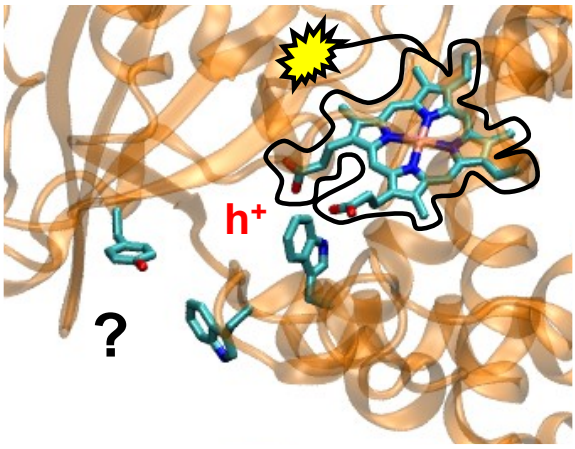
Mapping Hole Hopping Escape Routes in Proteins

Ruijie D. Teo,[†,§] Ruobing Wang,[†,‖,§] Elizabeth R. Smithwick,[†,§] Agostino Migliore,[†] Michael J. Therien,[†] David N. Beratan[†,‡,¶,*]*

[†] Department of Chemistry, Duke University, Durham, North Carolina 27708, United States
[‖] Current address: Science Center, Opera Solutions OPCO, LLC, San Diego, California 92130, United States
[‡] Department of Biochemistry, Duke University, Durham, North Carolina 27710, United States
[¶] Department of Physics, Duke University, Durham, North Carolina 27708, United States
[§] These authors contributed equally.

ABSTRACT

A recently proposed oxidative damage protection mechanism in proteins relies on hole hopping escape routes formed by redox-active amino acids. We present a computational tool to identify the dominant charge hopping pathways through these residues based on the mean residence times of the transferring charge along these hopping pathways. The residence times are estimated by combining a kinetic model with well-known rates expressions for the charge-transfer steps in the pathways. We identify the most rapid hole hopping pathways escape routes in cytochrome P450 monooxygenase ($P450_{BM3}$), cytochrome *c* peroxidase (Ccp1), and benzylsuccinate synthase (BSS). This theoretical analysis supports the existence of hole-hopping chains as a mechanism capable of supporting hole escape from protein catalytic sites on biologically relevant timescales. Furthermore, we find that pathways involving the [4Fe4S] cluster as the terminal hole acceptor in BSS are accessible on the millisecond timescale, suggesting a potential protective role of redox-active cofactors for preventing protein oxidative damage.



KEYWORDS. Electron-hole transfer, protein, oxidative damage, ROS, redox hopping pathway.

SIGNIFICANCE PARAGRAPH

Hole hopping pathways consisting of redox-active amino acids are proposed to direct potentially damaging holes away from active sites of proteins to the protein surface where they may be neutralized by reducing species. We developed an approach to rapidly identify hole hopping pathways and to quantify the time scale for their shepherding of holes to the protein surface. Using this theoretical methodology, we identify the dominant hopping pathways in cytochrome P450, cytochrome c peroxidase, and benzylsuccinate synthase (BSS). Hopping pathways that involve a high-potential [4Fe4S] cluster as the terminal hole acceptor in BSS were identified to be accessible on a millisecond timescale, highlighting for the first time an [4Fe4S] cluster that is kinetically poised to play an oxidative damage protective role in a protein.

INTRODUCTION

Proteins that incorporate oxygen into organic substrates from molecular oxygen and other potent oxidants, such as H_2O_2, are vital to biological function. In fact, these oxidants are often short-lived

products of cellular biochemical reactions. Redox reactions involving these oxidants are susceptible to forming reactive oxygen species (ROS), which can lead to oxidative damage of DNA and proteins (1). ROS damage in proteins can manifest as backbone oxidation (2), protein fragmentation (2), or amino acid side-chain oxidation; cysteine, methionine, tryptophan, and tyrosine are the most susceptible to oxidation (3, 4). For example, tyrosine oxidation can promote dityrosine-linked aggregate formation in the eye lens protein γβ-crystallin, tertiary structure destabilization, and cataractogenesis (5). Studies also link the effects of ROS damage to aging complications and to cancer (2). Protein oxidation can lead to unfolding and, ultimately, to the loss of function (6). Although proteasomes and lysosomes help to degrade damaged proteins (2), the accumulation of oxidized protein can contribute to Alzheimer's disease and diabetes (7). Here, we describe a computational method to identify hole hopping chains and to estimate the timescale with which these pathways may convey otherwise damaging oxidizing equivalents to a protein surface where they may be reduced by diffusible reducing species.

Recent studies of Gray and Winkler (3, 4, 8) investigated a ROS protective mechanism that may act within monooxygenases, dioxygenases, and peroxidases. They proposed that chains of redox-active amino acids, including Tyr, Trp, Cys, and Met (Fig. S1 and S2), can support hole hopping to direct strongly oxidizing electron holes to the protein surface in the absence of the enzyme substrate (3, 9), thus preventing damage to the protein active site. Once the excess charge is diverted to the exterior of the protein, reductants within the cell may be able to fill the holes. This hypothesis regarding molecular approaches to "defusing redox bombs" (10) was supported by a study of the Research Collaboratory for Structural Bioinformatics Protein Data Bank (PDB) (3), which revealed that one third of the surveyed proteins contained chains of 3-to-5 aromatic amino acids, with the highest occurrence in redox proteins such as oxidoreductases and hydrolases

(3).

In addition to protecting redox-active sites from oxidative damage, hole hopping pathways could also act to safeguard proteins from labilized hemes that may be produced in over-oxidized proteins like cytochrome *c* peroxidase (Ccp1). Labilized heme may be produced when H_2O_2 levels in Ccp1 (e.g., from *Saccharomyces cerevisiae*) increase ~10-fold with yeast respiration; in this mechanism, Ccp1 is activated via irreversible oxidation of the axial H175 ligand to transfer its heme to apo-catalase A, after which apo-Ccp1 escapes the mitochondria (11, 12). Irreversible oxidation of heme can therefore take place when H_2O_2 levels are high and hole escape routes are absent (12).

Solvent-exposed high-potential $[4Fe4S]^{2+/3+}$ clusters may serve as terminal hole acceptors for a through-protein oxidative damage protection chain (13). Although it was initially proposed that the Gly829$^{\bullet}$ radical cofactor could be reduced by the $[4Fe4S]^{2+}$ cluster in the glycyl radical enzyme benzylsuccinate synthase (BSS) (14), the 31.2 Å edge-to-edge separation between Gly829$^{\bullet}$ and the cluster (PDB 4PKF) suggested that this distance could not enable a HT-based protective mechanism (15). Despite the large transfer distance, experimental studies of other proteins (4) indicate that this hole transfer (HT) reaction may be accessible. Electron tunneling over distances of 15-20 Å can occur on a nanosecond to microsecond timescale in iron and copper proteins (3, 4), and hopping (multi-step tunneling) can enable charge transfer (CT) well over longer distances on a sub-millisecond time scale when the free energy profile is favorable (16).

