_{i,j}^{p} \chi_i e^{x_i} \chi_j e^{x_j}\right] \\ &\approx \int_{-\infty}^{\infty} \prod_{k=1}^{p} dx_k (2\pi\sigma^2)^{-1/2} \exp\left[-\sum_{i}^{p}\frac{x_i^2}{2\sigma^2} - t\left(\overline{\chi}^2 p^2 + \overline{\chi} p \sum_i \chi_i x_i (2 + x_i) + \sum_{ij}\chi_i\chi_j x_i x_j\right)\right],\end{aligned} \tag{10}$$

where $\overline{\chi} \equiv \sum_i^p \chi_i/p$ and we have used the fact that $\chi_i^2 = 1$. The Gaussian integral yields

$$\langle P(t)\rangle = \exp[-p^2\overline{\chi}^2 t + \tfrac{1}{2}\vec{b}\cdot\hat{A}^{-1}\cdot\vec{b} - \tfrac{1}{2}\ln \times \det(\sigma^2\hat{A})],$$

with the matrix $A_{ij} \equiv \delta_{ij}(\sigma^{-2} + 2\overline{\chi}\chi_i pt) + 2\chi_i\chi_j t$ and the vector $b_i = -2\overline{\chi}\chi_i pt$. Unfortunately, the determinant and inverse of A are not of a simple form. However, they can be brought into a simple form by looking at perturbations from various limits.

For the $|\overline{\chi}| = 1$ *limit* (all χ the same), $A_{ij} \equiv \delta_{ij}(\sigma^{-2} + 2pt) + 2t$ and vector $b_i = -2pt$. As the eigenvalues of a matrix of the form $a_0\delta_{ij} + a_1$ of dimension d are a single eigenvalue $a_0 + da_1$ and a $(d-1)$-degenerate eigenvalue a_0. Thus, the determinant is simply $a_0^{d-1}(a_0 + da_1)$. Also, the sum of the elements of the inverse of this matrix is $d/(a_0 + da_1)$. Thus,

$$\langle P(t)\rangle = \exp\left[-p^2 t + p^2 t\frac{2p\sigma^2 t}{1 + 4p\sigma^2 t}\right] \times [1 + 2p\sigma^2 t]^{-(p-1)/2}[1 + 4p\sigma^2 t]^{-1/2}. \tag{11}$$

In the $p \to 1$ limit, we retrieve the single path result (7). In the region in which we can safely ignore the power law contribution (small t limit), we can rescale (11) to get exactly the single path form by $p \to 1$ and $\sigma/p \to \sigma$. Thus, the multipath case has effectively a p^2 faster rate, but with a width shrunk by a factor of $1/p$.

For the small $\overline{\chi}$ *limit,* we have contributions to $\langle P(t)\rangle$ from the $\exp[-p^2\overline{\chi}^2 t]$ term and

$$\begin{aligned}\det[\sigma^2\hat{A}] &\approx [1 + 2p\sigma^2 t] + 2\overline{\chi}^2 p^2\sigma^2 t[1 + 2(p-1)\sigma^2 t] \\ \vec{b}\cdot\hat{A}^{-1}\cdot\vec{b} &\approx \overline{\chi}^2 p^2 t^2 \times\left[\frac{4p\sigma^2}{1 + 8p\sigma^2 t} - \frac{2p^2\overline{\chi}^2\sigma^4 t}{(1 + 2p\sigma^2 t)^2}\right].\end{aligned}$$

Note that for fixed $\overline{\chi} = 0$, $\vec{b}\cdot\hat{A}^{-1}\cdot\vec{b}$ and the exponential term rate $p^2\overline{\chi}^2$ vanish and the only contributions come from the power law $\langle P(t)\rangle \approx (1 + 2p\sigma^2 t)^{-1/2}$.

Thus, we find many distinct regimes depending on the width of disorder σ and the nature of interference $\overline{\chi}$.

This can be used, for example, to describe the properties of an ensemble of chains. Moreover, these regimes are fundamentally different from the pathways approximation and thus reflect this difference in $P(t)$.

(III) Many copies of a given disordered molecule.— In this section, we examine the behavior of a typical disordered molecule in which there are many paths between donor and acceptor. Moreover, we consider the ideal case in which each of these paths are like those described in section IIA, i.e., all Green's functions have the same sign and are taken from the distribution $\mathcal{P}(G)$ described in Eq. (4). With this model, we now examine the single to multipath transition. Experimentally, this can be realized (1) as biology provides us with the ability to make macroscopic amounts of a particular molecule (proteins for example) that for electron transfer purposes we can neglect differences among them; (2) by single molecule spectroscopy. In this case, we do not average over $P(t)$, as this models the best path of all possible molecules. Instead, we average over the logarithm of the rate (square of the Green's function). We appeal to the arguments of self averaging: since $\ln G/G_m = \sum_{i=1}^{l} \ln v_i$, $\ln G$ is the sum of many random variables and thus the width of $\ln G$ divided by l vanishes in the $l \to \infty$ limit.

