Long-Range Hopping Conductivity in Proteins

Siddharth Krishnan¹, Aleksei Aksimentiev¹†, Stuart Lindsay²,³,⁴† and Dmitry Matyushov²,³†

¹Department of Physics and Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL,

²Department of Physics, ³School of Molecular Sciences and ⁴Biodesign Institute, Arizona State University, Tempe, AZ

Abstract

Single molecule measurements show that many proteins, lacking any redox cofactors, nonetheless exhibit electrical conductance on the order of a nanosiemen, implying that electrons can transit an entire protein in less than a nanosecond when subject to a potential difference of less than 1V. In the conventional fast transport scenario where the free energy barrier is zero, the hopping rate is determined by the reorganization energy of approximately 0.8 eV, which sets the time scale of a single hopping event to at least 1μs. Furthermore, the Fermi energies of metal electrodes used in experiments are far-removed from the equilibrium redox states of the aromatic residues of the protein, which should additionally slow down the electron transfer. Here, we combine all-atom molecular dynamics (MD) simulations of non-redox active proteins (consensus tetratricopeptide repeats) with an electron transfer theory to demonstrate a molecular mechanism that can account for the unexpectedly fast electron transfer. According to our MD simulations, the reorganization energy produced by the energy shift on charging (the Stokes shift) is close to the conventional value of 0.8 eV. However, the nonergodic sampling of molecular configurations by the protein results in reorganization energies, extracted directly from the distribution of the electrostatic energy fluctuations, that are only ~ 0.2 eV, which is small enough to enable long-range hopping. Using the MD values of the reorganization energies we calculate a current decay with distance that is in agreement with experiment.

†Corresponding authors: dmatyus@asu.edu, stuart.lindsay@asu.edu, aksiment@illinois.edu
Significance

Electron transfer is fundamental to biology, facilitating a range of metabolic processes and efficient energy conversion. Conventionally, electron transfer through proteins is thought to occur via a chain of metal or organic co-factors connecting one side of the protein to another. Recent experiments, however, show that proteins lacking any co-factors can nonetheless transport electrons with high efficiency if properly connected to metal electrodes. This study provides a theoretical model of such cofactor-less transfer, showing that transient occupation of non-equilibrium states of the protein’s aromatic residues reduces the barrier to electron hopping, facilitating long range and rapid transport. Our results widen the pool of proteins potentially involved in biological electron transport and provide theoretical underpinning to design of protein molecular electronics.

Key Words: Electron transport, molecular dynamics, single molecule conductance, reorganization energy, molecular electronics, break junctions

Introduction

Proteins participating in the energy chains of biology (photosynthesis and respiration (1) and other enzymatic reactions) have to change the oxidation state of their active sites. Since amino acids are mostly redox inactive, the prevailing dogma in the field is that changes of the oxidation state are achieved by utilizing redox active cofactors intercalated into the protein fold. The possibility that this assumption is not always true has been considered in studies of electron relays in proteins (2-6) where protein residues effectively act as semiconductor elements to facilitate exchange of electrons between active sites. This conductivity mechanism adds versatility to redox active enzymes because direct tunneling in biology is limited to $\approx 1.4$ nm (7). If protein residues conduct, can they do so on the time scales of an enzymatic turnover?

It is often suggested that the barrier required for the electron to reach the tunneling configuration to hop between the localization sites is described by Marcus theory (8) with a “universal” protein reorganization energy $(7) \lambda \approx 0.8$ eV and a reaction free energy $\Delta G$. The latter should be close to
zero for hops between equal residues, in which case the barrier is mostly determined by the magnitude of $\lambda$. Some dependence of the tunneling probability on the bridging medium was found (9), but the experimental distinctions between intervening media are not yet sufficient to allow for selection of specific tunneling pathways through particular amino acid residues. These two components, a fairly large and uniform reorganization energy, and lack of knowledge of a specific tunneling path, are obstacles to calculating the conductivity occurring via hops between aromatic residues (mostly tryptophans, tyrosines, and histidines (5, 10)). In particular, the rates of individual hops are not sufficiently high to allow for a hopping conductivity of the magnitude observed experimentally (11-14).

The conductance of single protein molecules can be measured directly, provided that chemical contacts are utilized to inject charge (12), and a summary of one such set of measurements (13) for a series of linear proteins (the consensus tetra-tricopeptide repeat proteins – CTPR) is given in Figure 1. These proteins consist of a two-helix motif that is readily concatenated via recombinant techniques to form oligomers of a controlled length. The measured single molecule resistance is $\sim 1\, \Omega$ over a distance of 15 nm, fairly typical of many non-redox active proteins (12). This equates to a current of 0.1 nA at a bias of 0.1V, or about one electron passing from one contact to the other every ns. In contrast, an estimate (c.f. Equation 5) of the hopping time for a characteristic 0.6 nm distance with $\lambda = 0.8$ eV yields a time between hops of $\sim 1\, \mu s$, three orders of magnitude too slow. Eshel et al. (15) have pointed out that similar problems arise in an analysis of OmcS bacterial wires – specifically that in order to account for the measured conductivity, a reorganization energy below 0.2 eV and a fast diffusion constant ($\sim 20\, \text{nm}^2/\text{ns}$) are required.

A second problem lies in the electrochemical potential dependence of the conductance (16). This is measured either by changing electrode materials (to change the Fermi energy) or by changing the surface polarization of a given metal under electrochemical potential control (16). The conductance measured by either method shows a sharp maximum at $\sim 300$ mV vs. the normal hydrogen electrode, NHE, (Figure 1e). This is quite different from the $\sim 1V$ vs NHE required to oxidize tyrosine and tryptophan (17, 18). As we shall show, these challenges are largely resolved with the appropriate treatment of the reorganization energy.
Below, we describe the required modification to Marcus theory, then use all-atom molecular dynamics (MD) simulations to calculate the effective reorganization energy for charge transfers among the aromatic residues of CTPR monomers and dimers, and calculate the relative displacements of these residues, in order to calculate the hopping rates. A kinetic Monte Carlo simulation of the carrier diffusion, taken together with a model for charge injection, predicts both the magnitude of the currents and their decay with distance in agreement with experiment.

**Figure 1:** Experimental single-molecule conductance as a function of length of linear proteins: a) Five consensus teratricopeptide repeat (CTPR) proteins differing by the number of the repeat units as listed below the images. b) Schematic of a scanning tunneling microscope (STM) measurement of protein conductance. Cysteines (red dots) at the N- and C-termini form contacts between the protein and the metal electrodes. c) Electrical current versus electrode voltage difference. Examples of recordings from 3 different CTPR8 molecules are shown. Many such recordings yield a distribution of single-molecule resistances. d) The most probable resistance of a CTRP protein versus its length. The resistance does not increase linearly with L but rather as $L^2$ (red line) with an intercept of 0.22 GΩ. (data from ref. 12) e) The molecular conductance depends on charge injection potential, as controlled by the Fermi energies of the metals used (black squares) or by changing surface polarization under potential control (open circles) – data (from ref. 15) are for streptavidin (SA) in both cases. Stars are data for CTPR8, the protein considered here (data from ref. 12). For both proteins the conductance has a maximum at about 300 mV vs. the normal hydrogen electrode, NHE, (green arrow), well removed from the oxidation potentials of tyrosine and tryptophan at ~1V vs. NHE (red arrow).
**Modifications to Marcus Theory**

Here, we use the approach of Warshel (19) who amended the Marcus picture of crossing parabolas by specifying the energy gap $X$ between the donor and acceptor energies of the electron as the reaction coordinate. The probability of reaching the tunneling configuration, $X = 0$, is given by Gaussian statistics, as a consequence of long-range electrostatics of the electron interacting with many particles in the medium (*i.e.*, the central limit theorem). The combination of the Gaussian distribution of the reaction coordinate $X$ with the fluctuation-dissipation theorem (FDT) (20, 21)

![Figure 2: The hopping mechanism: a) Schematic free energy diagram of a donor (D)–acceptor (A) pair as a function of the electrostatic potential difference $X$ between the states before (“1”) and after (“2”) a charge transfer. The potential difference between the free energy minima is twice the Stokes shift for optical transitions, $\lambda^{\text{St}}$. Tunneling transitions occur where the parabolas intersect (green box). In Marcus theory, $\lambda^{\text{St}}$ also determines the width of the parabolas (dashed lines) through the fluctuation-dissipation theorem. However, when electron transfer is fast compared to the thermal equilibration time, the width of the parabolas (solid lines) is no longer determined by $\lambda^{\text{St}}$ and must be explicitly calculated from the distribution of energy fluctuations, leading to a reduced reorganization energy, $\lambda'$, with a corresponding reduction in the barrier for electron transfer (compare the crossing point of the solid parabolas vs. the crossing point of the dashed parabolas). b) Hopping occurs via aromatic residues. Two repeat units of CTPR protein are shown with tyrosines and tryptophans highlighted (numbering includes an N-terminal tail). The arrow shows one of the many paths through the two turns of protein, with W35 as the entry point and Y89 as the exit. c) Residues participating in the charge shift reaction change their oxidation state as illustrated (D = electron donor, A = electron acceptor).](image-url)
leads to a specific connection between the separation of distribution maxima in the initial and final states (the Stokes-shift reorganization energy, \( \lambda^{St} (22) \)) and the Gaussian width, 
\[ \sigma_x = \sqrt{2\lambda k_B T}, \]
specified by the Marcus reorganization energy \( \lambda \). The Marcus theory utilizes this connection to establish a single reorganization energy for electron transfer \( \lambda^{St} = \lambda \). This simplification leads to the activation barrier for hopping electron transfer \( \Delta G^+ \) in terms of the two parameters, the reorganization energy \( \lambda \) and the reaction free energy \( \Delta G \):

\[ \Delta G^+ = \frac{(\lambda + \Delta G)^2}{4\lambda} \quad (1) \]

