Published in final edited form as:

Isr J Chem. 2016 November ; 56(9-10): 816–824. doi:10.1002/ijch.201600040.

Mean First-Passage Times in Biology

Nicholas F. Polizzi^a, Michael J. Therien^b, and David N. Beratan^{a,b,c}

^aDepartment of Biochemistry, Duke University, Durham, NC 27710 (USA)

^bDepartment of Chemistry, Duke University, Durham, NC 27708 (USA)

^cDepartment of Physics, Duke University, Durham, NC 27708 (USA)

Abstract

Many biochemical processes, such as charge hopping or protein folding, can be described by an average timescale to reach a final state, starting from an initial state. Here, we provide a pedagogical treatment of the mean first-passage time (MFPT) of a physical process, which depends on the number of intervening states between the initial state and the target state. Our aim in this tutorial review is to provide a clear development of the mean first-passage time formalism and to show some of its practical utility. The MFPT treatment can provide a useful link between microscopic rates and the average timescales often probed by experiment.

Keywords

electron transfer; Markov chains; mean first-passage time; protein folding

Introduction

The cell is small enough for a chance event to determine its fate.^[1] For example, gene expression can be governed by a single protein binding event, which in turn may determine cellular phenotype.^[2] A deeper understanding of cellular events that happen *for the first time* may lead to a better understanding of biology at the cellular level. The mean first-passage time (MFPT) defines an average timescale for a stochastic event to first occur.^[3] The MFPT maps a multi-step kinetic process to a coarse-grained timescale for reaching a final state, having started at some initial state. The cellular machinery tailors the many MFPTs relevant to its biochemical networks in order to achieve orchestrated function. Indeed, the Michaelis-Menten rate equation for the reaction rate of enzyme catalysis can be cast as an inverse MFPT.^[4]

In what follows, we provide a pedagogical treatment of the MFPT to traverse a kinetic network, which depends on the number of steps through intervening states between the initial state and the target state. The target state may be a newly synthesized protein, a protein in its native fold,^[5] an exciton on an electron donor,^[6] or an electron on a final acceptor.^[7] We explore the question “How long does it take, on average, to arrive at a final

state F for the first time, having started at an initial state P ”, which is a quest for an effective rate constant $k_{P \rightarrow F}$. We derive the MFPT for a system of coupled kinetic rate equations that are linear and first-order. We then show the direct correspondence between rate equations and Markov chains. Finally, we use Markov chains to derive a general formula^[6a-c] for the MFPT. These formulae can also be directly applied to the analysis of time-resolved spectroscopic data: In a simple case, the timescale of ground-state recovery after photoexcitation can be viewed as a MFPT from the excited-state manifold to the ground-state manifold. Our aim in this tutorial review is to provide a clear development of the MFPT formalism and to show some of its practical utility.

The Mean First-Passage Time

A simple kinetic mechanism that captures key features of the MFPT formalism has three states (Figure 1, states 1, 2, and 3). The unidirectional arrows denote transitions to “trap” states. Any population that visits state 3 of Figure 1, even for the first time, is doomed to remain there. The dashed box of Figure 1 defines a subensemble where population can be redistributed: A random walker can bounce back and forth between states 1 and 2 without being trapped in either state. For concreteness, we have chosen states 1, 2, and 3 as representing states in a donor-bridge-acceptor (D-B-A) system. However, state 1 could also represent a free enzyme population ($[E]$), state 2 an enzyme-substrate complex ($[ES]$), and state 3 a product state ($[E]+[P]$), where $k_1 \rightarrow k_{\text{on}}[S]$, $k_2 \rightarrow k_{\text{off}}$, and $k_3 \rightarrow k_{\text{cat}}$. The derivation that follows for the MFPT from state 1 to state 3 is related to the Michaelis-Menten equation by $v/[E]_0 = 1/\text{MFPT}$, where v is the reaction velocity and $[E]_0$ is the total (conserved) enzyme concentration.^[4] The inverse of the MFPT is an effective rate of the overall reaction.

To find the MFPT ($\langle \tau \rangle_i$) for population to traverse from state i to the trap state (here, state 3), we need a probability density $f_i(t)$ of the first-passage time (FPT).^[3] Then

$$\langle \tau \rangle_i = \int_0^{\infty} t f_i(t) dt \quad (1.1)$$

Because of the irreversibility of a transition to state 3, the probability of *not having transitioned* to state 3 at time t is equal to the probability of being found in states 1 or 2 at time t : $P_1^{(1)}(t) + P_2^{(1)}(t)$, where the superscript denotes the state that is initially populated at $t = 0$. Assuming the population will eventually be found entirely in the trap state 3, the probability of making a first passage to state 3 at time t , having started in state 1, is then $P_{\text{fp}}^{(1)} = 1 - [P_1^{(1)}(t) + P_2^{(1)}(t)]$. The probability density is^[3]

$$f_i(t) = \frac{d}{dt} P_{\text{fp}}^{(i)}(t) = - \sum_{j \neq \text{trap}} \frac{d}{dt} P_j^{(i)}(t) \quad (1.2)$$

(We will present the MFPT formalism for this 3-state mechanism, but will use a notation that can be generalized to larger networks.)