The informatic-computational approach described here is used to identify the fastest hole hopping routes that may limit ROS formation in cytochrome P450 (from *Bacillus megaterium,* $P450_{BM3}$), Ccp1, and BSS. This methodology can be used more broadly for rapid database

screening in the context of oxidative protection pathways, and can also be used to explore hole hopping in signaling pathways that are proposed for protein-nucleic acid complexes (17, 18).

RESULTS AND DISCUSSION

We investigated hopping routes for $P450_{BM3}$, Ccp1, and BSS with transit times on biologically-relevant timescales using the search program described here (EHPath). The program identifies hole hopping routes and estimates the timescale for hole hopping transfer or escape. The Python-based program implements a kinetic model (17) outlined below. EHPath allows the choice of electron transfer kinetic parameters. The fastest hole escape routes (or, more generally, charge-transfer routes) are identified, and the transit time is estimated. Estimates of these times allow us to resolve the outstanding question of whether such hopping paths could establish viable signaling or oxidative damage protection mechanisms. The viability of the [4Fe4S] clusters to serve as hole traps in BSS is addressed, and the biological relevance of the differences in the escape times for $P450_{BM3}$ and Ccp1 are discussed.

The HT rate constant k_{DA} for each donor (D)-acceptor (A) pair in a HT route was estimated using the nonadiabatic CT rate expression in the high temperature (Marcus) limit (19):

$$k_{DA} = \frac{2\pi}{\hbar}\langle {V_{DA}}^2 \rangle \frac{1}{\sqrt{4\pi\lambda_{DA}T}} \mathrm{e}^{-\frac{(\Delta G^\circ + \lambda_{DA})^2}{4\lambda_{DA}k_B T}}. \qquad (1)$$

V_{DA} is the effective electronic coupling between the charge states localized on D and A, ΔG° is the reaction free energy, λ_{DA} is the reorganization energy, and T is the temperature. The reorganization energy for each CT step was approximated as (20, 21)

$$\lambda_{DA} = \frac{\lambda_{DD} + \lambda_{AA}}{2} = \lambda_D + \lambda_A \qquad (2)$$

Here, λ_{DD} and λ_{AA} are the reorganization energies for non-adiabatic self-exchange weak-overlap reactions in the *D*-*D* and *A*-*A* redox pairs, respectively. The contribution of redox species *S* (*S* = *D*, *A*) to the reorganization energy, λ_S is defined as half of the reorganization energy λ_{SS} for self-exchange in the *S*-*S* system. The λ_{SS} values for the redox-active residues and cofactors at the relevant center-to-center distances R_{DA} were derived from available literature values at selected center-to-center distances $R_{DA}{}'$ (summarized in Table S1) (22) via the equation (17)

$$\lambda_{ss} = \lambda'_{ss} + \left(\frac{1}{\varepsilon_o} - \frac{1}{\varepsilon_s}\right)(\Delta q)^2 \frac{R_{DA} - R_{DA}{}'}{R_{DA} R_{DA}{}'}. \qquad (3)$$

Here, ε_o and ε_s are the optical (2.2 (23)) and static (4.0 (17, 23-25)) dielectric constants describing the environments of the *D* and *A* species, and Δq is the transferring charge (equal to the magnitude of the electron charge here). To define the *D* and *A* moieties, the redox-active amino acids Cys, Met, Tyr, and Trp were truncated from their amide backbones and represented as methanethiol, thioether, phenol, and indole, respectively. Then, the *D* and *A* centers were defined as the averages of the heavy atom PDB coordinates in the truncated moieties. R_{DA} was calculated as the distance between the *D* and *A* centers.

$\Delta G°$ for the HT step from *D* to *A* was computed from the difference of the *A* and *D* oxidation potentials (Table S1). The oxidation potentials for the redox groups in this study were obtained as described in ref. (17), while the oxidation potential for heme was taken from ref. (3). V_{DA} was estimated using Hopfield's formula (26) (Eq. S2) for the amino acid residues and other semiempirical methods were used to estimate couplings for the heme (Eq. S3) and iron-sulfur cluster (Eq. S4) cofactors. Use of these semiempirical methods is justified and motivated by the following considerations: (i) the electronic coupling values previously obtained using

semiempirical methods showed acceptable order of magnitude agreement with more accurate DFT results (17); (ii) the aim here is to accomplish qualitative rapid screening of hopping pathways in large protein databases; (iii) the order of magnitude timescales computed for hole delivery to a protein surface do not depend strongly on the electronic couplings, while they can depend more critically on the free energy parameters, as is shown in Table 1 (without changing the conclusions of our analysis).

Points (i)-(iii), taken together, justify our use of semiempirical approaches to electronic coupling computations for rapid screening of vast databases with reasonable accuracy. However, more accurate evaluations of the charge-transfer parameters (including the electronic couplings) for specific systems are needed to discriminate between HT pathways relevant to the chemical function whose speeds differ by a few orders of magnitude. A starting point to improve on the coupling computation is proposed in ref. (17), where the ab initio calculation of the couplings for gas-phase redox-active groups is combined with a semiempirical treatment of the intervening medium.

Molecular motion may influence protein geometries and the nature of the predicted hopping pathways; if conformational changes are particularly large, these geometry changes may gate the hopping process, just as gating motions may limit single-step tunneling (27). Nonetheless, such motion is unlikely to affect the general conclusions on the feasibility of the identified charge-transport mechanisms, while the details of some pathways could change (17, 28).

We compared the k_{DA} values of Eq. 1 with values obtained using the empirical square barrier tunneling (SBT) rate expression (3),

$$k_{SBT} = 10^{13} \mathrm{e}^{-\beta(r-r_0)} \mathrm{e}^{-\frac{(\Delta G^\circ + \lambda_{DA})^2}{4\lambda_{DA} k_B T}} \qquad (4)$$

In Eq. 1, r_0 is a *D-A* contact distance of 3 Å, r is the *D-A* edge-to-edge distance, and β is the distance decay factor (1.1 $Å^{-1}$) for the *D-A* electronic coupling (29). The prefactor of 10^{13} s^{-1} was fit to kinetic measurements in Ru-modified azurins (29-31) (30). Rather than calculating the Marcus free energy parameters for every HT step (using Eqs. 2-4 and S2-S4), ref. (3) used approximate values for Tyr and Trp hopping in heme proteins: λ_{DA} = 0.8 eV, $\Delta G° = 0.1$ eV for HT from the heme to Trp or Tyr, $\Delta G° = 0$ eV for the intermediate HT step(s), and $\Delta G° = -0.1$ eV for the final HT to the terminal Trp or Tyr residue (3). This simple parameter choice was used with Eq. 4 to analyze Tyr/Trp hopping pathways. The simplicity of Eq. 4 enables the rapid evaluation of the CT rate constants. Comparing the results from different CT parameter choices and models (i.e., Eq. 1 and 4), we will examine the robustness of the hopping pathway analysis.