While the Gaussian distribution of the energy gap is a general consequence of statistics of many interacting particles, the application of the FDT to the statistics of the energy gap requires a Gibbsian ensemble based on the assumption that all “essential” configurations are accessible by the protein thermal bath on the reaction time scale. This is often not the case for a solvated protein, as the thermal bath, and many configurations are not accessible, either because they require times that are too long to be reached, or they are thermodynamically unstable (e.g., protein unfolding) or geometrically prohibited by closely packed folded protein. Overall, a folded protein is a frustrated (glassy (23)) medium with many dynamical and geometrical constraints imposed on the configurations coupled to the reaction coordinate. These frustrations lead to a non-Gibbsian statistics of thermal fluctuations and to violations of the FDT (22). When sampling is incomplete (nonergodic), the direct correspondence between \( \lambda^{St} \) and \( \lambda \) is broken. In particular, the parabolas connecting the free energy distributions to \( X \) can be broader (solid lines in Figure 2a) than predicted by Marcus theory (dashed lines in Figure 2a), leading to an effective (reaction) reorganization energy (22, 24)

\[ \lambda^r = \frac{(\lambda^{St})^2}{\lambda} \quad (2) \]
that replaces the Marcus \( \lambda \) in Eq. 1. Simulations of electron transfer between redox cofactors in electron-transfer proteins showed \( \lambda^{st} < \lambda \) and thus there is a lower barrier for hopping electron transfer than suggested in the Marcus framework (22, 24, 25). Low activation barriers achieved through \( \lambda^{T} < \lambda \) can potentially lead to sufficiently fast hops between aromatic residues, thus forming conductivity relays for protein electron transfer (3, 4). However, the required inequality \( \lambda^{st} < \lambda \) has been established only for protein cofactors (22), and it remains unclear if the inequality also applies to hopping conductivity through aromatic residues. According to our MD simulations (described below), a significant reduction of the activation barrier is indeed achieved for electron hops between tyrosine and tryptophan residues due to \( \langle \lambda^{st} / \lambda \rangle \approx 0.4 \). This ratio also enters the nonergodic reaction free energy which is modified from the thermodynamic value \( \Delta G_0 \) to

\[
\Delta G = (\lambda^{st} / \lambda) \Delta G_0
\]

(3)

Taking the ratio of the reorganization energies into account, the nonequilibrium redox potential of the tyrosines and tryptophans approaches the experimentally determine value of \(~300\) mV vs. NHE, Figure 1e.

The next important question is how individual random hops of charge carriers combine to generate the overall current and how the protein conductivity scales with the distance between the points of charge injection connected to external electrodes. The experimental data (13) summarized in Figure 1 (and data from other experiments (26)) show a very slow decrease of conductivity with the size of the protein. Previous studies suggested that this weak scaling reflects coherent tunneling of electrons through the entire protein (27, 28). Existing theories and calculations do not support this hypothesis (29) prompting the need for an alternative conductance mechanism.

Here, we propose that protein conductivity arises from random hopping of charge carriers over the network of tyrosine and tryptophan (Figure 2b) sites of the protein upon injection of the charge carriers at the electrode contact sites. The conductivity of redox-inactive proteins is mostly determined by the chemistry of the contact (12) and the externally imposed bias is dropped largely at the contacts. The random walk does not involve a significant external driving force, so the dependence of conductivity on the protein size is determined by the hopping sites
involved and the hopping rates. We develop a quantitative theory using all-atom simulations to calculate the electrostatic energy fluctuations of donor and acceptor sites in the initial and final states, thus deriving the effective reorganization energies and the free energies for the various hops between multiple pairs of aromatic residues located within the Dutton radius (7) (1.4 nm). The effective reorganization energies are used to calculate the hopping rates, from which a diffusion constant is derived, via a set of Monte Carlo simulations, leading to calculation of the experimentally-observed quantities.

**Extracting Local Hopping Rates from Molecular Dynamics Simulations**

A CTPR8 protein contains a total of 60 aromatic residues, out of which 52 are tyrosine and eight are tryptophan. The protein consists of eight two-helix repeat units with each unit being 34 amino acids in length and containing six tyrosines and one tryptophan (two such units are shown in Figure 2b). We hypothesized that, during the charge transfer, these residues are oxidized to their cationic free radical states (Figure 2c). We probed this possibility by conducting all-atom MD simulations of the protein having the aromatic residues in their neutral and oxidized states and using the fluctuations of the residues’ electrostatic energy to compute the rate of inter-residue electron transfer.

Figures 3a and b illustrate two all-atom systems used to probe a transfer of electron from residue Y77 to Y43. In the initial state of the hopping process (Figure 3a), Y43 is in its oxidized state whereas Y77 is charge neutral. At the end of the hopping process, Y77 is oxidized and Y43 is neutral. To evaluate the rate of electron transfer we computed equilibrium fluctuations of the electrostatic energy gap $X(t) = \sum_i \Delta q_i \Phi_i$, where the sum runs over all atoms of the two residues accepting or donating an electron, $\Delta q_i = q_i^{\text{final}} - q_i^{\text{initial}}$ is the change of the partial charge of atom $i$ upon transfer of the electron, and $\Phi_i$ is the local electrostatic potential at atom $i$. Knowing the partial charges on the oxidized forms of a tyrosine (30) and a tryptophan (31) residue (Tables S4 and S5), we determined the energy gap fluctuations by performing a 0.75 microsecond equilibration of each system and evaluating the instantaneous distribution of the electrostatic potential every 10 ps using the PMEpot plug-in of VMD (32, 33). Figure 3c plots the resulting $X(t)$ data for the initial and the final states of the Y77-Y43 transfer, whereas Figure 3d shows their statistical distributions. The corresponding free-energy distributions of the Y77-
Y43 pair, obtained from the electrostatic energy gap statistics as \( G_i = -k_B T \ln[P_i(X)] \), are shown in Figure 3e.

Figure 3: MD simulations of the electrostatic energy gap produced by a charge transfer process: a) All-atom simulation system representing the initial state of the Y77-to-Y43 charge transfer process, with Y43 being positively charged (oxidized) and Y77 being electrically neutral. b) Final state of the Y77-to-Y43 transfer, with Y43 being neutral and Y77 positively charged. c) Energy gap fluctuations at Y43 (\(X_1\)) and Y77 (\(X_2\)). The black traces indicate 10-ns block average of the 10-ps sampled raw data. d) Electrostatic energy probability distributions for the positively charged Y43 (triangles) and positively charged Y77 (circles). The separation between the distributions’ peaks is \(2\lambda_{St}\). e) Free energy as a function of the donor-acceptor electrostatic energy gap: the intersection at \(X=0\) defines the transition state.

Electron transfer occurs at the point of crossing of these two distributions (\(X=0\) in Figure 3e).

The variance of the distribution yields \(\lambda\) from \(\lambda = \sigma_X^2/2k_B T\). For this particular D-A pair, the full width of the \(X\) distribution is \(\sim 0.5\) eV yielding \(\sigma \sim 0.24\) eV and \(\lambda \sim 1.15\) eV. Thus, the free-energy parabolas are wider than the Stokes shift (\(\lambda_{St} \sim 0.52\) eV Figure 3d) would predict. For this
particular transfer reaction, Eq. 2 yields $\lambda^r = 0.24 \text{ eV}$, substantially less than the value derived either from the Stokes shift or from the distribution variance.

**Figure 4:** Calculation of reorganization energies: **a,b)** Reorganization energies for a tyrosine-tyrosine (panel a) and a tyrosine-tryptophan (panel b) electron transfer calculated as $\lambda = \sigma_X^2 / (2k_B T)$. The yellow and orange curves show cumulative average values of $\lambda$ for the two charge shift reactions. The two values should be equal, in principle, and hence their average, $\lambda^e$, is used in subsequent calculations. The blue curve indicates $\lambda^{St}$. As $\lambda^e$ is substantially greater than $\lambda^{St}$, $\lambda'^r = (\lambda^{St})^2 / \lambda$ is substantially smaller than $\lambda^{St}$. **c)** Distribution of $\lambda'^r$ values calculated for all pairs of aromatic residues located within the Dutton radius of each other. The modal value of the distribution is 0.2 eV, four times smaller than 0.8 eV derived from the Stokes shift. **d)** Distribution of the hopping times calculated from the distributions of $\lambda'^r$, the corresponding edge-to-edge donor-acceptor distances and using the polarizability adjusted reorganization energies. The modal time is $\sim 100 \text{ ps}$, consistent with the experimentally observed nA currents.

Similar calculations (see Methods) were carried out for all pairs of aromatic residues within the Dutton radius (1.4 nm edge to edge distance in the crystal structure) within two repeat units, which, together with the periodicity of the CTRP8 protein, encompassed all likely electron transfer paths. Test simulations carried out for equivalent pairs of residues from different protein units have explicitly verified the convergence of our rate calculations, Figure S1. Thus, taking
the repeat structure of the protein into account, we only needed to calculate the activation
ergies for residue pairs within one protein unit and between the two adjacent units.

Figure 4a and b show cumulative averages of the MD-determined reorganization energy values
for a representative Y-Y and W-Y residue pair. Note here that λ₁ (orange trace) is for the D-A⁺
state and λ₂ (green trace) is for the D⁺-A state (defined in Figure 2c). In either configuration, λ is
the solvation free energy of the electron transfer dipole, which is “-“ on the acceptor and “+” on
the donor. Since the charge distribution is the same, λ is expected to be the same in both states
and indeed λ₁ ≈ λ₂. The average of the two values is used in what follows.