The rate equations describing the kinetic mechanism of the 3-state system of Figure 1 are:

$$\frac{d}{dt}P_1^{(i)}(t) = -k_1P_1^{(i)}(t) + k_2P_2^{(i)}(t) \quad (1.3)$$

$$\frac{d}{dt}P_2^{(i)}(t) = k_1P_1^{(i)}(t) - (k_2 + k_3)P_2^{(i)}(t) \quad (1.4)$$

$$\frac{d}{dt}P_3^{(i)}(t) = k_3P_2^{(i)}(t) \quad (1.5)$$

The FPT probability density for this 3-state system can be found from eqn 1.2, by adding eqns 1.3 and 1.4:

$$\frac{d}{dt}P_1^{(i)}(t) = -k_1P_1^{(i)}(t) + k_2P_2^{(i)}(t) + \frac{\frac{d}{dt}P_2^{(i)}(t) = k_1P_1^{(i)}(t) - (k_2 + k_3)P_2^{(i)}(t)}{\frac{d}{dt}[P_1^{(i)}(t) + P_2^{(i)}(t)] = -k_3P_2^{(i)}(t)} \quad (1.6)$$

and multiplying by -1 , such that

$$f(t) = k_3P_2^{(i)}(t) \quad (1.7)$$

To find the MFPT, one can work with either expression (1.2) or (1.7) for $f(t)$. Substitution of eqn (1.2) into eqn (1.1), and integrating by parts leads to the concept of the MFPT as a sum of residence times (r_i) of each state leading to the final state:

$$\langle \tau \rangle_1 = t \sum_{i \neq 3} P_i^{(1)}(t) \Big|_{t=0}^{t=\infty} + \int_0^{\infty} \sum_{i \neq 3} P_i^{(1)}(t) dt \quad (1.8)$$

$$\langle \tau \rangle_1 = \sum_{i \neq 3} \int_0^{\infty} P_i^{(1)}(t) dt = \sum_{i \neq 3} r_i^{(1)}$$

The MFPT is the sum of residence times of the states within the subensemble defined by the box of Figure 1. The residence times are the shaded areas under the curves in Figure 2.

Noting that $P_i(t)$ of a linear, first-order kinetic mechanism is $\mathbf{p}_i = [\exp(\mathbf{K}t)\mathbf{p}_0]_i$ where $\mathbf{K}$ is the rate matrix, $\mathbf{p}$ is the state population vector, and $\mathbf{p}_0$ is the initial condition vector, the residence times are $r_i = [\int_0^{\infty} \exp(\tilde{\mathbf{K}}t) \tilde{\mathbf{p}}_0 dt]_i = [-\tilde{\mathbf{K}}^{-1} \tilde{\mathbf{p}}_0]_i$, where the tildas denote the subset

of the rate matrix and the initial condition vector that only includes the subensemble of states that are not traps (see eqn. 1.10). The kinetic mechanism of Figure 1, written explicitly in vector-matrix notation is

$$\frac{d}{dt}\mathbf{p}(t)=\mathbf{K}\mathbf{p}(t), \quad \mathbf{p}(t)=\begin{bmatrix} P_1^{(i)}(t) \\ P_2^{(i)}(t) \\ P_3^{(i)}(t) \end{bmatrix}, \quad (1.9)$$

$$\mathbf{K}=\begin{bmatrix} -k_1 & k_2 & 0 \\ k_1 & -k_2-k_3 & 0 \\ 0 & k_3 & 0 \end{bmatrix}$$

The MFPT formalism requires only the subensemble of states that are not traps:

$$\frac{d}{dt}\tilde{\mathbf{p}}(t)=\tilde{\mathbf{K}}\tilde{\mathbf{p}}(t), \quad \tilde{\mathbf{p}}(t)=\begin{bmatrix} P_1^{(i)}(t) \\ P_2^{(i)}(t) \end{bmatrix}, \quad \tilde{\mathbf{K}}=\begin{bmatrix} -k_1 & k_2 \\ k_1 & -k_2-k_3 \end{bmatrix} \quad (1.10)$$

The inverse of the rate matrix defined in eqn 1.10 can be used directly to calculate the MFPT:

$$\langle\tau\rangle_1=\sum_{i\neq 3}[-\tilde{\mathbf{K}}^{-1}\tilde{\mathbf{p}}_0]_i \quad (1.11)$$

with

$$-\tilde{\mathbf{K}}^{-1}=\begin{bmatrix} \frac{k_2+k_3}{k_1k_3} & \frac{k_2}{k_1k_3} \\ \frac{1}{k_3} & \frac{1}{k_3} \end{bmatrix}, \quad \tilde{\mathbf{p}}_0=\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (1.12)$$

such that

$$\langle\tau\rangle_1=\frac{k_1+k_2+k_3}{k_1k_3} \quad (1.13)$$

An effective rate from state 1 to state 3 is then $k_{1\rightarrow 3}=1/\langle\tau\rangle_1$. This derivation of the MFPT to state 3 assumes that state 3 is the only population trap. If other traps were included in the kinetic mechanism, the sum of the residence times defines the *mean exit time*^[3a] from the subensemble defined by the box in Figure 1.