Previous studies (22, 32) found that the reorganization energy for HT between redox-active residues can be much larger than 0.8 eV. One expects (33, 34) larger values because the size of the redox active groups is much smaller for amino acid residues than for most protein cofactors. Indeed, the λ_{DA} values for HT between amino acids are computed to range from 1.38 eV to 2.04 eV (see Table S6 of ref. (17)), while experimental estimates of λ_{DA} for these systems are yet to be determined. We will explore the sensitivity of our conclusions regarding the dominant hole escape pathways and timescales as we change the reorganization energy from ~0.8 eV to significantly higher values, up to ~2 eV, predicted using Eq. 3 in conjunction with theoretical estimates in the literature [see refs. (22, 35, 36) for further discussion of the reorganization energy parameters]. The EHPath program allows the user to define λ_{DA} in such ranges for all *D-A* pairs.

We calculated the mean residence time τ of the transferring charge in each pathway (that is, the inverse of the effective start-to-finish transfer rate) by inserting the k_{DA} or k_{SBT} values into the expression (17)

$$\tau = \sum_{n=0}^{N} \tau_n = \sum_{n=0}^{N-1} \frac{1}{k_{n\to n+1}} \left(\sum_{j=0}^{N-n-1} \prod_{i=n+1}^{N-j} \frac{k_{i\to i-1}}{k_{i\to i+1}} + 1 \right) + \frac{1}{k_{N\to N+1}} \qquad (5)$$

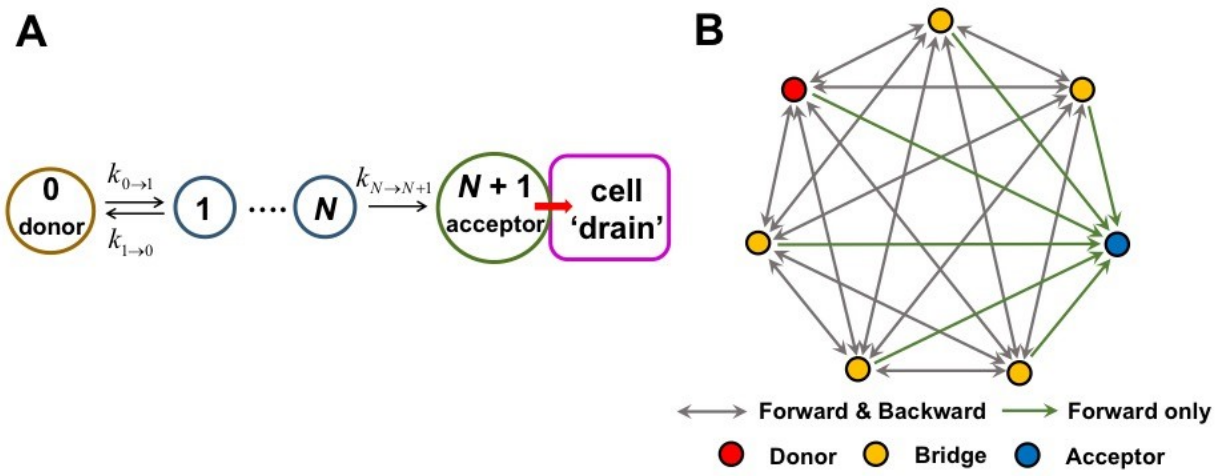
N is the total number of redox-active bridge sites in a pathway and $k_{n\to n\pm1}$ is the rate constant for CT between consecutive redox sites for any given realization of scheme 1B (i.e., for any given hopping pathway through a protein's 3D network of possible hole hopping routes). $n = 0$ denotes the initial charge donor, $n = 1$ to *N* indicates intervening redox-active residues, and site $n = N + 1$ is the terminal charge acceptor in contact with a charge 'drain'. This kinetic model includes both forward and backward HT steps in the hopping pathway, except for the charge-transfer step between the last redox-active site in the protein and the terminal acceptor, which is irreversible due to rapid scavenging of the charge by redox agents in the cellular environment (17) (see scheme 1A). When all backward CT rates can be neglected (i.e., $k_{n+1\to n} \ll k_{n\to n+1}$ for $0 \le n < N$), Eq. 5 reduces to the well-known expression (37)

$$\tau_{approx} \cong \sum_{n=0}^{N} \frac{1}{k_{n\to n+1}}, \qquad (6)$$

That is, the average time spent on the hopping chain by the transferring charge is approximately the sum of the mean residence times at the individual redox-active sites. Eq. 5 does not take into account the possibility of pathway branching, since the occupation probability of each redox group only depends on its coupling to the adjacent redox sites along a given HT pathway. We use this kinetic model since: (a) the protein motion can decouple the HT pathways, so that a given escape route is realized each time; (b) branching would appreciably mix only the most rapid HT routes with similar mean residence times (through protein regions rich in redox-active amino acids), and would not change our assessment of the feasibility for oxidative damage protection on sufficiently

rapid timescales. The kinetic model indicated in Scheme 1A (Eq. 5) and the pathway search indicated in Scheme 1B enable the fast screening of protein structural databases for proteins with possible protective structures, and will identify the most probable hole escape route in most situations.

In the following analysis, τ_M (τ_{SBT}) denotes the mean residence time obtained using the rate expression in Eq. 1 (Eq. 4), with the $\Delta G°$ values obtained from the oxidation potential differences and the reorganization energies of Eq. 2-3. The mean residence times using rate Eq. 4, with free energy parameters from ref. (3) are denoted τ'_{SBT}. Since the $\Delta G°$ estimates and analysis in ref. (3) were limited to pathways involving Tyr and Trp, we only calculated k_{SBT} and τ'_{SBT} for Tyr/Trp-based pathways.



Scheme 1. (**A**) Kinetic model leading to Eq. 5 (17). The cellular environment provides redox species that rapidly fill the hole that arrives at the final acceptor on the protein surface. Therefore, backward charge transfer from site $N+1$ to site N cannot occur. (**B**) Structure of a directed graph representing hopping routes in a protein, where the nodes represent the initial charge donor site (red), the intermediate hopping sites (yellow), and the acceptor site (blue). As the acceptor is in contact with a charge 'drain' in the kinetic model, there is no directed edge from acceptor to bridge (namely, no backward charge transfer to the bridge).