The distribution of λᵣ values, computed according to Eq. 2, is plotted in Figure 4c. The modal
number extracted from our simulations was 0.2 eV. A list of the values of λₛᵗ, λᵥ,a,r, Δ𝐺, λᵣ and
edge-to edge distances for all interacting pairs within a Dutton radius is given in Table S1. As
these simulations employed a nonpolarizable force field, they neglected possible contributions of
induced dipoles that would screen the electrostatic interactions, leading to even further reduction
of the reorganization energy. To account for such electronic screening, we introduced a
correction factor of 0.8, as discussed in detail in Supplementary Information. The correction
factor was applied to both λₛᵗ and λ, yielding an even lower modal value of λᵣ of 0.16 eV,
substantially smaller than 0.8 eV, a value commonly accepted for proteins. The activation
barrier (for Δ𝐺 = 0) is λᵣ/4, so λᵣ = 0.16 eV gives an activation barrier of 40 meV, close to
thermal energy at room temperature. Overall, these data yield an average value for ⟨λₛᵗ/λ⟩ ≈
0.4. Since the oxidation potential of tyrosine and tryptophan in solution includes a reorganization
energy between 0.5 and 2 eV (17, 18), the reduction of reorganization energy in a protein-
aqueous medium to 0.16 eV could account for the resonant injection of charge (Figure 1e and
Eq. 3) because the potential of the (non-equilibrium) oxidized state will be shifted towards the
vacuum by ~ 0.7 eV.

Our calculation of the hopping rates proceeds as follows. The charge transfer rate $k_{ET}$ is given
by

$$ k_{ET} = (1 + g(R))^{-1}k_{NA} \tag{4} $$
where the non-adiabatic transfer rate, $k_{NA}$, is given by

$$k_{NA} = \frac{V_{DA}^2}{\hbar} \sqrt{\frac{\pi}{k_B T \lambda r}} \exp \left[ \frac{1}{2} \gamma^2 \langle (\delta R)^2 \rangle \right] \frac{\Delta G^+}{k_B T}$$  \quad (5)$$

and the activation barrier $\Delta G^+$ is given by Eq. 1 in which $\lambda$ is replaced with $\lambda^r$ and the reaction free energy is from Eq. 3. $\langle (\delta R)^2 \rangle$ in is the variance of the donor-acceptor center of mass distance. The crossover parameter $g(R)$ (see Methods and Supplementary Information, Eqs. S1-S7) accounts for solvent dynamical control (34) of the rate pre-exponential factor (35-37). The factor $g(R)$ includes two dynamic components: the Stokes shift of the energy gap $X(t)$ and the donor-acceptor distance, $R(t)$, modulating the donor-acceptor electronic coupling $V_{DA}$. $V_{DA}$ decays exponentially with the donor-acceptor edge to edge distance and is calculated here from the Hopfield equation (38)

$$V_{DA} = \frac{2.7 \text{ eV}}{\sqrt{N_{D,A}}} \exp \left( -\gamma R \right)$$  \quad (6)$$

where $R$ is the edge-to-edge distance (in nm) and $N_{D,A} = 7$ for a tyrosine and 9 for a tryptophan. $\gamma/2$ is the distance decay parameter, equal to 7.7 nm$^{-1}$ (38). Taken together with the activation barriers obtained from the MD simulations, these equations yield the hopping times for all pairs within the Dutton radius in the protein, and the distribution of these times is plotted in Figure 4d and the forward and backward rates are listed in Table S2. Although the overall electric current is determined by the carrier velocities, and not by the transit times per se, the hopping times of $\sim 0.1$ ns (Figure 4d) are consistent with the nA currents measured in these molecules and are much smaller than 1μs estimated for $\lambda = 0.8$ eV. The dynamical crossover parameter $g(R) \propto V_{DA}^2$ gains its distance dependence from the electronic coupling. Therefore, for $g(R) > 1$, the term $V_{DA}^2$ cancels in Eq. 4 and the rate constant becomes insensitive to the electronic coupling. Many electron hops fall in this regime of dynamical control (Table S2): for instance, charge transfer between Y77 and Y43 shown in Figure 3 produces $g \approx 2\lambda$. The dynamical control introduces
the dependence of the electron-transfer kinetics on protein dynamics and elasticity which are absent in traditional theories.

Figure 5: Charge injection model of the current: a) An electron is extracted at the right contact 2 (energy $eF - eV_b$) at a rate $k_R^2$ via the oxidation of the closest aromatic residue. The oxidized state then diffuses, through a sequence of hops, within the body of the protein (dashed box, current density $j_d$) to be collected by the contact 1 on the left via reduction of the closest residue at a rate $k_R^1$. $eF$ is the Fermi energy of the electrodes, $V_b$ the applied bias, $\eta$ the potential drop at the contacts and $\phi_{ref}$ the potential of the electrodes with respect to a reference electrode. The arrows show the direction of electron motion. $K(L)$ is the experimentally-determined fraction of the bias that is dropped across the bulk of the protein (Figure S2). b) Calculated current-voltage dependences for $L=10$ nm (black) and $L=15$ nm (red) where the current is expressed in units of the saturation current $I_0$. The curves are linear in the ±200mV range (blue box), as observed in the experiments (Figure 1c). c) Calculated resistance vs length (red line) fits experimental data (dots) with $I_0 = 5.5$ nA, and the parameter $a = 3.8$ (see text for detail). The combination of the distance dependence of hopping diffusional current with the length-dependent potential drop across the contacts.

Calculation of currents and current decay rate from the hopping rates

The decay of current with distance is described by a parameter that depends on the electronic coupling between the electrodes and the entry/exit residues, the size of those residues, and the diffusion constant, as we will show below. The electronic coupling can be estimated from the measured contact resistance, and the size of the entry/exit residues is known, leaving the diffusion constant to be determined.

A model of the electrode-protein system is shown in Figure 5a, from which we derive an expression for the current through the complex, below. Electrons are extracted from the residue closest to the right (positive) electrode (2 on Figure 5a) and injected in a similar way from the residue closest to the left (negative) electrode (1 in Figure 5a). The experimentally determined contact resistance (Figure 1d) is a significant fraction of the overall resistance, particularly for
the shorter molecules. The fraction, $K(L)$, of the applied bias dropped across the bulk of the molecule was determined using an elementary circuit analysis, (13) and a fit to $K(L)$ is shown in Figure S2. However, even in the longest molecules, where a significant fraction of the bias is dropped across the protein, the applied field is small compared to the internal fields within the protein (see, for example, Figure 4 of Martin et al. (39)). Thus, once injected, the charges diffuse under the driving force of the carrier gradient owing to electron injection at contact 1, until captured by contact 2, a process characterized by the carrier diffusion constant.

**Diffusion constant.**

The diffusion constant connects the diffusion flux to the gradient of the volume density of charge carriers in the macroscopic Fick’s law. A one-dimensional Brownian walker stepping the distance $\Delta x$ in each step requiring time $\tau$ has the diffusion constant $D = (\Delta x)^2/(2\tau)$. The challenge of extending this equation to carrier diffusion in the protein is that both the step $\Delta x$ and the hop time $\tau$ become distributed variables. The problem is solved analytically by Derrida’s model (40) which sets up diffusion on a one-dimensional periodically replicated chain of sites with fixed $\Delta x$ and forward and backward transition rates specified at each lattice site. This solution does not allow distributed values of $\Delta x$ and provides no algorithm for choosing a path maximizing diffusion. The latter problem is resolved by kinetic Monte Carlo (KMC), which generates multiple paths through the network of Tyr and Trp residues and yields the time of first passage between the injection and collection sites. However, it does not provide a direct formalism for calculating the diffusion coefficient. The ability of the carriers to tranverse the 3D protein over many alternative paths is an essential aspect of the problem that has to be captured by a successful formalism. In the absence of such a formulation, we adopted a somewhat heuristic extension of the result for a Brownian walker (to which Derrida’s model reduces for symmetric forward and backward rates), calculating the diffusion constant in the form

$$D = \left< \frac{(\Delta x)^2}{2\tau} \right>$$

where both the step size $\Delta x$ and the step time $\tau$ are stochastic variables changing along the paths produced by KMC. Specifically, $\tau$ is calculated as the waiting time on a given site. In the limit of a single path, this result reduces to Derida’s calculation, where the average is now for the sites along the path. In the KMC formulation, the average was taken along a given path, with the most probable value of diffusion constant given by the peak of the distribution calculated for all paths.
The calculation is set up using the hopping rates and center-of-mass distances (Table S2) for all aromatic pairs within a Dutton radius in two repeats of the CTPR unit. We use two repeats so as to include the hops between neighboring repeats (c.f., Figure 2b). As shown in Figure S1, the properties of longer proteins are captured via translational symmetry.

Using the data listed in Table S2, a graph was constructed with edges that represent the hopping rates, with two edges between each pair of residues, one representing the forward hopping rate and the second the backward hopping rate. Electron hopping MC simulations were run until the electron reached the exit residue, with the value of $D_i$ calculated at the $i$th hop from the residence time, $\tau_i$, and the center of mass hopping distance $\Delta x_i$. On completion of the passage, an average of these values is stored (Eqns. S17-21).

In the experiments, N- and C-terminal cysteines were incorporated into the proteins in order to form defined chemical contacts to the electrodes. Examination of the structure (Figure 2b) shows that C-terminal injection or extraction is only possible at Y89. Injection at the N-terminal can occur at W35 (the closest residue in sequence to the N terminus) or Y36 which is at almost the same through-space edge to edge distance (0.9 nm) from the cysteine (W35 is 1 nm through-space edge to edge distance from the cysteine). We have therefore calculated the distribution of $D$ values for paths that start or end on either of these residues (the inclusion of backward and forward hops captures diffusion in either direction). 100000 simulations were run for each set of paths and the distributions of $D$ are shown in Figures 6a (W35 path) and 6b (Y36 path). The corresponding most probable $D$ values are 22.1 nm$^2$/ns and 22.8 nm$^2$/ns.