For more than one trap state, a *conditional* MFPT is defined as the MFPT to arrive at the trap state j assuming that the system has not first traversed to any other trap state. For the

conditional MFPT, the second definition of the FPT probability density (eqn. 1.7) is more convenient. The conditional MPFT is^[3a]

$$\langle \tau \rangle_1^{(3)} = \frac{k_3 \int_0^\infty t P_2^{(1)}(t) dt}{k_3 \int_0^\infty P_2^{(1)}(t) dt} = \frac{\int_0^\infty t P_2^{(1)}(t) dt}{\int_0^\infty P_2^{(1)}(t) dt} \quad (1.14)$$

Using $\int_0^\infty t e^{\tilde{\mathbf{K}}t} \tilde{\mathbf{p}}_0 dt = \tilde{\mathbf{K}}^{-2} \tilde{\mathbf{p}}_0$, the conditional MFPT to state 3, again having started at state 1, is

$$\langle \tau \rangle_1^{(3)} = - \frac{[\tilde{\mathbf{K}}^{-2} \tilde{\mathbf{p}}_0]_2}{[\tilde{\mathbf{K}}^{-1} \tilde{\mathbf{p}}_0]_2} \quad (1.15)$$

If more than one state transitions directly to the trap state j , the conditional MFPT to j is

$$\langle \tau \rangle_m^{(j)} = - \frac{\sum_i k_{i \rightarrow j} [\tilde{\mathbf{K}}^{-2} \tilde{\mathbf{p}}_0]_i}{\sum_i k_{i \rightarrow j} [\tilde{\mathbf{K}}^{-1} \tilde{\mathbf{p}}_0]_i} \quad (1.16)$$

where the initial state is m , and $k_{i \rightarrow j}$ is the rate constant from state i to state j . The conditional MFPT is always shorter than the MFPT of the same kinetic mechanism with only one trap state, since a random walker has to avoid the other trap site. For instance, if the initial state is a photoexcited electron donor, the first charge-separation reaction must compete with excited-state deactivation via internal conversion to the ground state, intersystem crossing to the triplet manifold, and emission. The population that eventually lands on state 3 (electron on the acceptor) must have done so faster than any of these competing processes that deactivate the population on the excited donor. Specifically for the mechanism of Figure 1, including the transition to state 4 (gray), the conditional MFPT to arrive at state 3 is

$$\langle \tau \rangle_1^{(3)} = \frac{k_1 + k_2 + k_3 + k_4}{k_1 k_3 + k_2 k_4 + k_3 k_4} \quad (1.17)$$

Eqn. 1.17 reverts to the three-state MFPT (eqn. 1.13) when k_4 is 0.

The quantum yield (QY) of the trap state j can be calculated from the residence times of all the states that transition directly to state j :^[6e]

$$\text{QY}_j = \sum_i k_{i \rightarrow j} r_i \quad (1.18)$$

With an additional rate k_4 that drains population from state 1 of Figure 1 (gray), e.g., to produce a neutral DBA system, the QY of state 3 is

$$QY_3 = k_3 r_2 = \frac{k_3 k_1}{k_3 k_1 + k_4 (k_2 + k_3)} \quad (1.19)$$

If the alternate trap state were not included in the kinetic mechanism ($k_4 = 0$), the QY of state 3 would be 1.

There is a general dependence of the MFPT on the number of states included in the kinetic mechanism. For example, we see that by adding another state (state 4) to a three-state mechanism, the MFPT to state 3 changes. In order to derive the general dependence of the MFPT on the number (N) of states in a linear chain, which is relevant to phenomena such as protein folding, exciton hopping, and charge hopping through proteins or DNA, we find a formal correspondence between kinetic rate equations and Markov chains. We then derive the mean number of steps (S_i) to a trap in a Markov chain that starts on state i . The MFPT is $S_i \cdot t$, where t is the timestep.

From Kinetic Rate Equations to Markov Chains

There is a formal correspondence between Markov chains and linear kinetic networks, which is useful to derive the chain-length dependence of the MFPT. Consider the following equations for the two populations $[X_i \ Y_i]^T$ at timestep i , which are functions only of the populations at timestep $i-1$ and the transition probabilities $P_{i \rightarrow j}$:

$$X_1 = P_{x \rightarrow x} X_0 + P_{y \rightarrow x} Y_0 \quad (1.20)$$

$$Y_1 = P_{x \rightarrow y} X_0 + P_{y \rightarrow y} Y_0$$

In vector-matrix form:

$$\mathbf{u}_1 = \mathbf{M} \mathbf{u}_0, \quad \mathbf{u}_1 = \begin{bmatrix} X_1 \\ Y_1 \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} P_{x \rightarrow x} & P_{y \rightarrow x} \\ P_{x \rightarrow y} & P_{y \rightarrow y} \end{bmatrix}, \quad \mathbf{u}_0 = \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix} \quad (1.21)$$

Moving to continuous time by using the definition of a derivative:

$$\frac{d}{dt} \mathbf{u}(t) = \lim_{\Delta t \rightarrow 0} \frac{\mathbf{u}_{t+\Delta t} - \mathbf{u}_t}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{\mathbf{M} \mathbf{u}_t - \mathbf{u}_t}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{\mathbf{M} - \mathbf{I}}{\Delta t} \mathbf{u}_t = \mathbf{K} \mathbf{u}(t) \quad (1.22)$$

with

$$\mathbf{K} \equiv \lim_{\Delta t \rightarrow 0} \frac{\mathbf{M} - \mathbf{I}}{\Delta t} \quad (1.23)$$