The *Pathways* method to estimate bridge mediated tunneling interactions uses graph theory to identify tunneling routes that maximize V_{DA} (38-40). Other CT pathway finders rely on QM/MM methods to identify tunneling pathways (41). EHPath focuses, instead, on the case of hopping transport. This hopping pathway finder: (i) identifies bidirected graphs that correspond to specific protein structures, (ii) estimates ΔG° and λ_{DA} values for all hopping steps, and (iii) calculates the mean times required for the charge hopping routes from D to A. The structure of a typical graph is shown in Scheme 1B. The rates for forward and reverse CT between nodes define the effective lengths of edges (in the two CT directions) in the corresponding graphs that are used in the minimum effective-length (minimum CT travel time) path search. The mean residence time for a charge hopping pathway is computed using Eq. 5 (Eq. 6 is also implemented) for user defined redox-active residues and cofactors (see Section S4). The pathway search uses NetworkX, a python package for network analysis based on graph theory (42), to rank the hopping pathways based on mean residence time (Section S4).

We tested the EHPath code by replicating the hopping pathways and mean residence times identified in earlier analysis [see Table 1 of ref. (17)]. The computations also identified weaker hole hopping routes that can be accessed on biologically relevant time scales, highlighting the capability of EHPath to map charge hopping routes in proteins.

Next, we examine the hopping routes in the heme domain of $P450_{BM3}$, a bacterial class II cytochrome P450 protein that catalyzes alkane hydroxylation. Cytochrome P450 is a known producer of ROS when the Fe-coordinated O_2 remains in the oxidized state in the absence of a substrate (43, 44). In this protein, hole-hopping pathways could divert holes from compound I (species D in Fig. 1) to the protein surface, thus preventing ROS formation. In the presence of a

substrate, these hopping pathways should not enable hole escape on the timescale for the conversion of compound I to compound II (D to E, Fig. 1). *Ab initio* MD simulations have found that D is converted to E in ~ 0.1 ps (45). Thus, protective HT to the protein surface should not approach this time scale, lest it defeat normal enzyme function, but could be useful to prevent ROS formation from compound I in the absence of substrate (the R-H substrate shown in Fig. 1 is not considered explicitly in our analysis, as it is not involved in hole hopping pathways). Two hole hopping pathways to the protein surface have been proposed in $P450_{BM3}$ (PDB 2IJ2) (4) (Fig. 2). Distance-based arguments suggested that Pathway 1 would support faster hole hopping (4).

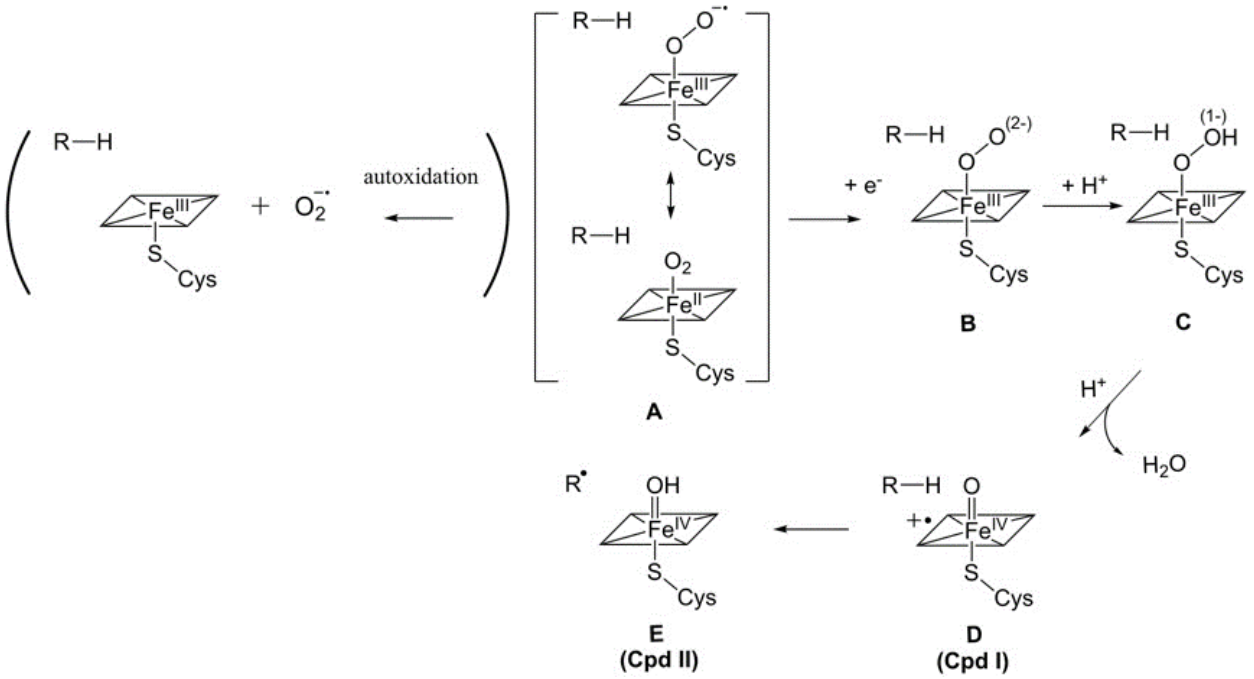


Fig. 1. Part of the catalytic cycle of cytochrome P450. The cycle begins with substrate entry to the active site (47), displacing an axial water, driving a low-to-high spin conversion at the metal center that causes the iron(III) center to dome the heme, and enabling the heme's action as an electron sink for the reductase domain (47). When the iron center is reduced from Fe(III) to Fe(II), O_2 can bind to the iron center to generate the superoxy form of the enzyme, A (47). This species may induce autoxidation, which could enable species A to decay to Fe(III) and a superoxide anion at a

rate of about 0.1 s^{-1} (46, 47) (indicated in the parentheses). Species A can avoid this fate by being converted into species B by electron transfer from the reductase domain at a rate faster than 99 s^{-1} in wild-type cytochrome $P450_{BM3}$ (46-48). This step is followed by proton transfer to B, to form species C. A second proton transfer results in an iron-oxo radical cation D (compound I) and a water molecule. Then, D abstracts a hydrogen atom from the substrate to form a Fe(IV)-hydroxide intermediate E (49). The substrate radical subsequently abstracts hydroxyl radical from compound II, regenerating the resting [Fe(III)] state of the enzyme and oxidized substrate.