**Figure 6:** Distribution of diffusion constant values for paths that connect W35 and Y89 (a) and Y36 and Y89 (b) calculated via a kinetic Monte Carlo algorithm. Fitting with Gamma function distributions.
(dashed lines) yields most probable values of \( D = 22.1 \, \text{nm}^2/\text{ns} \) for the W35 paths and 22.8 nm²/µs for the Y36 paths.

As a check on this heuristic approach, we used the Derrida formalism (40) to calculate values for two limiting cases. In one case we chose a single path that maximized forward rates (which is not the same as maximizing the diffusion constant) and calculated the corresponding diffusion constant (this path illustrated in Figure 2b). The calculation was carried out following Eqns. S13 and S14 using the rates shown in Table S3 (the final hop that closes the loop as required for the Derrida theory is Y89 to W35, equivalent to Y89 to W103 owing to translational symmetry – Figure S1). For this path we obtain \( D = 3.06 \, \text{nm}^2/\text{ns} \), an order of magnitude less than the multiple-path value obtained from the KMC calculations. An alternative approach, using the Derrida theory, is to find all the paths that connect entry and exit, calculate the corresponding \( D \) values for each, and summing them, on the assumption that each path adds a current proportional to its corresponding \( D \). The obvious fault with this approach is that paths are allowed to overlap – i.e., this approach allows a given site to be multiply oxidized (which is energetically impossible). This will yield an overestimate of \( D \). We find 433 such paths connecting Y36 to Y89 for which the total \( D = 578 \, \text{nm}^2/\text{ns} \), an order of magnitude larger than the value obtained from the MC calculations. These two limiting cases bound the results of the MC simulations, indicating that our heuristic calculation of \( D \) from KMC is likely of the correct order of magnitude.

**Charge injection.** We assume charge injection into a single site for which the injection and recombination rates depend only on the potential drop at the contacts, \( \eta \) (Figure 5a) and the electronic coupling between the electrodes and the entry and exit sites, assumed, for simplicity, to be the same, \( \Delta \, \text{eV} \), at each contact. Using empirical data to calculate \( \eta \) (Figure S2) and expressions for the injection rates (Eqns. S15) solved for stationary conditions (Eqns. S14-16) results in the following expression for the contact conductance:

\[
G_0 = \left. \frac{dI}{d\eta} \right|_{\eta=0} = \frac{2e^2\Delta}{\hbar \sqrt{\pi \lambda f k_B T}} \exp \left( -\frac{\lambda f}{4k_B T} \right). \tag{7}
\]
Evaluating this expression for $\lambda^r = 0.16$ eV and a measured contact conductance of 4.55 nS (1/220 MΩ) yields $\Delta = 5 \times 10^{-6}$ eV.

**Stationary current.** The stationary current through the protein can be written in terms of the fraction of oxidized states at each contact $n_i$ and the diffusion flux $J_d$ of charge carriers through the protein with $n_i$ satisfying the following current balance conditions:

\[
\begin{align*}
-n_1k_1^1c_s + (1 - n_1)k_0^1c_s - J_d &= 0 \\
-n_2k_2^2c_s + (1 - n_2)k_0^2c_s + J_d &= 0
\end{align*}
\]

where $J_d = -D \frac{\partial}{\partial x} \rho(x)$ is the diffusional flux due to the gradient of the bulk number density $\rho(x)$ and $D$ is the diffusion constant calculated above. Expressing the rates $k_1^1$, $k_2^2$, $k_0^1$, and $k_0^2$ (Eqns. S15) in dimensionless form leads to

\[
\frac{J_d}{J_0} = -a \frac{\delta}{L} [n_2 - n_1]
\]

where $J_0 = c_s \Delta / h = \Delta / (\hbar S)$ and $S$ is the contact area. The saturation current is given by $I_0 = eJ_0S = e\Delta / h \approx 1.2$ nA (using $\Delta = 5 \times 10^{-6}$ eV). The dimensionless parameter $a$ is given by

\[
a = \frac{D\hbar}{\delta^2 \Delta}
\]

Here, $\delta$ is the size of the entry/exit residue (adopted as 0.96 nm, representative of the two residues). With the values of $D$ calculated above (22.1-22.8 nm^2/ns), we obtain $a = 3.12 - 3.22$. 
Eqns. 8 and 9 can be solved for $n_1$ and $n_2$ to produce the diffusive current of holes and the corresponding conductivity:

$$n_1 = \frac{k^1_O + a \frac{k^1_O + k^2_O}{k^2}}{k^1 + a \frac{k^1}{(1 + \frac{k^1}{k^2})}}$$

$$n_2 = \frac{k^2_O + a \frac{k^1_O + k^2_O}{k^1}}{k^2 + a \frac{k^2}{(1 + \frac{k^2}{k^1})}}$$

(11)

where $k^1 = k_O(\eta) + k_R(\eta)$ and $k^2 = k_O(-\eta) + k_R(-\eta)$ as calculated from Eqns. S14, and $\ell = L/\delta$.

Equations 11 provide the solution for the stationary current through the protein:

$$\frac{l(V_b, \ell)}{I_0} = -\frac{a}{\ell} [n_2(\eta, \ell) - n_1(\eta, \ell)]$$

(12)

where the dependence of $\eta$ on $V_b$ and $\ell$ is given in the caption to Figure S2. Figure 5b plots $\frac{l(V_b, \ell)}{I_0}$ vs $V_b$ for two protein lengths (10 and 15 nm). With $I_0 \approx 1.2$ nA (as calculated from the contact resistance above) the calculated I-V curves are of the same magnitude as the examples given in Figure 1c, and feature a linear regime in the ± 200mV range as observed experimentally.

The resistance as a function of $\ell$ is given by

$$R = \left[ \frac{\partial l(V_b, \ell)}{\partial \eta} \frac{\partial \eta}{\partial V_b} \bigg|_{V_b=0} \right]^{-1}.$$  

(13)
This function is fitted (red line) to the experimental data (squares) in Figure 5c, with the best fit parameter $a = 3.8$, close to the values $(3.12 – 3.22)$ calculated above, showing that the present approach correctly predicts the current decay with distance. In a previous experimental work (13), the non-linear dependence of the molecular resistance on length was hypothesized to be intrinsic to the diffusion process. The analysis presented here predicts an inherently linear dependence, as it must be, given that the current is proportional to the carrier velocity, i.e., the ratio of diffusion time to molecular length. The observed non-linearity is mostly accounted for by the dependence of the potential distribution on the length of the molecule.

The absolute magnitude of the current is determined by $I_0$ which was calculated to be 1.2 nA via the estimate of the electronic coupling constant, $\Delta$, a parameter that the decay rate (Eqn. 10) also depends on. This estimate is based on the assumption of equal resistance of two contacts which can be merged into one residue to calculate the contact resistance. The fit to the resistance data (which are based on the peak of the distributions obtained from many current-voltage curves) is based instead on evaluating the diffusive current in Eqs. 9 and 12 and required $I_0 = 5.5$ nA, $\sim 4\times$ the value obtained from $\Delta$, but surprisingly close, given the limitations of the contact model.

The rather precise agreement between the calculated decay rate $a = 3.12 – 3.22$, and the value obtained from experimental data $(a = 3.8)$ is probably fortuitous, given both the lack of an exact theory for the diffusion process and other approximations used here: The force fields employed in MD are nonpolarizable and thus neglect screening of charges by induced molecular dipoles. An indication that this can cause minor inconsistencies is the fact that the mean ratio $\langle \lambda_{\text{ST}} / \lambda \rangle \approx 0.4$ is somewhat higher than $\approx 0.3$ from comparing Eq. 3 with the position of the conductance peak in Figure 1e. Dynamical heterogeneity, in particular at sites closer to the protein-water interface, can lead to faster dynamics and, correspondingly, to lower $g(R)$ in Eq. 4. Despite these limitations, the present theory demonstrates how long-range hopping is enabled through a reduced reorganization energy owing to non-ergodicity. The theory predicts the correct magnitude of current, and accurately predicts its decay with the length of the proteins.

**Conclusions**

The key result of this work is that proteins that do not contain redox cofactors can support long-range and rapid hopping of electrons, with implications for the role of aromatic amino acids in
facilitating charge transport. The same reduction of reorganization energy that accounts for this long-range transport also accounts for the alignment of the non-equilibrium charge shift states with the Fermi energy of some noble metals, important for the use of proteins as bioelectronic components. The reorganization energy appropriate to non-equilibrium oxidation states of tyrosine and tryptophan, ~ 0.16 eV, is significantly smaller than the Marcus reorganization energy derived from the Stokes shift (~0.8 eV). This reduced reorganization energy corresponds to a barrier (for $\Delta G = 0$) of 40 meV, which is close to available thermal energy. This reduction of reorganization energy has previously been demonstrated for charge transfer between cofactors (22, 24, 25) and is now extended to the aromatic amino acids. Even in the case of proteins rich in cofactors, the origin of the required reduction in reorganization energy remains obscure (15), so it is interesting to note that a reorganization energy below 0.2 eV and a fast diffusion constant (> 20 nm$^2$/ns), the conditions required for fast transport in OmcsS proteins, are obtained in the present case for transport a protein that lacks any redox cofactors. Finally, this reduction of the reorganization energy provides a plausible explanation for the fact that the current carrying states appear to be some 700 mV closer to the vacuum (i.e., at 300 mV vs NHE) than would be expected from the equilibrium oxidation potentials of tyrosine and tryptophan (1V vs NHE).