With eqn 1.23, we have a formal definition of the rate matrix $\mathbf{K}$ in terms of the Markov matrix $\mathbf{M}$:

$$\mathbf{K} = \begin{bmatrix} -k_{x \rightarrow y} & k_{y \rightarrow x} \\ k_{x \rightarrow y} & -k_{y \rightarrow x} \end{bmatrix} = \lim_{\Delta t \rightarrow 0} \begin{bmatrix} \frac{P_{x \rightarrow x} - 1}{\Delta t} & \frac{P_{y \rightarrow x}}{\Delta t} \\ \frac{P_{x \rightarrow y}}{\Delta t} & \frac{P_{y \rightarrow y} - 1}{\Delta t} \end{bmatrix} \quad (1.24)$$

Assuming the transition probabilities from each state are normalized:

$$\mathbf{K} = \lim_{\Delta t \rightarrow 0} \begin{bmatrix} \frac{-P_{x \rightarrow y}}{\Delta t} & \frac{P_{y \rightarrow x}}{\Delta t} \\ \frac{P_{x \rightarrow y}}{\Delta t} & -\frac{P_{y \rightarrow x}}{\Delta t} \end{bmatrix} \quad (1.25)$$

$$\text{where } k_{x \rightarrow y} = \lim_{\Delta t \rightarrow 0} \frac{P_{x \rightarrow y}}{\Delta t}.$$

The Mean First-Passage Time of a Biased Random Walker

We now derive the mean number of steps (S_j) that a random walker must take to first reach state 0 in a linear Markov chain (Figure 3), having started on state j . The MFPT $\langle \tau \rangle_j$ is related to the mean number of steps S_j by $\langle \tau \rangle_j = S_j t$, where t is the timestep. For the Markov chain of Figure 3, we start with the following recursion condition:

$$S_j = 1 + pS_{j+1} + qS_{j-1} \quad (1.26)$$

where p is the probability to transition away from state 0 and q is the probability to transition toward state 0. Equation (1.26) indicates that the mean number of steps to traverse to state 0 for the first time, having started on state j , is a function of the mean number of steps to state 0 from all possible states into which state j can transition.

Multiplying the lhs of eqn. (1.26) by 1:

$$(p+q)S_j = 1 + pS_{j+1} + qS_{j-1} \quad (1.27)$$

and rearranging:

$$S_{j+1} - S_j = \frac{q}{p}(S_j - S_{j-1}) - \frac{1}{p} \quad (1.28)$$

we find the recursion formula. The following examples illustrate this recursion:

$$S_1 - S_0 = S_1$$

$$S_2 - S_1 = \frac{q}{p}(S_1 - S_0) - \frac{1}{p} = \frac{q}{p}S_1 - \frac{1}{p} \quad (1.29)$$

$$S_3 - S_2 = \frac{q}{p}(S_2 - S_1) - \frac{1}{p} = \left(\frac{q}{p}\right)^2 S_1 - \left(\frac{q}{p}\right)\frac{1}{p} - \frac{1}{p}$$

In general:

$$S_{j+1} - S_j = \left(\frac{q}{p}\right)^j S_1 + \left(\frac{q}{p}\right)^j \frac{1}{p} - \sum_{k=0}^j \left(\frac{q}{p}\right)^k \frac{1}{p} \quad (1.30)$$

We find, by summation, the difference between the mean number of steps to state 0 from state $j+1$ and to state 0 from state 1:

$$S_{j+1} - S_1 = \sum_{m=1}^j (S_{m+1} - S_m) \quad (1.31)$$

Since $S_1 - S_0 = S_1$,

$$S_{j+1} = \sum_{m=0}^j (S_{m+1} - S_m) \quad (1.32)$$

Directly summing this series,

$$S_{j+1} = \left[q + (p-q)(pS_1 - j) - q\left(\frac{q}{p}\right)^j (qS_1 - pS_1 - 1) \right] (p-q)^{-2} \quad (1.33)$$

Using the boundary condition eqn (1.34) (i.e., the chain is finite and ends at state N), we can solve for S_1 as a function of N , p , and q :

$$S_N = 1 + qS_{N-1} + pS_N \quad (1.34)$$

$$S_1 = \frac{1 - \left(\frac{p}{q}\right)^N}{q - p} \quad (1.35)$$

We then substitute this formula for S_1 into eqn (1.33) to find:

$$S_j = (p - q)^{-2} \left(\frac{p}{q}\right)^N \left[j \left(\frac{q}{p}\right)^N (q - p) + (q - 1) \left(\left(\frac{p}{q}\right)^N - 1 \right) \right] \quad (1.36)$$

Choosing $j = N$, we have a general formula for the mean number of steps to one end of the Markov chain, having started on the opposite end of the chain (S_N):

$$S_N = \frac{1}{q} [aN + b(K^N - 1)]$$

$$K = \frac{p}{q} \quad (1.37)$$

$$a = (1 - K)^{-1}$$

$$b = K(1 - K)^{-2}$$

This equation is found in refs^[6a-c,7d], for $\langle \tau \rangle_N$, after one substitutes $p \rightarrow k_-$ and $q \rightarrow k_+$, and uses the definition $\langle \tau \rangle_N = S_N$. If $q \gg p$, then S_N is proportional to N , as in unidirectional hopping with no backward transitions.