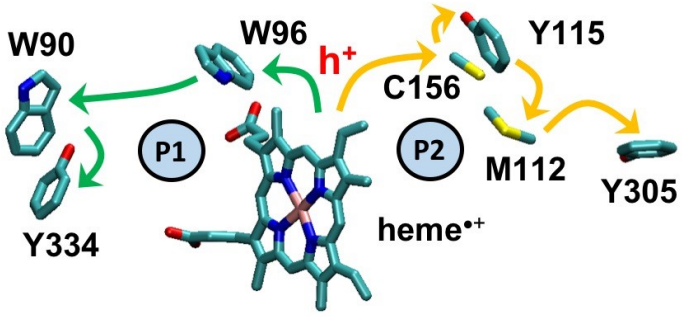


Fig. 2. Probable routes for hole hopping from the active site of $P450_{BM3}$ to the protein surface (4). Pathway 1 (P1, green arrows) is heme→W96→W90→Y334, with edge-to-edge distances (R_{ee}) between the redox-active groups of 7.32 Å, 8.35 Å, and 4.45 Å, respectively, based on PDB structure 2IJ2. R_{ee} was computed as the minimum distance between the heavy atoms for each redox-active site (that is, porphyrin ring of heme, indole group of Trp, phenol group of Tyr, methanethiol moiety of Cys, and dimethyl sulfide in Met). Pathway 2 (P2, orange arrows) is heme-→C156→Y115→M112-→Y305, with R_{ee} values of 10.03 Å, 3.99 Å, 3.70 Å, and 5.19 Å, respectively. 'h^+' denotes the hole that transfers from heme to W96 at the start of Pathway 1, or to C156 in Pathway 2, while the arrows in the figure indicate the direction of the protective hole transport.

EHPath was used to find the fastest hole hopping routes with Y334 or Y305 as the terminal hole acceptor in $P450_{BM3}$ (Fig. 2). The fastest HT route (with $\tau_M = 3.7 \times 10^{-2}$ s) is established by Pathway 1 shown in Fig. 2. This HT route does not compete with the much faster conversion of D to E on the sub-ps timescale (45) (Fig. 1). τ_M for Pathway 2 of Fig. 2 is 4.0×10^3 s (Table 1). Therefore, our results using EHPath exclude Pathway 2 (Fig. 2) as a plausible hole escape route.

Table 1 shows that the τ'_{SBT} values differ from the corresponding τ_M values by four orders of magnitude, although both are in the range that would lead to protection on the relevant biological time scale. The difference is a consequence of the large differences between the individual forward HT rates (see Table S2), resulting from the use of different free energy parameters in the two estimates of the timescale for hole escape. τ_M and τ_{SBT}, computed using rates derived from Eqs. 1 and 4 respectively, use the same $\Delta G°$ and λ_{DA} values and are therefore in much closer agreement; this reflects the similar conclusions derived from using Eq. 1 and 4. Importantly, irrespective of the specific model used for the nearest-neighbor hopping rate, we reach the same conclusion that the fastest hole escape routes are predicted to function on reasonable biological timescales (probably within the ranges shown in Table 1) and will not interfere with the catalytic function. The closer agreement between the τ_M and τ_{SBT} values (compared to τ_M and τ'_{SBT} values) results from using the same $\Delta G°$ and λ_{DA} values in the expressions for the mean residence times. Our theoretical finding that path 1 is the most probable oxidative damage escape route is consistent with the experimental finding that $W96^{\bullet+}$ is a critical intermediate in the hole hopping of Ru-modified $P450_{BM3}$ (8) (this amino acid is conserved in ~75% of the cytochrome P450 species (3)).

Table 1. Fastest 5 hole hopping routes in $P450_{BM3}$ (the heme is the hole donor, while Y305 or Y334 is the terminal hole acceptor), ranked by the mean residence time τ_M (Eq. 5, with rates derived from Eq. 1 and λ_{DA} from Eq. 2-3). τ_{SBT} was similarly calculated using Eq. 4. τ'_{SBT} also uses Eq. 4, but with the free energy parameters of ref. (3).

Hole Hopping Pathways	τ_M **(s)**	τ_{SBT} **(s)**	τ'_{SBT} **(s)**
HEM-W96-W90-Y334 (P1)	3.7×10^{-2}	9.3×10^{-3}	4.6×10^{-6}
HEM-Y115-Y305	6.1×10^{-2}	1.9×10^{-2}	6.1×10^{-6}
HEM-C156-Y115-Y305	1.4×10^{-1}	7.7×10^{-2}	-

HEM-W96-Y334	4.2×10^{-1}	5.7×10^{-2}	9.6×10^{-6}
HEM-C156-Y305	1.5×10^{1}	6.7	-
HEM-C156-Y115-M112-Y305 (P2)	4.0×10^{3}	2.2×10^{2}	-

A recent experimental study (12) offers an excellent framework in which to analyze the hole hopping routes that may prevent oxidative damage to the heme in Ccp1 of *Saccharomyces cerevisiae* (PDB 1ZBY (50)). Ref. (12) identifies protein zones with different densities of residues that can be oxidized by hole hopping from the heme. The three main zones are shown in Fig. 3: Zone 1 contains the highest percentage of oxidizable amino acid residues, followed by zones 2a and 2b.

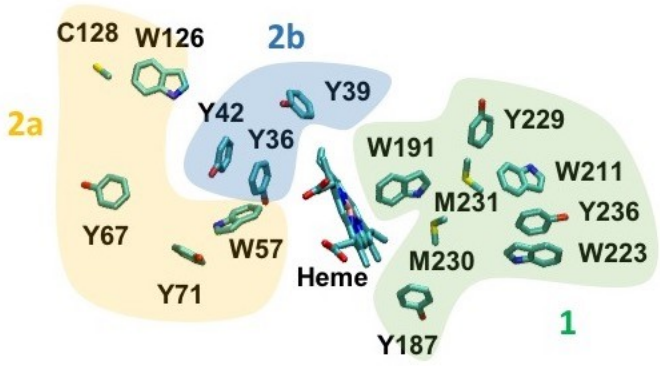


Fig. 3. Redox-active residues (C, M, W, Y) in zone 1 (green), 2a (yellow), and 2b (blue) of Ccp1 (PDB 1ZBY (50)). The zones lowest in oxidizable residues (zones 3 and 4 of ref. (12)) are not shown.

Table 2. Mean residence time τ_M of the hole for the 5 fastest hole hopping pathways in Ccp1 with selected terminal HT sites identified in ref. (12) (calculated using Eq. 1 and 5 for the HT rates and Eqs. 2-3 for the free energy parameters). Additional pathways are shown in Table S3.