Methods

**MD protocols.** All MD simulations were performed using NAMD2 (41) periodic boundary conditions, the CHARMM36 force field (42) for the protein and the TIP3P model (43) for water. Van der Waals and short-range electrostatic forces were evaluated using a 1.2 nm cutoff whereas long range electrostatic forces were evaluated using the particle mesh Ewald method computed over a 0.12 nm-spaced grid (32). The initial equilibration of the CTPR8 system was performed using a 2 fs time step. All simulations of the CTPR8 systems containing an oxidized residue were performed using a 4 fs integration time step and the hydrogen mass repartition scheme (33). The equilibration simulations were performed in the constant number of particles, pressure (1 bar), and temperature (310K) ensemble maintained using a Langevin dynamics thermostat and the Nose-Hoover Langevin pressure control (44, 45). Visualization and analysis were performed using VMD (32).
MD simulations of CTPR8 systems. The atomic coordinates of a CTPR8 protein were taken from the published crystal structure (46). To reproduce the protein construct used in experiment (13), the N and C termini of the protein were extended by 31 and 18 amino acids, respectively, both terminating with a cysteine residue. The partially folded structure of the N-terminal addition was generated using the Iterative Threading ASSEMBly Refinement (I-TASSER) server (Zhang lab, University of Michigan) and its default parameters (47), the C-terminal addition was assumed to be disordered. The resulting atomic model of the experimental CTPR8 protein is provided as supplementary material, CTPR8-Cys.pdb. Missing hydrogen atoms were added to the protein using the psfgen plugin of VMD. The protein was immersed in a 8.0 x 8.6 x 7.1 nm³ pre-equilibrated volume of water using the Solvate Plugin of VMD (48). Potassium and chloride ions were added using the Autoionize plugin to produce 150 mM KCl solution. The system was minimized using the conjugate gradient method for 100 steps, which was followed by a 500 ns equilibration run. During the first 25 ns of the equilibration, the $C\alpha$ atoms of the protein were restrained to their initial coordinates using harmonic potentials with the force constant of 100 kcal/(mol nm²). The microscopic configuration of CTPR8 system obtained at the end of the 500 ns equilibration was used to construct thirteen CTPR8 systems each containing a different tryptophan or tyrosine residue in its oxidized state. These tyrosine and tryptophan residues were assumed to be oxidized to their cationic free radical state, carrying an overall charge of one proton. The partial charges of an oxidized tyrosine or an oxidized tryptophan residue were taken previous semi-empirical (29) and quantum mechanics/molecular mechanics (QM/MM) (30) calculations, respectively. The patches used to oxidize the residues are provided as supplementary material (oxidize_Y.txt and oxidize_W.txt), the values of the partial charges are listed in Table S4 and S5. Each system containing an oxidized residue was then equilibrated for 750 ns. Atomic coordinates were saved every 10 ps.

Calculation of the electrostatic potential at the active residue.

To calculate the energy-gap trajectory, we evaluated the electrostatic potentials at the location of the atoms of the redox-active residue every 10 ps using the method (32) implemented in the PMEpot Plugin of VMD. Briefly, every partial charge in the system was approximated by a spherical Gaussian
\[ \rho_i(r) = q_i \left( \frac{\beta}{\sqrt{\pi}} \right)^3 \exp(-\beta^2 (r - r_i)^2), \]  

(19)

where parameter \( \beta \) is the width of the Gaussian. The instantaneous electrostatic potential corresponding to the charge density was obtained by solving the Poisson equation

\[ \nabla^2 \phi(r) = 4\pi \sum_i \rho_i(r). \]  

(20)

The resulting electrostatic potentials \( \Phi_j \) at the location of atoms \( j \) (atoms whose charge changes during the charge transition) were extracted from the MD simulation of the oxidized and reduced residues. The electrostatic (Coulomb) component of the donor-acceptor energy gap was calculated as

\[ X^C = -\sum_j \Delta q_j^D \Phi_j + \sum_j \Delta q_j^A \Phi_j, \]  

(21)

where \( \Delta q_j^a = q_j^{a,\text{Red}} - q_j^{a,\text{Ox}} \) are differences of atomic charges in the reduced and oxidized states of the donor, \( a=D \), and acceptor, \( a=A \), moieties.

The Stokes-shift reorganization energy is given by the difference of average values of the energy gap in the initial and final charge-transfer states

\[ 2\lambda^{St} = \langle X \rangle_1 - \langle X \rangle_2. \]  

(22)

The average \( \langle X \rangle_1 \) was calculated from MD simulation of the reduced donor and oxidized acceptor whereas \( \langle X \rangle_2 \) was calculated from MD simulation of the oxidized donor and oxidized acceptor (see Figure 2c for the reaction diagram). The variance reorganization energy \( \lambda \) was determined as the mean of the energy-gap variances \( \lambda_i = \sigma^2_{\lambda_i}/2k_BT; \lambda = (\lambda_1 + \lambda_2)/2 \). The hopping rates (Table S2) were calculated according to Eqns. 4-5.
Evaluation of the dynamic crossover parameter $g$

The parameter $g$ is given by (34)

$$g = \frac{2\pi v^2 \tau_X}{\sigma_X \hbar} \frac{\exp \left( \frac{2}{\gamma^2 ((\delta R)^2)} \right)}{\sqrt{2 \Delta G^+ / k_B T + 4 (\tau_X / \tau_R)} \gamma^2 ((\delta R)^2)} \tag{24}$$

where $\Delta G^+$ is given by Eq. 1 in which $\lambda$ is replaced with $\lambda^r$ and the reaction free energy is from Eq. 3, $\sigma_X^2 = 2k_B T \lambda$, and $\tau_X$ and $\tau_R$ are the relaxation times of the Stokes-shift dynamics and of the donor-acceptor distance, respectively. The integrated relaxation times $\tau_{X,R}$ were calculated using the time correlation functions from MD (Figs S3 and S3 and Eqns. S6 and S7).

Acknowledgements

We thank Eathen Ryan for generating the protein structures. This work was supported by grant 1R01HG011079 from the National Human Genome Research Institute, grant CHE-2154465 from the National Science Foundation, grant W911NF2010320 from the US Army and the Edward and Nadine Carson Endowment. AA acknowledges support via grants from the National Science Foundation (PHY-1430124 and DMR-1827346). The supercomputer time was provided through the Extreme Science and Engineering Discovery Environment (XSEDE) allocation MCA05S028.

Competing Interest

S.L. is a co-founder of a company using technology based on protein conductivity.

References


Supplementary Information for

Long-Range Hopping Conductivity in Proteins

Siddharth Krishnan¹, Aleksei Aksimentiev¹†, Stuart Lindsay²,³,⁴ and Dmitry Matyushov²,³†

¹Department of Physics, and Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana Champaign, Urbana, IL, ²Department of Physics, ³School of Molecular Sciences and ⁴Biodesign Institute, Arizona State University, Tempe, AZ

Figure S1. Energy gap fluctuations of equivalent residues in CTPR. A) Y43 and its equivalent residues in the eight repeat units of CTPR8. The sidechain of the equivalent residues are highlighted using green spheres. The unit number for each of the residues is indicated next to the residue. B) Electrostatic energy gap fluctuation for the oxidation half reaction computed for equivalent tyrosine residues from the eight repeat subunits of CTPR. The distributions have been normalized.
Figure S2: Linear fit to the experimentally-determined bias distribution. \( K(L) \) is the fraction of the applied bias that appears across the molecule. The potential drop at the contacts is \( \eta(L) = V_b \frac{1 - K(L)}{2} \approx V_b \left( 0.5 - 0.02042 \frac{L}{\text{Å}} \right) \).

Stokes-shift and Distance Dynamics

The dynamical crossover parameter \( g \) accounts for the solvent dynamical control of the rate pre-exponential factor. Equation S1 shows that \( g \) includes two relaxation times: the integral relaxation times \( \tau_X \) for the Stokes-shift dynamics and the relaxation time \( \tau_R \) for the dynamics of the donor-acceptor distance

\[
g = \frac{2 \pi V_{DA}^2 \tau_X}{\sigma_X h} \frac{e^{\frac{3}{2} y^2 \langle (\delta R)^2 \rangle}}{\sqrt{2 \beta F_{\text{act}} + 4(\tau_X/\tau_R)y^2 \langle (\delta R)^2 \rangle}}
\]  

(S1)

In addition, \( \sigma_X^2 = 2 k_B T \gamma \), \( \gamma \) is the parameter of the exponential fall-off of the squared donor-acceptor electronic coupling and \( \langle (\delta R)^2 \rangle \) is the distance variance.