The Mean First-Passage Time for an Unbiased Random Walker

For unbiased diffusion, $p = q = 1/2$, and eqn 1.28 becomes

$$S_{j+1} - S_j = S_j - S_{j-1} - 2 \quad (1.38)$$

leading to

$$S_{j+1} - S_j = S_1 - 2j \quad (1.39)$$

After substituting eqn 1.39 into eqns 1.31 and 1.32, we find

$$S_{j+1} = (S_1 - j)(j+1) \quad (1.40)$$

With the boundary condition,

$$S_N = 1 + \frac{1}{2}(S_{N-1} + S_N) \quad (1.41)$$

we find S_1 , S_j , and S_N :

$$S_1 = 2N \quad (1.42)$$

$$S_j = j(2N - j + 1) \quad (1.43)$$

$$S_N = N(N+1) \quad (1.44)$$

Using the definition $\langle \tau \rangle_N = S_N / k$, and $k = P / \tau = 1/(2 \tau)$, again we find

$$\langle \tau \rangle_N = \frac{N(N+1)}{2k} \quad (1.45)$$

in agreement with refs^[6a-c,7d]. The MFPT for unbiased, diffusive hopping through a chain of length N is proportional to N^2 . We note finally that the Markov chain for this derivation does not include probabilities to remain on state j . If the random walker finds itself on state j , at the next time step it will be located on state $j-1$ or $j+1$ with unit probability (except for the terminal state N). If a “staying” probability were included in the Markov chain of Figure 3, such that $p + q + g = 1$, with g representing the staying probability, the derived formulas for the MFPT will change accordingly.

Insights

The derivations discussed here allow the development of intuition concerning multi-step charge hopping processes in biology. In Figure 2, we show a kinetic mechanism with 3 bridging units between D and A. If the charge injection rate is slow, compared with the back rate, a sizeable population is never found on the bridge states. Under these conditions, the effective rate from D to A is well-approximated as the *residence time* of D.^[8] Note that the residence time of D is proportional to N , the number of hopping steps, for unbiased, diffusive hopping, whereas the MFPT from D to A for the same mechanism is proportional to N^2 . If one assumes very little accumulation of population on the bridge at any given time (e.g., in Figure 2B), then the effective rate from D to A is well-approximated as the inverse

residence time of the population on D (state 1), which will closely resemble the inverse MFPT from D to A in this case (see eqn. 1.8). An appropriate application of this approximation would be if the system studied were in the “activated” hopping regime, such as in DNA hole-hopping experiments, where the kinetics of hopping through the bridging units is rate-limited by charge injection to the bridge.^[7a-c,8] Indeed, a useful criterion for the validity of this assumption is that the backward rate of charge injection (i.e., ET from the bridge to D) should be an order of magnitude faster than forward charge injection into the bridge (or equivalently, using detailed balance, the change in free energy for charge injection to the bridge should be positive and $>2 k_B T$). When charge injection to the bridge is thermodynamically favorable, as in many photoinduced ET experiments, the inverse of the MFPT from D to A, which does not neglect the residence times of the bridge states, must be used as the effective ET rate from D to A. In this situation, the inverse residence time on D ($1/\tau_D$) would greatly overestimate the effective rate from D to A, since the population spends a non-negligible amount of time on the bridging states. Figure 2 shows the qualitative differences between these two regimes. In the activated hopping regime, with unbiased, diffusive hopping between bridge states, the effective rate between D and A decreases with distance approximately as $1/R_{DA}$, where $R_{DA} = aN$, where a is the distance between each site.^[9] If charge injection is thermodynamically favorable, as in many photoinduced charge-hopping experiments, this effective rate decreases with distance as $1/R_{DA}^2$, because the population on the bridge sites (and therefore the bridge-state residence times) is non-negligible.

Applications

Hole hopping in azurin

The blue-copper protein azurin has been a testbed for biological ET theory,^[10] where the protein medium lowers the tunneling barrier of an ET reaction governed by superexchange and to provide tunneling pathways between the donor and acceptor.^[11] Gray's group tagged a powerful Re-based photooxidant to a solvent-exposed histidine of azurin, which had enough driving force in its electronically excited triplet metal-to-ligand charge-transfer ($^3MLCT^*$) state to oxidize a nearby solvent-exposed tryptophan (Figure 4).^[12] The tryptophan radical cation then oxidized the Cu(I) to Cu(II) in the interior of the protein. An azurin protein with the Trp mutated to Tyr or Phe showed no formation of Cu(II). A Trp hole-hopping mechanism (Figure 4, dashed line) can be simplified to the kinetic scheme of Figure 1. As expected from the MFPT formalism presented above, the multi-step hole hopping from Re to Cu via tryptophan greatly accelerated the timescale of hole delivery to the Cu site: A predicted timescale of $\tau_{ET} > 1 \mu s$ for single-step ET from Re to Cu contrasts the measured, Trp-accelerated $\tau_{ET} \sim 50$ ns. Using the relaxed $^3MLCT^*$ state of Figure 4 as state 1 of Figure 1, the charge-separated state with the hole localized on Trp as state 2, and the charge-separated state with the hole localized on the Cu as state 3, with the rates $k_1 = (0.5 \text{ ns})^{-1}$, $k_2 = (1.4 \text{ ns})^{-1}$, and $k_3 = (31 \text{ ns})^{-1}$ (see Figure 4), the MFPT for hole-transfer from the relaxed triplet MLCT state of the Re-photooxidant to the Cu is (see eqn. 1.13) 43 ns.