Hole Hopping Pathways	τ_M **(s)**
HEM-W191-Y229	2.5×10^{-3}

HEM-W191-W211	2.5×10^{-3}
HEM-Y36	1.2×10^{-2}
HEM-Y187-W191-Y229	6.1×10^{-2}
HEM-Y187-W191-W211	6.1×10^{-2}

We find that 17 of the 20 fastest hole hopping pathways terminate in zone 1 (Table S3), providing theoretical support for the experimental data interpretation of ref. (12) (8 residues of the 24 oxidized residues in Ccp1 identified from LC-MS/MS analysis were mapped into zone 1, while the 16 other residues were distributed across zones 2a, 2b, 3, and 4). We also identify W191 as a key residue in four of the top five hole hopping routes. Even using the large reorganization energies when estimating τ_M, we find that a hole can transfer from the heme to Y229, through W191, on a millisecond timescale (Table 2). This result is consistent with the experimental observation of a role for $W191^{\bullet+}$ in mediating CT between Ccp1 and cytochrome c (51), and the hypothesis that the charge would most likely migrate from $W191^{\bullet+}$ to a tyrosine residue (which is identified to be Y229 in our analysis) (52). Interestingly, comparisons of the hole hopping pathways listed in Tables 1 and 2 show that the fastest pathways in Ccp1 have smaller τ_M values than the fastest pathways in $P450_{BM3}$. This difference may correlate with the differential functions of the two proteins. The main roles for Ccp1 are the oxidization of two ferrocytochrome c molecules through the initial two-electron reduction of H_2O_2 (53), and the eventual transfer of its heme to catalase during respiration when large mitochondrial levels of H_2O_2 are present (11). In the absence of ferrocytochrome *c* molecules, the Ccp1 pathways listed in Table 2 can help protect the active site from irreversible oxidative damage (and heme crosslinking to W51) (54) produced by highly oxidizing H_2O_2; the protection of the active site integrity allows the subsequent transfer of undamaged heme to catalase A (12). The $P450_{BM3}$ monooxygenase relies, instead, on the less

strongly oxidizing O_2 as the primary route to the hydroxylation of fatty acid substrates (55). Therefore, the reliance of $P450_{BM3}$ on a weaker oxidant may require fewer and less rapid hole hopping routes for sufficient heme active site protection from oxidative damage in $P450_{BM3}$ compared to the case with Ccp1.

τ_M (Eq. 5) and $\tau_{M,approx}$ (Eq. 6) values for the top five hole hopping pathways between Gly829$^\bullet$ and $[4Fe4S]^{2+}$ in BSS appear in Table 3. These pathways involve the key residues highlighted in Fig. 4. Since the backward HT rates are negligible in these hopping pathways, Eq. 6 provides an excellent approximation to the mean residence time (Table 3). The pathways from G829$^\bullet$ to $[4Fe4S]^{2+}$ via Y533, M713, and W58 are predicted to be traversed on a millisecond timescale (see Table 3) despite their lengths (15) as a consequence of their downhill free energy landscapes. Such landscapes dictate the unidirectional character of the hole transport to the iron-sulfur cluster, which is quantified by the very similar τ_M and $\tau_{M,approx}$ values in Table 3. Since our analysis demonstrates that $[4Fe4S]^{2+}$ may reduce G829$^\bullet$ (note that the G829$^\bullet$ reduction potential used in our analysis accounts for the coupled proton transfer (56)), the iron-sulfur cluster could act as a reducing agent in other systems (e.g., in the experimentally observed reversible reduction of the Gly radical in pyruvate formate lyase, another glycyl radical enzyme (15, 57)). The BSSβ small subunit of BSS can coordinate a [4Fe4S] cluster by adopting a local folding similar to that of high-potential [4Fe4S] proteins (15). However, in contrast to the latter, BSSβ does not provide the substantial shielding from the solvent that stabilizes $[4Fe4S]^{3+}$ (15). This fact, and the thermodynamic driving force that prevents the electron hole from returning to the protein, might cause the iron-sulfur cluster to serve as a transient hole trap prior to hole delivery outside of a protein.

In a previous study, we showed that through-protein hole transfer from an oxidized system (a nucleic acid double strand) to high-potential [4Fe4S] clusters may generally occur on a millisecond

to microsecond timescale (17). This theoretical result agrees with previous experimental findings on the viability of redox signaling involving [4Fe4S] proteins relevant to DNA replication and repair (18, 58). The present analysis may indicate an (additional) interpretation of the CT producing the signaling in terms of hole trapping in iron-sulfur clusters.

Table 3. Mean residence times τ_M (Eq. 5) and $\tau_{M,approx}$ (Eq. 6) of the hole in the 5 fastest hole-hopping escape pathways of BSS, where Gly829$^\bullet$ is the initial hole donor and $[4Fe4S]^{2+}$ is the final hole acceptor. Pathways 1 and 2 are delineated in Fig. 4.

Hole Hopping Pathways	τ_M **(s)**	$\tau_{M,approx}$ **(s)**
G829-M713-Y533-[4Fe4S] (1)	4.5×10^{-3}	4.5×10^{-3}
G829-M713-Y533-W58-[4Fe4S] (2)	4.5×10^{-3}	4.5×10^{-3}
G829-M713-W55-[4Fe4S]	7.8×10^{-3}	7.8×10^{-3}
G829-M713-W55-Y533-[4Fe4S]	8.0×10^{-3}	8.0×10^{-3}
G829-M713-W55-W531-[4Fe4S]	1.0×10^{-2}	9.1×10^{-3}

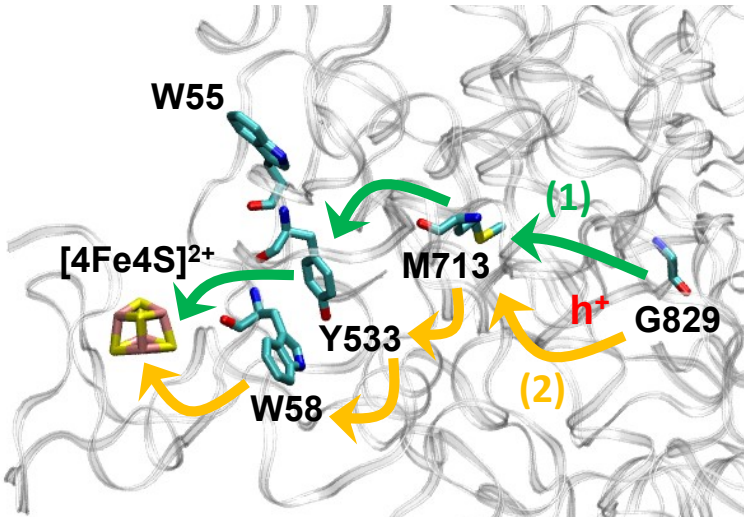


Fig. 4. Locations of the BSS redox-active residues involved in the five most rapid hole hopping routes from G829$^\bullet$ to the terminal hole acceptor $[4Fe4S]^{2+}$ (Table 3). The two fastest paths (1) and

(2) are highlighted with green and orange arrows, respectively.