The Hopfield equation\(^8\) used to calculate \( V_{DA} \) is based on the edge-to-edge distance \( (R_{\text{edge}} \text{, in nm}) \) between the donor and acceptor

\[
V_{DA} = \frac{2.7 eV}{\sqrt{N_{DA} N_{D}}} e^{-7.2 R_{\text{edge}}}
\]  

(S2)

Here \( N_{DA} = 7 \) for tyrosine and \( N_{DA} = 9 \) for tryptophan. In addition, the activation energy \( \Delta G^\dagger \) in Eq. 1 is calculated as

\[
\Delta G^\dagger = \frac{(\lambda^r \pm \Delta G)^2}{4 \lambda^r}
\]  

(S3)

with “+” and “-” referring to the forward and backward transitions. In Eq. S3, \( \lambda^r \) is the reaction (superscript “\( r \)”) reorganization energy calculated from the equation\(^9\)

\[
\lambda^r = (\lambda^{St})^2 / \lambda
\]  

(S4)
where the variance reorganization energy $\lambda = \sigma_X^2 / 2k_BT$ is calculated from the variance of the energy gap reaction coordinate $X$. We take the mean of the forward and backward $\lambda$ to calculate the reaction reorganization energy $\lambda^R$ in Eq. S4. The Stokes-shift reorganization energy $\lambda^{St}$ is given by the difference of average values of the energy gap in the initial and final electron-transfer states

$$\lambda^{St} = \frac{\langle X_1 \rangle - \langle X_2 \rangle}{2}$$  \hspace{1cm} (S5)

Calculations of the Stokes-shift and donor-acceptor distance dynamics were based on MD simulations using NAMD$^2$ with a 4 fs integration time step and with hydrogen mass repartition.$^{10}$ Simulations in the NPT ensemble were performed using a Langevin dynamics thermostat and the Nose-Hoover Langevin$^{4,5}$ piston pressure control set at 310 K and 1 atm. The Stokes-shift dynamics were calculated using two MD trajectories, one being 1 ns in duration and sampled every 12 fs and another being 250 ns in duration and sampled every 10 ps. Both simulation systems probed the same electron transfer reaction between oxidized residue Y43 and neutral Y77. The energy gap was calculated as a function of time for the short and long trajectories. The two trajectories were combined while disposing the first 1ns of the longer trajectory. The Stokes-shift correlation function $C_X(t) = \langle \delta X(t) \delta X(0) \rangle$ was normalized and fitted to a sum of a Gaussian ballistic decay and two exponential functions

$$S_X(t) = C_X(t) = (1 - a - b) e^{-t^2 / \lambda_1} + ae^{-t / \lambda_2} + be^{-t / \lambda_3}$$ \hspace{1cm} (S6)

The MD data and the fit are shown in Fig S3. The fit was obtained with the parameters: $a = 0.085$, $b = 0.662$, $\lambda_1 = 0.0777$ ns$^2$, $\lambda_2 = 3.8 \times 10^{-3}$ ns, and $\lambda_3 = 6 \times 10^{-5}$ ns. These fitting parameters give the integral relaxation time of 20 ps.

The relaxation time for the dynamics of the donor-acceptor distance was calculated from a 65 ns trajectory. The center-of mass distance $R(t)$ between Y55—Y76 residue pair was sampled with the time step of 2.5 ps producing the correlation function

$$S_R(t) = \frac{C_R(t)}{C_R(0)} = e^{-2.953t^{0.27}}$$ \hspace{1cm} (S7)

where $C_R(t) = \langle \delta R(t) \delta R(0) \rangle$. The result is shown in Fig S4. The fit to Eq. S7 yields the integral relaxation time $\tau_R = 281$ ps.
**Figure S3**: Stokes-shift time correlation function. The dots and stars are the data from MD simulation, while the dashed line indicates the fit to Eq. 6. The grey stars indicate correlation data from 10ps sampling step and the black dots indicate correlation data during the 12fs step.

**Figure S4**: Distance time correlation function. The dots are the data from MD simulation, while the dashed line indicates the fit to Eq. 7.
Polarizability

We conduct the MD simulations in non-polarizable force fields, which do not account for screening of electrostatic interaction by electronic induced dipole moments. The standard dielectric theories account for the medium polarization in terms of the Pekar factor $c_0 = \frac{1}{\epsilon_\infty} - \frac{1}{\epsilon_s}$ which results in the reduction of the reorganization energy by a factor $\approx 2$ when the dielectric constant due to induced dipole $\epsilon_\infty$ is accounted for ($\epsilon_s$ is the static dielectric constant). Microscopic simulations and analytical theories$^{11,12,13}$ indicate that the reduction is smaller than suggested by continuum theories, but the problem has not been studied in depth for protein electron transfer.$^{14}$ With $\epsilon_\infty$ of water, the reduction compared to nonpolarizable solvents is 0.8 and this factor was used here to rescale the reorganization energies from MD simulations. The corrected $\lambda$ and $\lambda^g$ are the MD values (Table S1) multiplied by a factor of 0.8. As the result of this correction, the reaction reorganization energy in Eq 5 is also multiplied by 0.8.

Rate Calculations

Applying both relaxation times to Eq S1, we find that, due to a large relaxation time of distance fluctuations $\tau_R = 281$ ps, the term including $\tau_x/\tau_R$ in the denominator of Eq. 1 can be dropped. A simpler result for the crossover parameter $g$ follows

$$g = \frac{\pi \nu_{DA}^2 \tau_x}{\hbar \sqrt{2 \pi \tau_{act}}} e^{2\gamma^2 (\delta R)^2}$$  \hspace{1cm} (S8)

We also find that the parameter $g$ is large, $g \approx 55$ at $R \approx 0.6$nm, when $\gamma = 14.4$ nm$^{-1}$ and $\langle (\delta R)^2 \rangle = 0.041^2$ nm$^{-1}$ are adopted. The parameter $\gamma$ is taken from the Hopfield equation (Eq. S2) used to calculate the electronic coupling and the variance of the donor-acceptor distance is from MD simulations of the Y55-Y76 pair adopted for all other donor-acceptor pairs. This value of $g$ implies that, at relatively short distances, electron hops between the residues are in the dynamics controlled regime when the rate constant is not affected by the electronic coupling. This happens because the total reaction rate includes the parameter $g$ in the denominator of the following relation

$$k = k_{NA}/(1 + g)$$  \hspace{1cm} (S9)

Here, $k_{NA}$ is the non-adiabatic rate constant given by the equation

$$k_{NA} = \frac{\nu_{DA}^2}{\hbar} \sqrt{\frac{\pi}{k_B T \tau_x}} \exp\left[\frac{1}{2} \gamma^2 (\delta R)^2 - \frac{\Delta G^+}{k_B T}\right]$$  \hspace{1cm} (S10)

The dynamics-controlled electron-transfer rate constant becomes

$$k = \frac{1}{\tau_x} \sqrt{\frac{\Delta G^+}{k_B T \pi}} e^{-\gamma^2 (\delta R)^2 - \frac{\Delta G^+}{k_B T}}$$  \hspace{1cm} (S11)
The rate calculations between pairs of residues (Table S1) produce the parameter \( g \) spanning a broad range of values such that both the dynamics-controlled and nonadiabtic limits apply to specific intra-protein electron hops. The appearance of the crossover parameter \( g \) in the form given by Eqs. S1 and S8 leads to slight deviations from the detailed balance condition requiring the ratio of the forward and backward rate constant to be equal to the Boltzmann factor \( \exp[-\beta \Delta G] \).

Given that the Derrida model used here for the calculation of the carriers diffusion constant requires detailed balance, we used a somewhat simplified form of the parameter \( g \) replacing \( \Delta G \) in Eqs. 1 and 9 with \( \lambda^r/4 \). The correction is mostly insignificant for the rate values, but strictly preserves the detailed balance.

### Diffusion constant from the Derrida model

Following Derrida\(^1^\)\(^7\)

\[
D = \frac{N \Delta x^2}{\sum_{n=1}^{N} r_n} \sum_{n=1}^{N} k_{n+1,n} r_n u_n \tag{S12}
\]

where the step size \( \Delta x \) is calculated as the average distance between the residues and

\[
r_n = \frac{1}{k_{n+1,n}} \left[ 1 + \sum_{i=1}^{N-1} \prod_{j=1}^{i} k_{n+j, n+j+1} \right]^{-1}
\]

\[
u_n = \frac{1}{k_{n+1,n}} \left[ 1 + \sum_{i=1}^{N-1} \prod_{j=1}^{i} k_{n-j, n+1-j} \right]^{-1} \tag{S13}
\]

Applied to the path listed in Table S3, these equations yield \( D = 3.06 \text{ nm}^2/\text{ns} \). Similar numbers are found for other single paths. Applied to the 433 such paths connecting Y36 to Y89, we obtain a sum of the 433 \( D \) values of 578 nm\(^2/\text{ns} \).

### Charge Injection Calculation

The rate constants for the two contact sites, 1, 2 are

\[
k_R^{1,2} = \frac{\Delta}{\hbar} \text{erfc} \left( \frac{\lambda^r \mp e\eta}{\sqrt{4 \lambda^r k_B T}} \right) \tag{S14}
\]

\[
k_O^{1,2} = \frac{\Delta}{\hbar} \text{erfc} \left( \frac{\lambda^r \pm e\eta}{\sqrt{4 \lambda^r k_B T}} \right)
\]
where R and O refer to reduction or oxidation and we assume the same electronic coupling to the contacts, $\Delta$, and the potential drop at the contacts, $\eta$, is given in the caption to Figure S2. “−” and “+” refer to 1 and 2, respectively.