This situation is analogous to the averaging of the free energy in spin glass systems [7]. Thus, we can make *another* physical analogy (completely different from that of the previous section), as we can view the different paths are the states, G_i is the Boltzmann weight of the path, and $\langle \ln G_t \rangle = \langle \ln[\sum_i^p \exp(x_i)] \rangle$ is the quenched free energy. If we were to average G instead, the best paths from the best molecules would dominate, analogous to the annealed free energy.

Thus, to average the "free energy" over Gaussian disorder, we make a cumulant expansion,

$$\begin{aligned}\langle 2 \ln G_t \rangle &= 2 \ln p + \frac{1}{p} \sum_i \langle 2x_i + x_i^2 \rangle - \sum_{i,j} \langle x_i x_j \rangle \\ &= 2 \ln p + \sigma^2 (p-1)/p, \end{aligned} \tag{12}$$

where $G_t \equiv \sum_i^p G_i$ is the total Green's function and $G_i = \exp(x_i)$ is the Green's function for path i.

This form is directly analogous to the free energy in the random energy model [11]. Thus, we have the competition between the entropy $\ln p$ versus the benefits of low energy (i.e., high rate) of a few paths $\sigma^2(p-1)/2p$. We see that for large enough σ^2 compared to the "entropy" $\ln p$, the main contribution comes from a few $\mathcal{O}(1)$ fast paths instead of the $\mathcal{O}(p)$ paths with the mean rate.

While this simple model qualitatively demonstrates a single to multiple pathway transition, it is far from quantitative predictions. In this direction, one needs considerably more realistic descriptions for the distribution of disorder $\mathcal{P}(G)$. This consists of many factors, including the distribution of lengths of paths (which is essentially of a purely geometrical nature) as well as the disorder of the values for the couplings v. The consideration of the regime in which different Green's functions may have different signs (for the type of disorder of this section) will also lead to a new level of complexity.

While the pathways model has been successful in describing the electron transport in many proteins, there are some proteins which are not well described. In the previous sections, we have begun to detail how multiple pathway tubes and disorder, in both the length of paths as well as the couplings between sites, can lead to behavior very different from the pathways regime. The importance of these effects for understanding electron transfer in proteins is unquestionable. For example, recently we have computed electron transfer between donor and acceptor bound to different strands in azurin, a β-barrel protein [3]. Because of the multiple hydrogen bonds coupling the two strands, if one considers all the couplings equivalents, several equivalent pathways (via different hydrogen bonds) dominate the tunneling matrix element. Is this the correct picture or does disorder in the couplings (the hydrogen bonds are not exactly the same, for example) choose one or a few of these paths to be the most important? Similar behavior may appear for α-helical motifs. These questions have to be answered if a more quantitative understanding, necessary for new ET protein design, is to be achievable.

The work was supported by the National Science Foundation (Grant No. MCB 96-16186) and the UC/Los Alamos Research (CULAR) Initiative. V. S. P. acknowledges support through MIT's CMSE (NSF Grant No. DMR 90-22933).

[1] I. Bertini, H. B. Gray, S. Lippard, and J. Valentine, *Bioinorganic Chemistry* (University Science Books, Mill Valley, CA, 1994).
[2] D. N. Beratan, J. N. Betts, and J. N. Onuchic, Science **252**, 1285 (1991).
[3] J. J. Regan *et al.*, Chem. Biol. **2**, 489 (1995); J. J. Regan, S. M. Risser, D. N. Beratan, and J. N. Onuchic, J. Phys. Chem. **97**, 13 083 (1993).
[4] A. Kuki and P. G. Wolynes, Science **236**, 1647 (1987); J. M. Grushus and A. Kuki, J. Phys. Chem. **97**, 581 (1993); P. Siddarth and R. A. Marcus, J. Phys. Chem. **97**, 13 078 (1993); A. Broo and S. Larsson, J. Phys. Chem. **95**, 4925 (1991); A. A. Stuchebrukhov, J. Chem. Phys. **104**, 8424 (1996); J. M. Lopez-Castillo and J. P. Jay-Gerin, J. Phys. Chem. **100**, 14289 (1996); D. J. Bicout and M. J. Field, J. Phys. Chem. **100**, 11505 (1996).
[5] S. M. Risser, D. N. Beratan, and J. N. Onuchic, J. Phys. Chem. **97**, 4523 (1993).
[6] G. M. Stewart and M. A. Fox, J. Am. Chem. Soc. **118**, 4354 (1996).
[7] M. Mezard, G. Parisi, and M. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).
[8] J. R. Miller, J. V. Beitz, and R. K. Huddleston, J. Am. Chem. Soc. **106**, 5057 (1984).
[9] W. B. Curry *et al.*, J. Bioenerg. Biomembr. **27**, 285 (1995).
[10] S. S. Skourtis, D. N. Beratan, and J. N. Onuchic, Chem. Phys. **176**, 501 (1993).
[11] B. Derrida, Phys. Rev. Lett. **45**, 79 (1980).