CONCLUDING REMARKS

The hole transfer kinetic modeling and analysis incorporated in the EHPath program allow us to identify, assess, and understand the role of hole hopping pathways as oxidative damage escape routes in three important biological systems, taking into account microscopic models for the HT rates among hopping stations. The quantitative results of this theoretical study support the general viability of the proposed hole-transfer protective mechanism in redox proteins (3, 4). Our analysis identifies the predicted dominant hole escape route in $P450_{BM3}$. We find that the mean residence time computed for the charge on the fastest pathway is sufficiently rapid to be biologically relevant (and thus to help preventing ROS formation in the absence of substrate) and is longer than the timescale of the functional transition from the iron-oxo radical cation D to the Fe(IV)-hydroxide complex E in the presence of the substrate. The identification of W191 as a key site for hole hopping away from the active site in Ccp1 and the termination of the most rapid hopping are both consistent with the EPR spectral signature of $W191^{\bullet+}$ (12, 51). A [4Fe4S] cluster can serve as the terminal hole acceptor in BSS, reducing the Gly radical through hopping pathways that involve the M713, Y533 and W58 residues. The software and modeling described here provide a robust approach to map and to characterize CT hopping pathways.

Mutational studies of $P450_{BM3}$, Ccp1, and BSS, combined with laser spectroscopy, would help to evaluate the hole escape routes explored here. For example, the Y533 residue in BSS, which is shared between the two fastest hole hopping pathways (Table 3), may be mutated to a structurally similar Phe residue in order to examine the robustness and detailed pathway of the oxidation protection mechanism in BSS. As well, W90, which appears in pathway P1 of $P450_{BM3}$, may be mutated to His in order to explore the importance of the W90 residue for mediating hole escape.

Our approach should enable further studies of hopping pathways in proteins, DNA, and protein-DNA complexes, including the primase-polymerase complex (17, 18). Future applications of this approach to oxygen-utilizing/evolving proteins, including ribonucleotide reductase, cytochrome *c* oxidase, and photosystem II (3), may help to reveal how nature has evolved charge hopping protection mechanisms.

Supporting Information. The Supporting Information includes Tables S1-S3 and Fig. S1-S2, as well as details about the EHPath code and its use. The EHPath.py code can be accessed and downloaded at https://github.com/etransfer/EHPath.

Corresponding Authors

*Email: david.beratan@duke.edu (D.N.B), agostino.migliore@duke.edu (A.M.)

Notes

The authors declare no competing financial interests.

Acknowledgement

We thank the National Institutes of Health (Grant GM-48043) and the Blue Waters sustained-petascale computing project (R. D. T) – funded by the National Science Foundation (awards OCI-0725070 and ACI-1238993) and the State of Illinois, and the National Science Foundation (CHE-1709497) for research support.

References and Notes

1. Davies MJ (2005) The oxidative environment and protein damage. *BBA-Proteins Proteomics* 1703(2):93-109.
2. Berlett BS & Stadtman ER (1997) Protein oxidation in aging, disease, and oxidative stress. *J. Biol. Chem.* 272(33):20313-20316.
3. Gray HB & Winkler JR (2015) Hole hopping through tyrosine/tryptophan chains protects proteins from oxidative damage. *Proc. Natl. Acad. Sci. U. S. A.* 112(35):10920-10925.

4. Winkler JR & Gray HB (2015) Electron flow through biological molecules: does hole hopping protect proteins from oxidative damage? *Q. Rev. Biophys.* 48(4):411-420.
5. Kanwar R & Balasubramanian D (1999) Structure and stability of the dityrosine-linked dimer of γB-crystallin. *Experimental Eye Research* 68(6):773-784.
6. Hawkins CL & Davies MJ (2005) The role of aromatic amino acid oxidation, protein unfolding, and aggregation in the hypobromous acid-induced inactivation of trypsin inhibitor and lysozyme. *Chem. Res. Toxicol.* 18(11):1669-1677.
7. Stadtman ER & Levine RL (2003) Free radical-mediated oxidation of free amino acids and amino acid residues in proteins. *Amino Acids* 25(3-4):207-218.
8. Ener ME, Gray HB, & Winkler JR (2017) Hole hopping through tryptophan in cytochrome p450. *Biochemistry* 56(28):3531-3538.
9. Choi M, Shin S, & Davidson VL (2012) Characterization of electron tunneling and hole hopping reactions between different forms of MauG and methylamine dehydrogenase within a natural protein complex. *Biochemistry* 51(35):6942-6949.
10. Polizzi NF, Migliore A, Therien MJ, & Beratan DN (2015) Defusing redox bombs? *Proc. Natl. Acad. Sci. U. S. A.* 112:10821-10822.
11. Kathiresan M, Martins D, & English AM (2014) Respiration triggers heme transfer from cytochrome c peroxidase to catalase in yeast mitochondria. *Proc. Natl. Acad. Sci. U. S. A.* 111(49):17468-17473.
12. Kathiresan M & English AM (2017) LC-MS/MS suggests that hole hopping in cytochrome c peroxidase protects its heme from oxidative modification by excess H_2O_2. *Chem. Sci.* 8(2):1152-1162.
13. Martins BM*, et al.* (2011) Structural basis for a Kolbe-type decarboxylation catalyzed by a glycyl radical enzyme. *J. Am. Chem. Soc.* 133(37):14666-14674.
14. Knappe J, Neugebauer FA, Blaschkowski HP, & Gänzler M (1984) Post-translational activation introduces a free radical into pyruvate formate-lyase. *Proc. Natl. Acad. Sci. U.S.A* 81(5):1332-1335.
15. Funk MA, Judd ET, Marsh ENG, Elliott SJ, & Drennan CL (2014) Structures of benzylsuccinate synthase elucidate roles of accessory subunits in glycyl radical enzyme activation and activity. *Proc. Natl. Acad. Sci. U.S.A* 111(28):10161-10166.
16. Gray HB & Winkler JR (2005) Long-range electron transfer. *Proc. Natl. Acad. Sci. U.S.A* 102(10):3534-3539.
17. Teo RD*, et al.* (2019) Charge transfer between [4Fe4S] proteins and DNA is unidirectional. Implications for biomolecular signaling. *Chem* 5:122-137.
18. O'Brien E*, et al.* (2017) The [4Fe4S] cluster of human DNA primase functions as a redox switch using DNA charge transport. *Science* 355:eaag1789.
19. Marcus RA & Sutin N (1985) Electron transfers in chemistry and biology. *Biochim. Biophys. Acta* 811(3):265-322.
20. Marcus RA (1963) On the theory of oxidation-reduction reactions involving electron transfer. *J. Chem. Phys.* 67(4):853-857.
21. Marcus RA (1968) Theoretical relations among rate constants, barriers, and Broensted slopes of chemical reactions. *J. Chem. Phys.* 72(3):891-899.
22. Heck A*, et al.* (2012) Charge transfer in model peptides: obtaining Marcus parameters from molecular simulation. *J. Phys. Chem. B* 116(7):2284-2293.
23. Krishtalik LI, Kuznetsov AM, & Mertz EL (1997) Electrostatics of proteins: description in terms of two dielectric constants simultaneously. *Proteins* 28(2):174-182.

24. Pethig R (1979) *Dielectric and electronic behavior of biological materials* (John Wiley & Sons, New York).
25. Teo RD, Smithwick ER, Migliore A, & Beratan DN (2019) A single AT-GC exchange can modulate charge transfer-induced p53-DNA dissociation. *Chem Commun* 55(2):206-209.
26. Hopfield JJ (1974) Electron transfer between biological molecules by thermally activated tunneling. *Proc. Natl. Acad. Sci. U. S. A.* 71(9):3640-3644.
27. Davidson VL (2008) Protein control of true, gated, and coupled electron transfer reactions. *Accounts of Chemical Research* 41(6):730-738.
28. Goodpaster JD (2019) Theoretical Chemistry Answers Bimolecular Signaling Debate in [4Fe4S] Proteins. *Chem* 5(1):12-14.
29. Langen R*, et al.* (1995) Electron tunneling in proteins: coupling through a beta strand. *Science* 268(5218):1733-1735.
30. Winkler JR & Gray HB (2014) Electron flow through metalloproteins. *Chem. Rev.* 114(7):3369-3380.
31. Regan JJ*, et al.* (1995) Electron tunneling in azurin: the coupling across a β-sheet. *Chem. Biol.* 2(7):489-496.
32. Zhao Y-Y, Ma J-Y, Zhao X-J, & Li X-Y (2008) Solvent reorganization energy of intramolecular electron transfer in peptides involving tryptophan and tyrosine. *Chin. J. Chem.* 26(11):2003-2008.
33. Petersen RA & Evans DH (1987) Heterogeneous electron-transfer kinetics for a variety of organic electrode-reactions at the mercury acetonitrile interface using either tetraethylammonium perchlorate or tetraheptylammonium perchlorate electrolyte. *J Electroanal Chem* 222(1-2):129-150.
34. Migliore A & Nitzan A (2011) Nonlinear charge transport in redox molecular junctions: a Marcus perspective. *Acs Nano* 5(8):6669-6685.
35. Kubar T & Elstner M (2009) Solvent reorganization energy of hole transfer in DNA. *J Phys Chem B* 113(16):5653-5656.
36. Khan A (2013) Reorganization, activation and ionization energies for hole transfer reactions through inosine-cytosine, 2-aminopurine-thymine, adenine-thymine, and guanine-cytosine base pairs: a computational study. *Comput Theor Chem* 1013:136-139.
37. Leake MC (2013) *Single-molecule cellular biophysics* (Cambridge University Press).
38. Balabin IA, Hu X, & Beratan DN (2012) Exploring biological electron transfer pathway dynamics with the pathways plugin for VMD. *J. Comput. Chem.* 33(8):906-910.
39. Beratan DN, Betts JN, & Onuchic JN (1991) Protein electron transfer rates set by the bridging secondary and tertiary structure. *Science* 252(5010):1285-1288.
40. Betts JN, Beratan DN, & Onuchic JN (1992) Mapping electron-tunneling pathways - an algorithm that finds the minimum length maximum coupling pathway between electron-donors and acceptors in proteins. *J Am Chem Soc* 114(11):4043-4046.
41. Guallar V & Wallrapp F (2008) Mapping protein electron transfer pathways with QM/MM methods. *J. Royal Soc. Interface* 5(Suppl 3):233-239.
42. Hagberg A, Swart P, & S Chult D (2008) *Exploring network structure, dynamics, and function using networkx* (; Los Alamos National Lab. (LANL), Los Alamos, NM (United States)) p Medium: ED.
43. Gottlieb RA (2003) Cytochrome p450: major player in reperfusion injury. *Arch. Biochem. Biophys.* 420(2):262-267.
44. Jezek P & Hlavata L (2005) Mitochondria in homeostasis of reactive oxygen species in

cell, tissues, and organism. *Int. J. Biochem. Cell Biol.* 37(12):2478-2503.
45. Elenewski JE & Hackett JC (2015) Ab initio dynamics of the cytochrome p450 hydroxylation reaction. *J. Chem. Phys.* 142(6):9.
46. Ost TWB*, et al.* (2003) Oxygen activation and electron transfer in flavocytochrome p450 bm3. *J. Am. Chem. Soc.* 125(49):15010-15020.
47. Meunier B, de Visser SP, & Shaik S (2004) Mechanism of oxidation reactions catalyzed by cytochrome p450 enzymes. *Chem. Rev.* 104(9):3947-3980.
48. The ligand-reduction step from A to B in Fig. 1 corresponds to the step from D to E in Fig. 7 of ref. 47. The rate-determining step in the p450 catalytic cycle corresponds to the heme reduction step (that is, the step from B to C in Fig. 7 of ref. 47), which has a rate of 99 s^{-1} in the wild-type protein (see ref. 46). Therefore, the rate of the A-to-B step in Fig. 1 is larger than 99 s^{-1}.
49. Huang X & Groves JT (2018) Oxygen activation and radical transformations in heme proteins and metalloporphyrins. *Chemical Reviews* 118(5):2491-2553.
50. Bonagura CA*, et al.* (2003) High-resolution crystal structures and spectroscopy of native and compound I cytochrome c peroxidase. *Biochemistry* 42(19):5600-5608.
51. Payne TM, Yee EF, Dzikovski B, & Crane BR (2016) Constraints on the radical cation center of cytochrome c peroxidase for electron transfer from cytochrome c. *Biochemistry* 55(34):4807-4822.
52. Barrows TP, Bhaskar B, & Poulos TL (2004) Electrostatic control of the tryptophan radical in cytochrome c peroxidase. *Biochemistry* 43(27):8826-8834.
53. Volkov AN, Nicholls P, & Worrall JAR (2011) The complex of cytochrome c and cytochrome c peroxidase: the end of the road? *BBA-Bioenergetics* 1807(11):1482-1503.
54. Murphy EJ, Metcalfe CL, Basran J, Moody PCE, & Raven EL (2008) Engineering the substrate specificity and reactivity of a heme protein: creation of an ascorbate binding site in cytochrome c peroxidase. *Biochemistry* 47(52):13933-13941.
55. Lee D-S*, et al.* (2003) Substrate recognition and molecular mechanism of fatty acid hydroxylation by cytochrome p450 from Bacillus subtilis: crystallographic, spectroscopic, and mutational studies. *J. Biol. Chem.* 278(11):9761-9767.
56. Armstrong DA, Rank A, & Yu D (1995) Solution Thermochemistry of the Radicals of Glycine. *J Chem Soc Perk T 2* (3):553-560.
57. Nnyepi MR, Peng Y, & Broderick JB (2007) Inactivation of E. coli pyruvate formate-lyase: role of AdhE and small molecules. *Arch. Biochem. Biophys.* 459(1):1-9.
58. Arnold Anna R, Grodick Michael A, & Barton Jacqueline K (2016) DNA charge transport: from chemical principles to the cell. *Cell Chemical Biology* 23(1):183-197.