The kinetic equation for the fraction of holes $n$ at the single injection site is

$$\partial_t n(x) = k_R c_s n(x) - k_O c_s \left(1 - n(x)\right)$$  \hspace{1cm} (S15)

Where $k_R = k_R^1 + k_R^2$ and $k_O = k_O^1 + k_O^2$ and $c_s = I/\pi$ is the surface concentration determined by the area of the contact, $S$. At the stationary condition determined by $\partial_t n(x) = 0$, one gets $n = \frac{1}{2}$ and the following expression for the total current $I$ through the contact

$$I = \frac{e\Delta}{\hbar} \left[ \text{erfc}\left(\frac{\lambda^r - e}{\sqrt{4\lambda^r k_B T}}\right) - \text{erfc}\left(\frac{\lambda^r + e\eta}{\sqrt{4\lambda^r k_B T}}\right) \right]$$  \hspace{1cm} (S16)

**Kinetic Monte Carlo Model and Brownian diffusion**: We perform single charge kinetic Monte Carlo based on a previous routine implemented in python\textsuperscript{21} to simulate charge hopping in the graph based representation of the protein. Each node in the graph is connected by two edges, one for forward and one for backward rates. Every iteration of the Monte Carlo involves determining the lifetime of the charge on the node $i$ which it currently resides on and the node which the charge hops to next. The charge can only hop to the site which are connected to its current node by an edge. First, the lifetime of the charge $\tau$ is determined from an exponential random variable $T$ as $\tau = T/\Gamma_i$, where $\Gamma_i$ is the sum of the weights of the edges which originate at node $i$.

$$\Gamma_i = \sum_j \Gamma_{ij}$$  \hspace{1cm} (S17)

$\Gamma_{ij}$ is the weight of the edge connecting nodes $i$ to $j$ which is essentially the rate of charge transfer between the residues represented by $i$ and $j$. Next, to determine which residue, the charge hops to, we calculate the probability of hopping to that residue/node $j$. 
A random number $u$ is drawn from a uniform distribution and the next site $k$ is determined from the inequality.

$$\sum_{j<k} p_{ij} \leq u < \sum_{j<k} p_{ij}$$  \hspace{1cm} (S19)

The Monte Carlo (MC) simulation ends when the charge arrives at the residue in the protein which is in contact with the electrodes. We store the residence time $\tau$ for every hop and the distance $\Delta x_i$ between the center of mass of the residue sidechains travelled during every hop. We then calculate the average Brownian diffusion for that path.

$$D = \langle \Delta x_i^2 / 2\tau_i \rangle$$  \hspace{1cm} (S20)

The diffusion constant is calculated for 100000 MC with the starting residue as W35. Another 100000 MC runs are performed for Y36 as the starting residue, summarized Figure 6 of the main text. The resultant diffusion constant values are divided into 300 bins to obtain a distribution, which is fit to a gamma function probability distribution function given by the expression

$$PDF(x) = \frac{e^{-\frac{x}{\beta}}x^{\alpha-1}\beta^{-\alpha}}{\Gamma(\alpha)}$$  \hspace{1cm} (S21)

Here $\alpha$ and $\beta$ are parameters which determine the skewness and peak location of the distribution respectively. We assume $\beta$ to be the effective diffusion constant for the transport. This gives us 22.13 nm$^2$/ns for W35 and 22.8 nm$^2$/ns for Y36.
<table>
<thead>
<tr>
<th>Acceptor</th>
<th>Donor</th>
<th>$\lambda^{st}$ (eV)</th>
<th>$\lambda_{var}$ (eV)</th>
<th>$\Delta G$ (eV)</th>
<th>Edge Distance (Å)</th>
<th>$\lambda^r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>36</td>
<td>0.55</td>
<td>1.422</td>
<td>0.18</td>
<td>4.51</td>
<td>0.213</td>
</tr>
<tr>
<td>35</td>
<td>42</td>
<td>0.561</td>
<td>1.482</td>
<td>0.18</td>
<td>9.097</td>
<td>0.212</td>
</tr>
<tr>
<td>35</td>
<td>43</td>
<td>0.632</td>
<td>1.834</td>
<td>0.18</td>
<td>11.542</td>
<td>0.217</td>
</tr>
<tr>
<td>35</td>
<td>54</td>
<td>0.543</td>
<td>0.881</td>
<td>0.18</td>
<td>3.437</td>
<td>0.335</td>
</tr>
<tr>
<td>35</td>
<td>55</td>
<td>0.756</td>
<td>1.3</td>
<td>0.18</td>
<td>5.45</td>
<td>0.44</td>
</tr>
<tr>
<td>36</td>
<td>42</td>
<td>0.407</td>
<td>1.634</td>
<td>0.00</td>
<td>9.529</td>
<td>0.101</td>
</tr>
<tr>
<td>36</td>
<td>43</td>
<td>0.514</td>
<td>1.3</td>
<td>0.00</td>
<td>8.49</td>
<td>0.203</td>
</tr>
<tr>
<td>36</td>
<td>54</td>
<td>0.484</td>
<td>1.561</td>
<td>0.00</td>
<td>6.674</td>
<td>0.15</td>
</tr>
<tr>
<td>36</td>
<td>55</td>
<td>0.477</td>
<td>0.584</td>
<td>0.00</td>
<td>2.421</td>
<td>0.39</td>
</tr>
<tr>
<td>42</td>
<td>43</td>
<td>0.456</td>
<td>1.249</td>
<td>0.00</td>
<td>4.85</td>
<td>0.166</td>
</tr>
<tr>
<td>42</td>
<td>48</td>
<td>0.383</td>
<td>1.756</td>
<td>0.00</td>
<td>6.683</td>
<td>0.083</td>
</tr>
<tr>
<td>42</td>
<td>54</td>
<td>0.301</td>
<td>0.582</td>
<td>0.00</td>
<td>7.124</td>
<td>0.156</td>
</tr>
<tr>
<td>42</td>
<td>55</td>
<td>0.492</td>
<td>1.253</td>
<td>0.00</td>
<td>7.069</td>
<td>0.193</td>
</tr>
<tr>
<td>43</td>
<td>48</td>
<td>0.348</td>
<td>1.312</td>
<td>0.00</td>
<td>3.419</td>
<td>0.092</td>
</tr>
<tr>
<td>43</td>
<td>54</td>
<td>0.55</td>
<td>1.351</td>
<td>0.00</td>
<td>8.177</td>
<td>0.224</td>
</tr>
<tr>
<td>43</td>
<td>55</td>
<td>0.411</td>
<td>0.617</td>
<td>0.00</td>
<td>4.136</td>
<td>0.274</td>
</tr>
<tr>
<td>48</td>
<td>54</td>
<td>0.51</td>
<td>1.871</td>
<td>0.00</td>
<td>10.181</td>
<td>0.139</td>
</tr>
<tr>
<td>48</td>
<td>55</td>
<td>0.518</td>
<td>1.115</td>
<td>0.00</td>
<td>5.98</td>
<td>0.24</td>
</tr>
<tr>
<td>54</td>
<td>55</td>
<td>0.542</td>
<td>1.232</td>
<td>0.00</td>
<td>4.737</td>
<td>0.239</td>
</tr>
<tr>
<td>36</td>
<td>69</td>
<td>0.631</td>
<td>1.42</td>
<td>-0.18</td>
<td>8.594</td>
<td>0.28</td>
</tr>
<tr>
<td>36</td>
<td>70</td>
<td>0.382</td>
<td>1.211</td>
<td>0.00</td>
<td>7.384</td>
<td>0.12</td>
</tr>
<tr>
<td>36</td>
<td>88</td>
<td>0.542</td>
<td>1.589</td>
<td>0.00</td>
<td>9.799</td>
<td>0.185</td>
</tr>
<tr>
<td>43</td>
<td>70</td>
<td>0.465</td>
<td>1.512</td>
<td>0.00</td>
<td>7.35</td>
<td>0.143</td>
</tr>
<tr>
<td>43</td>
<td>76</td>
<td>0.528</td>
<td>1.235</td>
<td>0.00</td>
<td>7.441</td>
<td>0.225</td>
</tr>
<tr>
<td>43</td>
<td>77</td>
<td>0.524</td>
<td>1.15</td>
<td>0.00</td>
<td>6.308</td>
<td>0.239</td>
</tr>
<tr>
<td>43</td>
<td>88</td>
<td>0.529</td>
<td>1.364</td>
<td>0.00</td>
<td>10.328</td>
<td>0.205</td>
</tr>
<tr>
<td>43</td>
<td>89</td>
<td>0.645</td>
<td>1.521</td>
<td>0.00</td>
<td>6.903</td>
<td>0.274</td>
</tr>
<tr>
<td>48</td>
<td>70</td>
<td>0.328</td>
<td>1.529</td>
<td>0.00</td>
<td>12.525</td>
<td>0.07</td>
</tr>
<tr>
<td>48</td>
<td>76</td>
<td>0.465</td>
<td>1.677</td>
<td>0.00</td>
<td>7.543</td>
<td>0.129</td>
</tr>
<tr>
<td>48</td>
<td>77</td>
<td>0.462</td>
<td>1.535</td>
<td>0.00</td>
<td>8.233</td>
<td>0.139</td>
</tr>
<tr>
<td>55</td>
<td>69</td>
<td>0.809</td>
<td>1.3</td>
<td>-0.18</td>
<td>6.901</td>
<td>0.503</td>
</tr>
<tr>
<td>55</td>
<td>70</td>
<td>0.544</td>
<td>0.983</td>
<td>0.00</td>
<td>9.022</td>
<td>0.301</td>
</tr>
<tr>
<td>55</td>
<td>76</td>
<td>0.654</td>
<td>1.087</td>
<td>0.00</td>
<td>6.001</td>
<td>0.393</td>
</tr>
<tr>
<td>55</td>
<td>77</td>
<td>0.45</td>
<td>0.964</td>
<td>0.00</td>
<td>8.64</td>
<td>0.21</td>
</tr>
<tr>
<td>55</td>
<td>88</td>
<td>0.697</td>
<td>1.126</td>
<td>0.00</td>
<td>7.775</td>
<td>0.432</td>
</tr>
<tr>
<td>55</td>
<td>89</td>
<td>0.768</td>
<td>0.949</td>
<td>0.00</td>
<td>8.734</td>
<td>0.621</td>
</tr>
</tbody>
</table>