These azurin experiments point out that charge hopping can be very efficient to move charge quickly at long distance, consistent with the algebraic distance dependence prediction of the effective hopping rate,^[9] even with minimal expenditure of driving force. (Gray's experiments find a driving force of ~ 1 $k_B T$ for hole hopping from Re to Trp, and about 40 $k_B T$ driving force from Trp to Cu(I).) Yet, biochemical catalytic events, such as water oxidation, typically take ms to seconds to complete, necessitating a stable charge-separated state that survives for at least this long. This requirement seems to be the main factor that drives the large expenditure of free energy^[13] for the electron hopping steps in photosynthesis: a reaction with a large change in free energy can be safely approximated as unidirectional. Yet, might proteins manipulate forward and backward ET rates by mechanisms that do not involve such large "wastes" of free energy?

Can Proteins Uniquely Manipulate Forward and Backward Hopping Rates with a Low Expenditure of Free Energy, in Order to Bias Multi-Step Hopping?

The dielectric environment of a solvent or protein can dramatically impact the rate of a charge-separation (CS) and charge-recombination (CR) reaction.^[14] The effective static dielectric constant of a protein is typically assumed to be the same before and after a CS event.^[15] However, a mechanism to speed CS (that occurs in the Marcus normal regime) and slow charge recombination (that occurs in the Marcus inverted regime) would be to switch between a high and low dielectric environment before and after the CS event. To directly determine the effective dielectric environment within a protein interior that mediates photoinduced electron transfer of a donor-bridge-acceptor molecule, we computationally designed a 4-helix bundle to bind the donor-bridge-acceptor molecule PZn-Ph-NDI (Figure 5, A and B).^[16] We measured the impact of dielectric environment on photoinduced CS by studying electron transfer in a wide range of solvents (Figure 5, C), and found that the dielectric environment characteristic of the 4-helix bundle was unique: it could not be described by a single static dielectric constant (ϵ_S). To appropriately describe the CS and CR dynamics that occur in the protein bundle, the effective dielectric of CS was $\epsilon_S \sim 9$, whereas that of CR was $\epsilon_S \sim 3$. The measured time constants for CS and CR (7.5 ps and 70 ps, respectively) are similar to those of the initial electron-transfer event in photosynthesis, begging the question if such a "dielectric switching" mechanism, as described for this simple 4-helix bundle, is at play in the more complex photosynthetic reaction center. In the simple de novo designed bundle shown in Figure 5, the rate of CR is slowed by ~ 4 -fold relative to what would be predicted by the static dielectric environment that describes the forward CS reaction (Figure 5, C, green). Dielectric switching in proteins may be a unique mechanism to bias the direction of charge hopping to a catalytic site, and, furthermore, may be a means to localize charge on the terminal trap site in an otherwise unbiased (or minimally biased) hopping chain.

Conclusions

The MFPT formalism is widely applicable throughout chemistry, physics, and biology.^[17] Using Markov chains, we have provided derivations for the general cases of the MFPT of a biased and unbiased random walker in a 1-dimensional hopping network, which can be applied directly to the analysis of time-resolved spectroscopic data of long-range ET,^[12] as

well as current flow through mesoscopic structures such as bacterial nanowires.^[18] The MFPT to deactivate a highly reactive catalytic intermediate, perhaps through hole-hopping via Trp and Tyr amino acids,^[19] may also place constraints on the timescale of catalysis.^[20] Mathematical modeling of multi-step kinetic mechanisms is indispensable to understand and predict underlying physical behavior. The appropriate application of a MFPT analysis can elucidate the underlying hopping kinetics that give rise to a measured timescale of trapping, such as in exciton hopping to a photosynthetic reaction center. For example, knowing that the MFPT of exciton trapping at a photosynthetic reaction center is ~40 ps, the MFPT formalism constrains the individual exciton hopping times to be < 200 fs.^[21] Appropriate application of the MFPT formalism described here will continue to provide essential insight into biological, chemical, and physical processes by linking microscopic and macroscopic timescales.

Acknowledgments

D. N. B. acknowledges National Institutes of Health grant GM-48043. N. F. P. and M. J. T. acknowledge National Institutes of Health grant GM-071628.

Biographies

Nicholas F. Polizzi received his B.S. in Biology at Cornell University. He is currently a Ph.D. candidate in the Department of Biochemistry at Duke University, working in the labs of David N. Beratan and Michael J. Therien to investigate photo-induced PCET reactions both inside and outside of proteins.

Michael Therien received his undergraduate education at UCLA. His doctoral dissertation research with Bill Trogler at UCSD focused on the reaction mechanisms and spectroscopy of organometallic radicals; his postdoctoral training with Harry Gray at Caltech examined long-range through-protein electron transfer reactions. In 1990, Michael joined the faculty at the University of Pennsylvania; in 2008, he moved to Duke University where he is now the

William R. Kenan, Jr. Professor of Chemistry. Michael's interests in light-driven charge transfer processes, developed during his time with Harry, define a central research theme in his laboratory. Michael's previous honors include Dreyfus (1997) and Sloan (1995) Foundation fellowships, as well as young investigator awards from the Journal of Porphyrins and Phthalocyanines (2002), NSF (1993), Beckman Foundation (1992), and Searle Scholars Program (1991). He has received the ACS Philadelphia Section Award (2004), and the Francqui Medal (Belgium) in the Exact Sciences (2009), and is a Fellow of the American Association for the Advancement of Science (2005) and the Flemish Academy of Arts and Sciences (2009).

David Beratan received his B.S. in Chemistry from Duke University and Ph.D from Caltech. Following postdoctoral and staff appointments at NASA's Jet Propulsion Lab and Caltech's Beckman Institute, where he had intensive collaborative interactions with the Gray group, he moved to the University of Pittsburgh and later to Duke University. At Duke, David is the R. J. Reynolds Professor of Chemistry, Biochemistry, and Physics. David is an elected Fellow of the American Chemical Society, Royal Society of Chemistry, American Association for the Advancement of Science, and American Physical Society. He was awarded a J.S. Guggenheim Foundation Fellowship, the Feynman Prize in Nanotechnology, and the Herty Metal of the American Chemical Society Georgia Section.

References

1. Phillips, R., Kondev, J., Theriot, J., Garcia, HG. Physical Biology of the Cell. Second. Garland Science; New York, NY: 2013.
2. Choi PJ, Cai L, Frieda K, Xie XS. Science. 2008; 322:442–446. [PubMed: 18927393]
3. a) Van Kampen, NG. Stochastic Processes in Physics and Chemistry. Elsevier; Amsterdam, Netherlands: 1992. b) Weiss, GH. Adv. Chem. Phys. John Wiley & Sons, Inc; 2007. p. 1-18.
4. Cleland WW. Biochemistry. 1975; 14:3220–3224. [PubMed: 1148201]
5. a) Zwanzig R. Proc. Natl. Acad. Sci. USA. 1995; 92:9801–9804. [PubMed: 7568221] b) Zwanzig R, Szabo A, Bagchi B. Proc. Natl. Acad. Sci. USA. 1992; 89:20–22. [PubMed: 1729690]
6. a) Bar-Haim A, Klafter J. J. Lumin. 1998; 76–77:197–200. b) Bar-Haim A, Klafter J. J. Phys. Chem. B. 1998; 102:1662–1664. c) Bar-Haim A, Klafter J, Kopelman R. J. Am. Chem. Soc. 1997; 119:6197–6198. d) Montroll EW. J. Math. Phys. 1969; 10:753–765. e) Ritz T, Park S, Schulten K. J. Phys. Chem. B. 2001; 105:8259–8267.
7. a) Berlin YA, Burin AL, Ratner MA. J. Am. Chem. Soc. 2001; 123:260–268. [PubMed: 11456512] b) Berlin YA, Grozema FC, Siebbeles LDA, Ratner MA. J. Phys. Chem. C. 2008; 112:10988–11000. c) Berlin YA, Ratner MA. Radiat. Phys. Chem. 2005; 74:124–131. d) Jortner J, Bixon M, Langenbacher T, Michel-Beyerle ME. Proc. Natl. Acad. Sci. USA. 1998; 95:12759–12765. [PubMed: 9788986]
8. Nitzan A. Isr. J. Chem. 2002; 42:163–166.

9. Venkatramani R, Keinan S, Balaeff A, Beratan DN. *Coord. Chem. Rev.* 2011; 255:635–648. [PubMed: 21528017]
10. a) Beratan DN, Onuchic JN, Winkler JR, Gray HB. *Science*. 1992; 258:1740–1741. [PubMed: 1334572] a) Beratan DN, Skourtis SS. *Curr. Opin. Chem. Biol.* 1998; 2:235–243. [PubMed: 9667934] c) Onuchic JN, Beratan DN, Winkler JR, Gray HB. *Annu. Rev. Biophys. Biomol. Struct.* 1992; 21:349–377. [PubMed: 1326356]
11. Gray HB, Winkler JR. *Q. Rev. Biophys.* 2003; 36:341–372. [PubMed: 15029828]
12. Shih C, Museth AK, Abrahamsson M, Blanco-Rodriguez AM, Di Bilio AJ, Sudhamsu J, Crane BR, Ronayne KL, Towrie M, Vl ek A, Richards JH, Winkler JR, Gray HB. *Science*. 2008; 320:1760–1762. [PubMed: 18583608]
13. Feher G, Allen JP, Okamura MY, Rees DC. *Nature*. 1989; 339:111–116.
14. Hopfield, JJ. *Protein Structure: Molecular and Electronic Reactivity*. Austin, R.Buhks, E.Chance, B.Dutton, P.De Vault, D.Frauenfelder, H., Gol'danskii, V., editors. Springer; New York: 1987. p. 167-185.
15. Schutz CN, Warshel A. *Proteins: Struct. Funct. Bioinform.* 2001; 44:400–417.
16. Polizzi NF, Eibling MJ, Perez-Aguilar JM, Rawson J, Lanci CJ, Fry HC, Beratan DN, Saven JG, Therien MJ. *J. Am. Chem. Soc.* 2016; 138:2130–2133. [PubMed: 26840013]
17. a) Condamin S, Benichou O, Tejedor V, Voituriez R, Klafter J. *Nature*. 2007; 450:77–80. [PubMed: 17972880] b) Szabo A, Schulten K, Schulten Z. *J. Chem. Phys.* 1980; 72:4350–4357.
18. a) Beratan DN, Liu C, Migliore A, Polizzi NF, Skourtis SS, Zhang P, Zhang Y. *Acc. Chem. Res.* 2015; 48:474–481. [PubMed: 25307316] b) Polizzi NF, Skourtis SS, Beratan DN. *Faraday Discuss.* 2012; 155:43–62. [PubMed: 22470966]
19. Gray HB, Winkler JR. *Proc. Natl. Acad. Sci. USA*. 2015; 112:10920–10925. [PubMed: 26195784]
20. Polizzi NF, Migliore A, Therien MJ, Beratan DN. *Proc. Natl. Acad. Sci. USA*. 2015; 112:10821–10822. [PubMed: 26290579]
21. Fleming GR, van Grondelle R. *Physics Today*. 1994; 47:48–55.

Figure 1.

A donor-bridge-acceptor (DBA) kinetic network. The states 1, 2, and 3 define a simple hopping network. Inclusion of a 4th state (gray), which may represent a side-reaction, along with the rate k_4 , influences both the mean first-passage time from state 1 to 3, as well the quantum yield of state 3. The dashed box defines a subensemble of states with finite residence times. States 3 and 4 are “trap” states.

Figure 2.

Simple hopping kinetic networks that display very different effective rates from donor to acceptor. (A) Charge injection to the first bridge state (red) is thermodynamically favorable (here, by ~ 1 k_BT), with a rate $k_{\text{inj}} = 2 \text{ ns}^{-1}$. $k_{\text{back}} = 0.5 \text{ ns}^{-1}$, $k = 1 \text{ ns}^{-1}$, $k_{\text{out}} = 2 \text{ ns}^{-1}$. The residence times of the bridge states (purple, yellow, and green shaded regions) are non-negligible compared to the residence time of the donor state (blue shaded region). (B) Charge injection into the bridge is thermodynamically unfavorable (here, by ~ 2 k_BT), with a rate $k_{\text{inj}} = 0.1 \text{ ns}^{-1}$. $k_{\text{back}} = 1 \text{ ns}^{-1}$, $k = 1 \text{ ns}^{-1}$, $k_{\text{out}} = 10 \text{ ns}^{-1}$. The residence times of the

bridge states (purple, yellow, and green shaded regions) are negligible compared to the residence time of the donor state (blue).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 3.

A Markov chain representing biased, diffusive hopping. Hops to states with a lower number occur with probability q ; hops to states with a higher number occur with probability p . State 0 is the first state in the chain, and state N is the last state in the chain. The total transition probabilities emerging from each state sum to 1, i.e. $p + q = 1$. For unbiased, diffusive hopping, $p = q = 1/2$.

Figure 4.

A kinetic scheme depicting hole hopping between a Re-based photo-oxidant, a nearby tryptophan, and the Cu(I) active site of azurin. The states considered in the MPFT derivation discussed in the text are circled by a green, dashed line. Time constants from states outside the green line were explicitly considered by Gray *et al.*^[12] in their kinetic analysis of hole transfer. The figure was adapted with permission from AAAS.

Figure 5.

Photoinduced electron transfer elicits a change in the static dielectric constant of a de novo designed protein. (A) A donor-bridge-acceptor molecule was bound in the interior of a computational designed protein (B). (C) ET time constants τ_{CS} and τ_{CR} of PZn-Ph-NDI as a function of solvent static dielectric constant (ϵ_s). All solvents include 1% N-methylimidazole (NMI). ϵ_s is a volume-fraction weighted sum of component ϵ_s values. Shaded line widths indicate confidence intervals of the fitted time constants (bounded at 16 % and 84 % percentiles of the bootstrapped lifetime distributions). For the protein holo-monomer, τ_{CR} is denoted with a green X. Green vertical lines mark the range of dielectric constants consistent with the measured value of τ_{CS} in the protein holo-monomer. Purple arrows display an exemplary reduction of ET time constant magnitudes that occur upon NMI coordination, shown explicitly for anisole solvent. Inset: expected dielectric dependence of τ_{CS} (blue) and τ_{CR} (red) according to ET rate theory. DMM = dimethoxymethane, THF = tetrahydrofuran, 2-MeTHF = 2-methylTHF, DCM = dichloromethane, PhCN = benzonitrile, DMF = N,N-dimethylformamide, DMSO = dimethyl-sulfoxide. This figure was adapted with permission from ACS.