**Table S1:** Acceptor (first column) and donor (second column) residue numbers and the reorganization energies calculated from directly from the MD simulations without accounting for the polarization. Also listed (sixth column) the edge-to-edge distance between the residue pairs used in the Hopfield equation. The last column lists the reaction reorganization energies.
<table>
<thead>
<tr>
<th>Donor</th>
<th>Acceptor</th>
<th>Forward Rate (/ns)</th>
<th>Backward Rate (/ns)</th>
<th>g</th>
<th>R (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>35</td>
<td>36</td>
<td>0.024</td>
<td>25.621</td>
<td>325.759</td>
</tr>
<tr>
<td>2</td>
<td>35</td>
<td>42</td>
<td>0.007</td>
<td>7.848</td>
<td>0.441</td>
</tr>
<tr>
<td>3</td>
<td>35</td>
<td>43</td>
<td>0.0</td>
<td>0.327</td>
<td>0.012</td>
</tr>
<tr>
<td>4</td>
<td>35</td>
<td>54</td>
<td>0.023</td>
<td>24.311</td>
<td>975.706</td>
</tr>
<tr>
<td>5</td>
<td>35</td>
<td>55</td>
<td>0.015</td>
<td>15.964</td>
<td>41.042</td>
</tr>
<tr>
<td>6</td>
<td>36</td>
<td>42</td>
<td>3.183</td>
<td>3.183</td>
<td>0.628</td>
</tr>
<tr>
<td>7</td>
<td>36</td>
<td>43</td>
<td>3.072</td>
<td>3.072</td>
<td>1.42</td>
</tr>
<tr>
<td>8</td>
<td>36</td>
<td>54</td>
<td>6.589</td>
<td>6.589</td>
<td>26.219</td>
</tr>
<tr>
<td>9</td>
<td>36</td>
<td>55</td>
<td>1.702</td>
<td>1.702</td>
<td>4661.247</td>
</tr>
<tr>
<td>10</td>
<td>42</td>
<td>43</td>
<td>6.31</td>
<td>6.31</td>
<td>327.05</td>
</tr>
<tr>
<td>11</td>
<td>42</td>
<td>48</td>
<td>8.476</td>
<td>8.476</td>
<td>45.609</td>
</tr>
<tr>
<td>12</td>
<td>42</td>
<td>54</td>
<td>6.165</td>
<td>6.165</td>
<td>13.142</td>
</tr>
<tr>
<td>13</td>
<td>42</td>
<td>55</td>
<td>5.081</td>
<td>5.081</td>
<td>11.552</td>
</tr>
<tr>
<td>14</td>
<td>43</td>
<td>48</td>
<td>8.481</td>
<td>8.481</td>
<td>4563.516</td>
</tr>
<tr>
<td>15</td>
<td>43</td>
<td>54</td>
<td>3.125</td>
<td>3.125</td>
<td>2.025</td>
</tr>
<tr>
<td>16</td>
<td>43</td>
<td>55</td>
<td>3.502</td>
<td>3.502</td>
<td>559.671</td>
</tr>
<tr>
<td>17</td>
<td>48</td>
<td>54</td>
<td>1.095</td>
<td>1.095</td>
<td>0.18</td>
</tr>
<tr>
<td>18</td>
<td>48</td>
<td>55</td>
<td>4.176</td>
<td>4.176</td>
<td>44.744</td>
</tr>
<tr>
<td>19</td>
<td>54</td>
<td>55</td>
<td>4.293</td>
<td>4.293</td>
<td>269.791</td>
</tr>
<tr>
<td>20</td>
<td>36</td>
<td>69</td>
<td>11.08</td>
<td>0.01</td>
<td>0.692</td>
</tr>
<tr>
<td>21</td>
<td>36</td>
<td>70</td>
<td>7.127</td>
<td>7.127</td>
<td>11.684</td>
</tr>
<tr>
<td>22</td>
<td>36</td>
<td>88</td>
<td>1.104</td>
<td>1.104</td>
<td>0.237</td>
</tr>
<tr>
<td>23</td>
<td>43</td>
<td>70</td>
<td>6.422</td>
<td>6.422</td>
<td>10.354</td>
</tr>
<tr>
<td>24</td>
<td>43</td>
<td>76</td>
<td>3.963</td>
<td>3.963</td>
<td>5.817</td>
</tr>
<tr>
<td>25</td>
<td>43</td>
<td>77</td>
<td>4.153</td>
<td>4.153</td>
<td>28.054</td>
</tr>
<tr>
<td>26</td>
<td>43</td>
<td>88</td>
<td>0.469</td>
<td>0.469</td>
<td>0.099</td>
</tr>
<tr>
<td>27</td>
<td>43</td>
<td>89</td>
<td>3.208</td>
<td>3.208</td>
<td>10.422</td>
</tr>
<tr>
<td>28</td>
<td>48</td>
<td>70</td>
<td>0.104</td>
<td>0.104</td>
<td>0.011</td>
</tr>
<tr>
<td>29</td>
<td>48</td>
<td>76</td>
<td>6.703</td>
<td>6.703</td>
<td>8.679</td>
</tr>
<tr>
<td>30</td>
<td>48</td>
<td>77</td>
<td>5.376</td>
<td>5.376</td>
<td>2.989</td>
</tr>
</tbody>
</table>
Table S2: Acceptor (second column) and donor (third column) residue numbers and forward (4th column) and backward rate (5th column). The 6th column lists the crossover parameter g and the last column lists the average COM-COM distance between the residues (averaged over 3000 MD frames). The rates have been calculated with the adjusted reorganization energy to account for the non-polarizable MD forcefield which was used.

<table>
<thead>
<tr>
<th></th>
<th>Residue Pairs</th>
<th>Forward Rate (ns&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>Backward Rate (ns&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>g</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>35-36</td>
<td>0.024</td>
<td>25.621</td>
<td>325.759</td>
</tr>
<tr>
<td>2</td>
<td>36-43</td>
<td>3.072</td>
<td>3.072</td>
<td>1.42</td>
</tr>
<tr>
<td>3</td>
<td>43-89</td>
<td>3.208</td>
<td>3.208</td>
<td>10.422</td>
</tr>
<tr>
<td>4</td>
<td>89-35</td>
<td>9.752</td>
<td>0.009</td>
<td>4.441</td>
</tr>
</tbody>
</table>

Table S3: Forward and backward rates and the crossover parameter g for transitions in a path from the start of first unit of CTPR8 protein to the last active residue in the second CTPR8 protein. The path was calculated using Dijkstra's shortest path algorithm<sup>13</sup>. Derrida's model<sup>16</sup> is applied to the path to calculate the diffusion constant of 3.06 nm<sup>2</sup>/ns.
Table S4: Partial charges for Y-H⁺ in units of the elementary charge¹⁹. Atom names are consistent with the CHARMM36 convention. The total sum of the partial charges is +1. The partial charges are applied to the residue and the protein is simulated using NAMD. The trajectory is used to calculate the electrostatic potential map of either the initial or final state during the charge transfer process. The initial state has the electron acceptor residue positively charged while the final state has the donor residue positively charged.

<table>
<thead>
<tr>
<th>Atom</th>
<th>Charge (e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CB</td>
<td>-0.18000000715255737</td>
</tr>
<tr>
<td>HB1</td>
<td>0.09000000357627869</td>
</tr>
<tr>
<td>HB2</td>
<td>0.09000000357627869</td>
</tr>
<tr>
<td>CG</td>
<td>0.42800000309944153</td>
</tr>
<tr>
<td>CD1</td>
<td>-0.23000000417232513</td>
</tr>
<tr>
<td>HD1</td>
<td>0.22300000488758087</td>
</tr>
<tr>
<td>CE1</td>
<td>-0.23000000417232513</td>
</tr>
<tr>
<td>HE1</td>
<td>0.22300000488758087</td>
</tr>
<tr>
<td>CZ</td>
<td>0.6209999918937683</td>
</tr>
<tr>
<td>OH</td>
<td>-0.4300000071525574</td>
</tr>
<tr>
<td>HH</td>
<td>0.4090000092983246</td>
</tr>
<tr>
<td>CD2</td>
<td>-0.23000000417232513</td>
</tr>
<tr>
<td>HD2</td>
<td>0.22300000488758087</td>
</tr>
<tr>
<td>CE2</td>
<td>-0.23000000417232513</td>
</tr>
<tr>
<td>HE2</td>
<td>0.11500000208616257</td>
</tr>
</tbody>
</table>
Table S5 Partial charges for W-H\(^+\) in units of the elementary charge\(^{20}\). Atom names are consistent with the CHARMM36 convention. The total sum of the partial charges is +1. The partial charges are applied to the residue and the protein is simulated using NAMD. The trajectory is used to calculate the electrostatic potential map of either the initial or final state during the charge transfer process. The initial state has the electron acceptor residue positively charged while the final state has the donor residue positively charged.

<table>
<thead>
<tr>
<th>Atom</th>
<th>Charge (e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CB</td>
<td>0.07000000029802322</td>
</tr>
<tr>
<td>HB1</td>
<td>0.009999999776482582</td>
</tr>
<tr>
<td>HB2</td>
<td>0.009999999776482582</td>
</tr>
<tr>
<td>CG</td>
<td>0.11999999731779099</td>
</tr>
<tr>
<td>CD1</td>
<td>0.12999999523162842</td>
</tr>
<tr>
<td>HD1</td>
<td>0.00499999888241291</td>
</tr>
<tr>
<td>NE1</td>
<td>-0.14000000059604645</td>
</tr>
<tr>
<td>HE1</td>
<td>0.4699999988079071</td>
</tr>
<tr>
<td>CE2</td>
<td>-0.00999999776482582</td>
</tr>
<tr>
<td>CD2</td>
<td>-0.07000000029802322</td>
</tr>
<tr>
<td>CE3</td>
<td>0.02999999329447746</td>
</tr>
<tr>
<td>HE3</td>
<td>0.00499999888241291</td>
</tr>
<tr>
<td>CZ3</td>
<td>0.15000000059604648</td>
</tr>
<tr>
<td>HZ3</td>
<td>0.02999999329447746</td>
</tr>
<tr>
<td>CZ2</td>
<td>0.12999999523162842</td>
</tr>
<tr>
<td>HZ2</td>
<td>0.00499999888241291</td>
</tr>
<tr>
<td>CH2</td>
<td>0.05000000074505806</td>
</tr>
<tr>
<td>HH2</td>
<td>0.00499999888241291</td>
</tr>
</tbody>
</table>

References:


