Introduction

Many vital cellular processes such as photosynthesis, respiration, and a variety of metabolic reactions rely on long-range electron transfer (ET) between macromolecules. In a biological context, the intermolecular ET involves formation of a redox complex, electron tunneling across a biomolecular interface, and dissociation of the reduced and oxidized products. While the electron tunneling is, arguably, the best described step in this sequence of events [1,2], our understanding of the first and the last steps is often obscured by the macromolecular dynamics and the uncertainty of the binding geometries in reactive complexes [3]. Experimental study of the biomolecular ET in solution is further complicated by two factors. First, in order to maximize the overall turnover in the ET chain, redox complexes have short lifetimes and, consequently, low binding affinities, making structural characterization of these weak, transient, highly dynamic complexes difficult [4]. Second, unlike in other biomolecular interactions where formation of a functional complex requires precise alignment of the partner molecules, an efficient intermolecular ET can occur from more than one protein-protein orientation. This possible multitude of active binding geometries in redox complexes, and not the electron tunneling event itself, is the rate-limiting step in the intermolecular ET. The resulting picture of the ET interactome reveals that most ET-competent binding geometries are located in electrostatically favorable regions, 2) addition, electrostatic interactions can guide the partner proteins to stabilize transient protein-protein complexes shows little or no effect on the binding affinity and the ET rate [14–16]. Thus, contrary to the conclusions of early modeling studies [5,6], it appears that ET complexes are not optimized for intermolecular electrostatic interactions, which is thought to facilitate rapid dissociation required for a high turnover [13]. The current view is that charged residues – often located at the periphery of the binding sites – play a critical role in the early stages of protein-protein interactions by steering binding partners to form a loose encounter complex, but are less important for the ensuing formation of the specific, ET-active form, dominated by short-range forces [4,13,17,18]. In addition, electrostatic interactions can guide the partner proteins to productive docking geometries, without contributing to the overall binding affinity [19].

The discrepancies between the experimental results and the outcome of the electrostatics-based simulations highlight the need for alternative computational approaches. In this work, we used a
Author Summary

A number of vital cellular processes such as respiration, photosynthesis, and multifarious metabolic conversions rely on a long-range electron transfer (ET) among protein molecules. Full understanding of the biological ET requires accurate description of the redox protein complexes, which is hampered by their pronounced mobility and short lifetimes. Here we used a simple computational approach to predict the ET properties of the physiological protein complexes of cytochrome c (Cc)—a small electron carrier that coordinates the electron flow among different redox partners. By performing extensive conformational sampling of the possible binding geometries, we mapped out functional epitopes in the Cc complexes and assessed their ET properties. Our study suggests that protein dynamics within redox complexes is the rate-limiting step in the intermolecular ET and indicates that the functional epitope size correlates with the extent of dynamics in the Cc complexes. We believe that the latter finding can be used as a diagnostic tool for protein mobility and expect that this work will engender future studies of the intermolecular ET in biological networks.

Methods

Molecular Modeling

Protein coordinates were taken from the PDB; cofactor topologies and parameters were obtained from the HIC-Up database [27]. The ScErv1 structure was modeled with Modeller [28] based on the X-ray coordinates of the homologous protein AtEnv1 [29]. Missing protein loops were built with Modeller [28]. Hydrogen atoms were added in Xplor-NIH using standard topology and parameter sets [30]. The ensembles of Fcb2 and SOX domain-domain orientations were generated in Xplor-NIH by simulated annealing in torsion angle space with standard protein geometry and vdW parameters and knowledge-based dihedral angle potentials [31]. Two domains were treated as rigid bodies, while the intervening linker segment (residues 99–100 or 89–99 for Fcb2, and residues 83–94 for SOX) was allowed full flexibility. Detailed description of the systems studied is provided in the Table S1.

Conformational Search

Conformational sampling was performed in Xplor-NIH [30] using a modified version of the published protocol [32]. The molecular system was oriented such that the centers of mass (CM) of the two interacting proteins appeared at the origin of the coordinate system and on the positive z axis. In a typical run, the Cc binding partner was kept stationary, while Cc was systematically rotated around x and z axes, corresponding to the θ and φ rotations in the spherical coordinate space (Figure 1A). The desired spatial resolution, δd—defined here as the separation between neighboring CMs of the rotated protein and typically set to 1 Å—determines the δθ and δφ rotation increments and the total number of steps required to sample the entire conformational space (Equations 1–3).

\[
\delta \theta = a \cos \left[ 1 - 0.5 \left( \frac{\delta d}{\delta} \right)^2 \right],
\]

where δd and d are the desired spatial resolution and the distance between CMs of the two proteins in the starting structure. Number of δθ rotational increments is given by \( n_\theta = \pi/\delta \theta \), with the value rounded to the last digit before the decimal so that \( n_\theta \notin \mathbb{I} \). For θ rotation around the x axis, the rotation matrix is given by:

\[
R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{bmatrix}
\]

Upon θ rotation, the ordinate of the Cc CM becomes \( y' = d \cdot \sin \theta \). By symmetry with the Equation 1, the δφ rotational increment at each θj is given by:

\[
\delta \phi_j = a \cos \left[ 1 - 0.5 \left( \frac{\delta d}{\delta} \right)^2 \right] = a \cos \left[ 1 - 0.5 \left( \frac{\delta d}{d \cdot \sin \theta_j} \right)^2 \right]
\]

At each \( \theta_j \), the number of δφ rotational increments is given by \( n_\phi = 2\pi/\delta \phi \), with the value rounded to the last digit before the decimal so that \( n_\phi \notin \mathbb{I} \). For φ rotation around the z axis, the rotation matrix is given by:

\[
R_z(\phi) = \begin{bmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{bmatrix}
\]

The total number of the CMs defining the coverage of the conformational space is given by

\[
N = \sum_{i=1}^{n_\theta} \sum_{j=1}^{n_\phi} n_{i,j}
\]

To sample flat, irregular surfaces of the selected CCO and Cbc1 regions, the protocol was modified to include the initial translation of Cc along the long axis of the molecular system, followed by rotation around the same axis, both in increments required to produce 1 Å separation between neighboring Cc CMs.

To simulate the rotational freedom, Cc was rotated around orthogonal \( x, \psi, \zeta \) axes originating at its CM (Figure 1A). The rotational axes were defined relative to the plane determined by the two proteins’ CMs and the Fe atom of Cc with the corresponding Cartesian coordinates \((x_{0},y_{0},z_{0})\), \((x_{1},y_{1},z_{1})\), and \((x_{2},y_{2},z_{2})\) (Equation 6):

\[
Ax + By + Cz + D = 0
\]

\[
\begin{bmatrix} x-x_{0} & y-y_{0} & z-z_{0} \\ x_{1}-x_{0} & y_{1}-y_{0} & z_{1}-z_{0} \\ x_{2}-x_{0} & y_{2}-y_{0} & z_{2}-z_{0} \end{bmatrix} = 0
\]
During subsequent refinement, the frontal orientation of the Cc heme was ensured by the separation between the ET cofactors (black circles in Figure 1B). In the starting structure (black circles) the Cc in the starting protein-protein orientation was rotated around its CM in the starting structure (black circles) is rotated to obtain the shortest separation between the ET cofactors (arrows). This ensures the frontal orientation of the Cc heme during subsequent refinement.

Proteins’ CMs are shown as grey spheres, one of which is at the origin of the coordinate system. The moving protein (typically Cc) possesses three translational (u, v, r) and three rotational (y, z, ζ) degrees of freedom. For the definition of the rotational axes see text. (B, C) Conformational search in the reduced rotational space (B) Cc in the starting structure (black circles) is rotated to obtain the shortest separation between the ET cofactors (arrows). This ensures the frontal orientation of the Cc heme during subsequent refinement.

The reduced rotational space $-45^\circ \leq \chi, \psi \leq 45^\circ$ traced by a point on the surface of a sphere (bold shape). Location of the other conformational sampling points in the reduced rotational space (grey circle). (C) The reduced rotational space $-45^\circ \leq \chi, \psi \leq 45^\circ$ is indicated. (B, C) Conformational search was performed with composite rotation potential and side-chain optimization for the entire complex at the value between 5.5 and 8.5 Å. The vdW potential was set to zero for all side-chain atoms extending beyond Cβ, and Cc translated along the vector joining two proteins' CMs until the vdW energy reached the value between zero and a chosen cut-off, thus producing a rigid-body mimic of the protein complex. The separation between the proteins’ CMs determines the other translational coordinate, r. In control runs, side-chain optimization with full vdW potential was performed after each translational step (Figure 2).

Figure 1. Illustration of the conformational search used in this work. (A) Definition of the search space in the spherical coordinates. Proteins’ CMs are shown as grey spheres, one of which is at the origin of the coordinate system. The moving protein (typically Cc) possesses three translational (0, 0, 0) and three rotational (y, z, ζ) degrees of freedom. For the definition of the rotational axes see text. (B, C) Conformational search in the reduced rotational space (B) Cc in the starting structure (black circles) is rotated to obtain the shortest separation between the ET cofactors (arrows). This ensures the frontal orientation of the Cc heme during subsequent refinement.

Solving the determinant and noticing that one of the CMs is at the origin of the coordinate system, i.e. (u₀, v₀, r₀) = (0, 0, 0), gives $A = -v₀z₀ - h₃$, $B = v₀z₀ - h₃$, $C = v₀y₀ - h₄$, and $D = 0$. The $\chi$ and $\zeta$ axes are defined along the (A, B, C) normal to the plane and the (y₀, z₀, z₁) vector joining the two CMs, respectively, while the y axis is perpendicular to the latter two and is given by the vector (A’, B’, C’)= (C₀ - B₀, A₀ - C₀, B₀ - A₀).

Two rotational sampling protocols were employed in this work. In the first protocol, the rotational coordinates were systematically varied in $\Delta \chi = \delta \chi = \delta \zeta = 5^\circ$ increments in the full rotational space ($0^\circ \leq \chi, \psi < 360^\circ$, $0^\circ \leq \zeta < 180^\circ$). In the second protocol, the conformational search was performed with composite-rotation increments ($\Delta \chi = \delta \chi = \delta \zeta = 15^\circ$) in the reduced rotational space $(-45^\circ \leq \chi, \psi < 45^\circ$, $0^\circ \leq \zeta < 360^\circ$) (Figure 1B,C). In the latter case, the Cc in the starting protein-protein orientation was rotated around its CM in $\Delta \chi = \delta \chi = \delta \zeta = 5^\circ$ increments ($0^\circ \leq \chi, \psi < 360^\circ$, $0^\circ \leq \zeta < 180^\circ$) in the search for the position with the smallest separation between the ET cofactors (black circles in Figure 1B).

In this way, the frontal orientation of the Cc heme was ensured during subsequent refinement, and the ET cofactors (grey circle in Figure 1B), enabling sampling of short intermolecular ET distances during the reduced rotational search (Figure 1C).

For each rotamer, the intermolecular vdW energy term $f_{vdW}$ was calculated with a repulsive quartic potential (Equation 7) [33,34]:

$$f_{vdW} = \sum_{ij} \left\{ \begin{array}{ll} 0, & r_{ij} \geq r_{min} \\ k_{vdW} (s^2 - r_{min}^2)^2, & r_{ij} < r_{min} \end{array} \right.$$  (7)

where $r_{ij}$ is the interatomic distance, $r_{min}$ adopts the value of the standard vdW radii as represented by the Lennard-Jones potential [33], $s$ is the scaling factor (the value of 0.75, frequently used in Xplor rigid-body docking [33,34], was employed throughout), and $k_{vdW}$ is a force-constant (20 kcal/mol Å$^{-4}$). The sum is carried over all pairwise intermolecular interactions between the atoms that are below a specified non-bonded cut-off, typically set to a value of 5.5 and 8.5 Å. The vdW potential was set to zero for all side-chain atoms extending beyond Cβ, and Cc translated along the vector joining two proteins’ CMs until the vdW energy reached the value between zero and a chosen cut-off, thus producing a rigid-body mimic of the protein complex.

ET Rate Calculations

The ET rate constants were calculated from Equation 8 [2]:

$$k_{ET} = k_0 \exp \left[ -\beta (r - r_0) \right] \exp \left[ -\frac{(\Delta G^o + \lambda)^2}{4\lambda k_B T} \right].$$  (8)

where $k_0 = 10^{13}$ s$^{-1}$ is the nuclear frequency, $\beta = 1.4$ Å$^{-1}$ is the decay coefficient of the electronic coupling, $r$ is the edge-to-edge distance between the redox centers, $r_0 = 3.6$ Å is the vdW contact distance, $\Delta G^o$ is the free energy difference between reactant and product states, $\lambda$ is the reorganization energy, $k_B$ is the Boltzmann constant, and $T$ is the temperature. The values of $\Delta G^o$ and $\lambda$ were taken from the literature (see Text S2).

Results/Discussion

Conformational Sampling

To define the conformational space available to the interacting molecules and determine the shortest distance between redox cofactors at each sampled point, Cc was used as a probe to explore the surface of a partner protein in search of possible binding geometries. Proteins were treated as rigid bodies with van der Waals (vdW) potential set to zero for all atoms extending beyond Cβ to allow partial side-chain overlap, mimicking the uncertainty of the side-chain positions due to local binding-induced reorganization. Steric properties of the generated complexes were assessed by calculating the intermolecular vdW energy term.

To determine the optimal parameter set for the conformational search protocol, we performed a number of control runs (Figure 2). In the first control series (where only translational freedom of Cc was simulated to accelerate the calculations), we examined the choice of the vdW cut-offs (Figure 2A,B). The value of 10 kcal/mol, found to yield realistic mimics of protein complexes in our earlier work [32], and a more permissive 300 kcal/mol cut-off were assessed. First, we performed a run with the full vdW potential and side-chain optimization for the entire complex at
each translation step. The ET distance distribution in this run (black profile in Figure 2A) was used as a reference for evaluating the protocols without side-chain optimization and the full vdW potential (vdW cut-off values of 300 kcal/mol and 10 kcal/mol; red and orange, respectively) and reduced vdW term, set to zero for all atoms extending beyond Cb (vdW cut-off values 300 kcal/mol and 10 kcal/mol; blue and green, respectively). In these runs, only translational degrees of freedom of Cc were simulated. (B) Correlation plots of heme-heme distance occurrences in the side-chain optimization run (horizontal axis) and the other control runs (vertical axis), color-coded as in (A). Pearson correlation coefficients are indicated in the plots. (C) Distribution of the heme-heme distances in the runs with the reduced vdW term (vdW cut-off 10 kcal/mol, green dataset in A) and a given rotational increment (δχ = δψ = 5°, 0° ≤ ψ ≤ 360°, 0° ≤ ζ ≤ 180°). For the composite run, δχ = δψ = 5°, 25° ≤ χ ≤ 45°, 0° ≤ ζ ≤ 360°. (D) Number of rotamers for each Cc CM, the total number of structures sampled per run, and the total run time (for 1 CPU).
doi:10.1371/journal.pcbi.1002807.g002

Figure 2. Conformational sampling in the Cc-Cb5 complex. (A) Distribution of the heme-heme distances in the simulation runs with the full vdW potential and side-chain optimization (black); full vdW potential without side-chain optimization (vdW cut-off values of 300 kcal/mol and 10 kcal/mol; red and orange, respectively); and reduced vdW term, set to zero for all atoms extending beyond Cb (vdW cut-off values 300 kcal/mol and 10 kcal/mol; blue and green, respectively). In these runs, only translational degrees of freedom of Cc were simulated. (B) Correlation plots of heme-heme distance occurrences in the side-chain optimization run (horizontal axis) and the other control runs (vertical axis), color-coded as in (A). Pearson correlation coefficients are indicated in the plots. (C) Distribution of the heme-heme distances in the runs with the reduced vdW term (vdW cut-off 10 kcal/mol, green dataset in A) and a given rotational increment (δχ = δψ = 5°, 0° ≤ ψ ≤ 360°, 0° ≤ ζ ≤ 180°). For the composite run, δχ = δψ = 5°, 5° ≤ χ ≤ 360°, 0° ≤ ζ ≤ 360°. (D) Number of rotamers for each Cc CM, the total number of structures sampled per run, and the total run time (for 1 CPU).

Having established the optimal vdW parameters, we proceeded to evaluate the choice of rotational increments to be used in the conformational search protocol. As can be seen in Figure 2C, including Cc rotations around orthogonal axes centered on its CM allows sampling shorter ET distances, with progressively shorter distances accessed at decreasing rotational increments. This is accompanied by a steady increase in the number of rotamers, reaching ca. 187,000 per (θ, ϕ) position for δχ = δψ = δζ = 5° (Figure 2D). Decreasing the rotational increments further quickly leads to an explosion of structures (e.g. 1.8·10^11 structures per run with δχ = δψ = δζ = 1°), rendering the computations prohibitively time-consuming. Besides, the number of shorter distances accessed is expected to reach a plateau at smaller increments, making the choice of δχ = δψ = δζ = 5° a good compromise between the density of the rotational space sampling and the computational time invested.

Finally, we noticed that in protein-protein complexes with short ET distances, Cc is oriented with its heme group facing the binding partner (Figure 1B). In addition, control runs with the stationary Cc and Cb5 or CcP as the sampling probes showed that the Cc functional epitope maps onto a well-defined patch of the protein surface surrounding the exposed heme edge (Figures S6 and S7). Based on these observations, we devised an accelerated
sampling scheme in a reduced rotational space, consisting of defining the frontal Cc orientation in the starting structure, usual sampling of $\theta$ and $\varphi$ dimensions, and the rotational search in a reduced $\chi$, $\psi$, $\xi$ space. As can be seen in Figure 2C (compare red and black traces), the ET distance distributions obtained with the full sampling scheme ($0 \leq \chi, \psi < 360^\circ$, $0 \leq \xi < 180^\circ$; $\delta \chi = 4^\circ \psi = 5^\circ \xi = 5^\circ$) and the reduced rotational search ($-45^\circ \leq \chi$, $\psi < 360^\circ$, $0 \leq \xi < 180^\circ$; $\delta \chi = 5^\circ \psi = 5^\circ \xi = 15^\circ$) are virtually identical. At the same time, confining rotations to a smaller space of $-30^\circ \leq \chi, \psi < 360^\circ$, $0 \leq \xi < 360^\circ$ leads to reduced occurrences of small ET distances, indicating insufficient sampling of the rotational space. Thus, the former protocol — affording a 24-fold gain in the computational time (Figure 2D) — was used in all subsequent simulations.

The conformational search used in this work is not optimized for speed, but rather tailored for a fine sampling of the entire conformational space in the protein complexes. However, significant reduction in the computational time can be achieved by performing the rotational search and outputting the solutions only for the CMs with short separations between the ET cofactors (e.g. < 20 Å). This in effect will discard most of the ET-inactive orientations (i.e. blue areas in the CM distribution maps in Figures 3–6, see below), while sampling the functional epitopes. A further gain in the calculation speed can be achieved by decreasing the resolution of the conformational search, either by increasing the rotational increments or decreasing the spatial resolution, $\delta d$. For instance, setting $\delta \chi$ and $\delta \psi$ to $15^\circ$ instead of $5^\circ$ and increasing the spacing between the neighboring CMs by 1 Å afford, respectively, 9- and 4-fold decrease in the number of sampled orientations. Combining the two strategies and restricting the search to the ET-competent areas only (see above) are expected to yield a 40- to 60-fold decrease in the computational time, depending on the system. Though the suggested scheme will miss out some of the ET-competent orientations (thus lowering the accuracy of the upper-limit $k_{ET}$ estimates), it should still provide robust functional epitope maps.

### ET Rate Calculations

The ET rate constants can be estimated from the distances between redox centers as described by Dutton and co-workers [2,35] or calculated from the ET pathway theory developed by Beratan and Onuchic [36]. The latter, computationally more demanding, approach requires the knowledge of the side-chain contacts across the interface, while the former relies solely on the separation between the electron donor and acceptor sites and appears to be more suitable to the rigid-body search method used here. In a broad range of biological systems, the ET rate constants are described by an exponential dependence on the distances between the redox centers (Equation 8 in the Methods section) [2]. Being a measure of the electronic coupling between the electron donor and acceptor, the exponential decay constant, $\beta$, is a property of the medium through which electron tunneling occurs, with the values varying from $\beta = 1.0$ Å$^{-1}$ and $\beta = 1.3$ Å$^{-1}$ for $\beta$-strands and $\alpha$-helices, respectively, to $\beta = 2.8$ Å$^{-1}$ for the vacuum [2,3]. Analysis of biological ET systems showed that electron tunneling through a protein can be accurately described with a single value of $\beta = 1.4$ Å$^{-1}$ [2], which is used throughout this work.

**Electron Transfer Interactome of Cytochrome c**

**Cc-CeP.** Yeast CeP is a heme-containing enzyme catalyzing two-electron reduction of peroxides using reducing equivalents provided by Cc. Initial reaction with H$_2$O$_2$ yields compound I, containing Fe(IV) = O heme oxoferryl and a cationic indole radical W191”, followed by a two-step reduction to generate the resting-state enzyme [21]. CeP possesses two Cc binding sites with markedly different binding affinities and ET properties. The crystal structure of CeP bound to the high-affinity site (Figure 3A) suggested an ET pathway leading from the Cc heme to the W191 group of CeP [37], while flash photolysis kinetics studies of Hoffman and co-workers [38,39] demonstrated fast heme-heme ET from the Cc heme at the low-affinity site, presumably located in a negatively-charged CeP region identified in Brownian dynamics simulations (Figure 3B) [8]. The intermolecular ET in this complex has been studied in great detail; however, the relative contributions of the heme-W191 and heme-heme ET to the compound I reduction are still a matter of debate [21].

We used the Ce-CeP system to validate our computational approach and, in particular, verify whether it can predict the two ET-active sites. The results are presented in Figure 3. Panels (C) and (E) show the distributions of Cc centers of mass (CM) around CeP colored by the shortest heme-W191 (C) or heme-heme (E) distances determined from 7,776 rotamers at each Cc CM position (see the Conformational Sampling section above). The hotter the color, the smaller the distance and, consequently, the higher the intermolecular ET rate in that particular protein-protein orientation. A set of Ce CMs for the entire conformational space sampled determines the ET-competent epitope in the protein complex. The distributions of redox distances and the corresponding ET rates used for color-coding in panels (C) and (E) are given in plots (D) and (F).

Our analysis shows that protein-protein orientations exhibiting short heme-W191 distances and, consequently, fast heme-W191 ET rates are indeed located around the crystallographic, high-affinity binding site (Figure 3C,D). Moreover, most of these geometries are inaccessible in the ternary complex (compare red and black traces in Figure 3D; also Figure S1A), suggesting that the low-affinity site is incapable of sustaining fast heme-W191 ET. In contrast, the heme-heme ET-active orientations, found primarily in the low-affinity binding region, are virtually unaffected by Ce binding to the high-affinity site (Figures 3E,F and S1B). Thus, the present analysis confirms the findings of the experimental studies and correctly predicts the ET properties of the two binding sites.

When compared to the ET pathway coupling maps of Hoffman and co-workers [40], the heme-W191 functional epitope matches the CeP region exhibiting strong coupling, while the area supporting high heme-heme ET activity differs. This seeming discrepancy stems from the fact that the surface coupling maps, obtained for individual proteins, reflect the intrinsic ET properties of a given macromolecule [40], while the functional epitopes derived in this work report on the intermolecular ET in the protein complex.

Finally, we note that the experimentally measured heme-W191 ET rate is an order of magnitude higher than that calculated from the X-ray structure of the complex [41]. As the crystallographic binding orientation is preserved in solution [42], this discrepancy is not due to differences between the crystal and solution structures. Interestingly, a number of protein-protein orientations, located in the vicinity of the crystallographic binding site, exhibit shorter heme-W191 distances than that observed in the X-ray structure (Figure 3C,D), suggesting the possibility of a faster ET.

**Cc-Cbc1 and Ce-CCO.** Cbc1 and CCO are homodimeric, multisubunit, integral membrane protein complexes that mediate ET in the terminal branch of the mitochondrial respiratory chain, thereby generating proton gradient and membrane potential required for ATP synthesis. In the final step of the aerobic respiration, CCO channels the electron flow from Cc to O$_2$ via a
binuclear CuA centre, heme α, and heme α3/CuB site, the place of oxygen reduction [43]. The CuA centre is located in the subunit II facing the mitochondrial intermembrane space (Figure 4A). Experimental work showed that neighboring residue W104 is critical for the rapid ET from Cc to CuA [44,45] and may serve as an electron entry port [46]. At present, the structure of the Cc-CCO complex is not known. However, modeling studies suggest that Cc binds to a negatively-charged patch on the front side of subunit II, in close proximity to W104 [10,47].

Cc transfers electrons from ubiquinol to Cc via a Q-cycle mechanism involving four redox cofactors: two b-type hemes in the cytochrome b subunit, 2Fe2S iron-sulfur cluster in the Rieske protein, and the heme group in cytochrome c1 (Cc1) – the actual electron donor to Cc [48,49]. The crystal structure of the Cc-Cbc complex from yeast shows a single Cc molecule bound to the Cbc dimer (Figure S3A). Recent molecular dynamics simulations suggest that the interaction between the Cc1 monomers triggers conformational changes, which alter the binding affinity for the second Cc molecule and lead to the observed asymmetric binding [11].

Simulating the interaction of Cc with the intramitochondrial part of the CCO monomer (blue surface in Figure 4A) reveals that the ET-competent binding geometries are located in a funnel-shaped area on the negatively-charged front face of subunit II (Figure 4B,C). We find a protein-protein orientation with the Cc heme and CCO W104 in vdW contact, enabling a remarkably fast ET. The edge-to-edge distance between the two redox centers (3.8 Å) is in a striking agreement with that reported for the best Cc-CCO docking solution in an earlier modeling study [10]. The direct ET from the Cc heme to the CCO CuA site is expected to be much slower due to the larger separation between the two groups (compare red and black traces in Figure 4D). The experimentally measured intermolecular ET rate is a million-fold slower than the upper limit estimated here (Table 1), suggesting that the ET occurs from a less active protein-protein orientation or is conformationally gated.

A steady-state kinetics study of CCO suggested the presence of two ET-active Cc binding sites with different affinities and catalytic properties [50], and an earlier modeling work identified a likely low-affinity binding region [10]. To assess the ET properties of the putative low-affinity site, we simulated the interaction of the second Cc molecule with the Cc-CCO binding geometry showing the shortest separation between the Cc heme and CCO W104 (see above). As most of the ET-competent protein-protein orientations are then effectively excluded by the steric interactions with the first Cc molecule, the resulting docking solutions show large separations between redox cofactors (Figure S2B), suggesting that the low-affinity binding site is not ET active.

We also modeled the interaction of Cc with the intramitochondrial part of the Cbc1 dimer (blue surface in Figure S3A). The Cc binding sites on the two Cbc1 monomers were treated as equivalent, and the data for a single, crystallographic site (Figure S3B) are reported. All ET-competent Cc-Cbc1 orientations are located in a well-defined patch around the crystallographic binding site, in a negatively-charged region of the Cbc1 surface (Figure S3C,D). As the reduction potentials of Cc and Cc1 hemes are nearly identical [49], i.e. activation energy $\Delta G^{\text{act}} = 0$, the Cc-Cbc1 system presents a good example of endergonic tunneling in
biological ET. The experimentally measured intermolecular ET rate is 3–4 orders of magnitude smaller than our estimated \( k_{ET} \) limit (Table 1) and 20–350 times slower than the ET rate calculated from the X-ray structure (Figure S3E, left). Differences between the protein–protein orientations in crystal and solution or possible conformational gating of the ET have been advanced as possible explanations for the observed discrepancy [49].

**Cc-Fcb\(_2\) and Cc-SOX.** Yeast Fcb\(_2\) is a homotetrameric lactate dehydrogenase, with each subunit composed of a b\(_2\)-like (Cb\(_5\)) heme domain and a flavin mononucleotide (FMN)-binding domain, separated by a flexible linker. The FMN group undergoes a two-step reduction by the substrate lactate, with the electrons sequentially transferred to the Fcb\(_2\) heme group and Cc [23]. In the Fcb\(_2\) crystal structure [51], the two domains are in an orientation favorable for the intramolecular ET between the FMN and heme groups (Figure 5A,B). Interestingly, only two heme domains are observed in the tetramer, suggesting that the other pair is motionally disordered.

Found in animals, plants, and bacteria, SOX is a Mo-containing enzyme catalyzing sulfite oxidation in the final step of cysteine and methionine degradation [22]. The crystal structure of chicken SOX [32] shows a protein dimer, with each subunit consisting of a Cb\(_5\) heme domain and a Mo-binding domain, joined by a flexible tether (Figure S5A). Following the two-electron reduction of the Mo center by the substrate, the electrons are transferred to the SOX heme group, followed by the intermolecular ET to Cc [22]. In contrast to Fcb\(_2\), the crystallographic SOX orientation is not favorable for an interdomain ET.

In both Fcb\(_2\) and SOX, the domain-domain mobility is essential for the protein function [22,23]. Albeit linked by a covalent tether, the two domains behave as a transient protein complex, with the heme domain alternating between the catalytic domain-bound and the Cc accessible states as exemplified by the X-ray structures of Fcb\(_2\) and SOX, respectively. Depending on the tether length and flexibility, the relative domain motions can sample a range of kinetic regimes spanning the “unimolecular” and “bimolecular” modes [33]. In order to assess the influence of the interdomain mobility on the intermolecular ET, we generated ensembles of SOX and Fcb\(_2\) domain-domain orientations and assessed their ET properties. For the crystallographic Fcb\(_2\) conformation, most of the Cc-Fcb\(_2\) binding geometries are not ET competent (Figure 5C and E, red trace). However, allowing flexibility of just two residues in the interdomain linker leads to an ensemble of structures with augmented solvent accessibility of the heme group (Figure 5B) and, consequently, markedly increased intermolecular ET rates (Figure 5D and E, black traces). With the shortest heme-heme distances of 13.7 Å in the crystallographic orientation and 5.7 Å in the modeled ensemble, such limited interdomain mobility affords five orders of magnitude enhancement in the ET upper limit (Table 1). Allowing higher degree of linker flexibility leads to more ET competent geometries, but does not further increase the \( k_{ET} \) value (Figure S4). Thus, a restricted interdomain motion pivoted on just two linker residues appears to be sufficient to enable fast intermolecular ET underlying the Fcb\(_2\) function.

In contrast to Fcb\(_2\), the crystallographic SOX heme domain orientation is favorable for the rapid ET to Cc (Figure S5C and E, red trace), with the majority of ET-competent docking solutions found in a negatively-charged interdomain cleft (Figure S5B). To simulate the largest extent of the domain-domain motion, we allowed full flexibility of the interdomain tether (residues 83–94) and generated an ensemble of heme domain orientations constrained only by the linker geometry and the vdW interactions with the other domain (Figure S5D). The shortest Cc-SOX heme–heme distances in the crystallographic orientation and the modeled ensemble are 6.1 Å and 5.5 Å, respectively, converting into very similar \( k_{ET} \) values (Table 1). Thus, though domain mobility leads to an increase in the number of ET-competent Cc binding geometries (Figure S5E, black trace), it does not change the upper limit \( k_{ET} \) value, suggesting that the relative domain orientation observed in the SOX crystal structure is sufficient for the rapid intermolecular ET necessary for the protein function.

**Cc-Cb\(_5\) and Cc-Erv1.** In what follows, we present the results for the Cc-Cb\(_5\) complex; the analysis of the Cc-Erv1 pair is given in the Text S3. Cb\(_5\) is a small hemoprotein carrying electrons between Cb\(_5\) reductase and a number of partners, including fatty acid desaturase, cytochrome P450 and methemoglobin [25]. Given the conflicting reports in the literature [20,25,26], it is not clear whether Cb\(_5\) interacts with Cc in vivo; however, in vitro the proteins form a complex, which has become a paradigm for the study of the intermolecular ET [54,55]. At present, the Cc-Cb\(_5\) structure is not known. A hypothetical binding model of Salemme [5], featuring four charge-charge interactions across the interface.
and nearly coplanar heme groups (Figure 6A), is supported by ample experimental evidence [54,55]. However, recent NMR studies revealed pronounced dynamics within the complex, suggesting the presence of multiple binding geometries [17,56]. Depending on the source of Cb and the experimental conditions, formation of binary Cb-Cc and ternary Cb-(Cc)$_2$ complexes has been reported [54–58].

In a binary complex, the ET-competent protein-protein orientations map out onto a broad patch of the Cb surface (Figure 6C), containing an area surrounding the Salemme binding site and comprising the binding geometries identified by Brownian dynamics [9] and experiment-driven docking [56]. Many of those orientations are still available in a ternary complex (Figure 6D), suggesting that Cc bound to the low-affinity Cb site can be highly ET-active (Figure 6E). The fastest experimentally measured ET rate is five orders of magnitude lower than the upper-limit $k_{ET}$ value and ca. 1,700 times slower than that calculated for the Salemme structural model (Figure 6E, bottom panel). It was suggested before that the ET in this system might be “gated”, i.e. limited by protein dynamics within the complex [55]. If this is indeed the case, re-orientation of protein molecules to a more active bound form is expected to greatly decrease the overall observed ET rate, bringing it closer to the experimental value.

**ET Interactome of Cc**

The functional epitopes of the physiological Cc complexes derived in this work delineate the conformational space available to all ET-competent protein-protein orientations. For the complexes studied, the epitopes map onto the negatively-charged surfaces of the Cc partner proteins. As confirmed by the simulations where the Cc surface is sampled with the Cb and CcP molecules (Figures S6 and S7), Cc maintains frontal orientation in all ET-competent conformations, which is in excellent agreement with the experimental studies showing that the Cc binding interfaces in redox complexes comprise the surface-exposed heme edge and the surrounding ring of lysines (Figure 7A). Taken together, these findings highlight the importance of the electrostatic steering in the ET complexes and confirm that the exposed heme edge serves as the electron exit/entry port in Cc. It is noteworthy that the charge complementarity in Cc
more than one protein-protein orientation and might not require formation of a well-defined, stereospecific complex. Indeed, a recent experimental study of the interactions among a small copper protein amicyanin and several bacterial cytochromes has shown that a highly efficient ET can take place without formation of specific complexes, but rather via loose protein encounters guided by complementary electrostatics [59]. We hypothesize that the Cc ET interactome—particularly complexes with broad functional epitopes such as those in Cb5, Fcb5, SOX, and Erv1—can exhibit a similar behavior.

It is well known that dynamics regulate the ET within proteins and, together with the fluctuations in the surrounding medium, enable ET gating [60]. Similarly, heterogeneity of the binding geometries in the ET complexes—arising from multiple protein-protein orientations sampled during the initial, encounter state of the biomolecular association [61]—is expected to modulate the intermolecular ET rates. Indeed, this notion has been gaining an increasing experimental support. For example, computational analysis of the kinetics data for the interaction of bacterial cytochrome c5 (Cc5) with the photosynthetic reaction centre (RC) revealed the presence of multiple ET-active binding geometries explored in the encounter state [62,63]. In addition, detailed molecular dynamics simulations of the Cc5-RC encounter complex showed that strong ET coupling pathways between the two proteins include interstitial water molecules and illustrated that structural fluctuations within the encounter ensemble enable fast ET rates, in agreement with the experiment [63]. In another series of studies, Hoffman, Beratan and co-workers showed that interacting Cb5 and myoglobin sample multiple ET-active protein-protein orientations, constituting a subset of the binding geometries explored in the encounter state [19,64]. Here, we use the functional epitope mapping to delineate the extent of the conformational space containing all ET-geometries, including those populated in the transient encounter complex.

The functional epitopes in Cc complexes differ in size and exhibit a large variation in the number of ET-competent binding geometries (Figure 7B). For example, a broad epitope in the Cc-Cb5 complex contains ten times more protein-protein orientations with \(k_{ET}>10^6\) s\(^{-1}\) than a narrow, funnel-shaped ET patch in the Cc-CbC1 complex. Interestingly, the size of the ET epitopes in Cc complexes correlates with the extent of protein dynamics. For instance, being a highly dynamic system consisting of multiple binding geometries [56], the Cc-Cb5 complex features a broad functional epitope, while the Cc-CcP complex—where proteins spend 70% of the time in the crystallographic binding orientation and 30% in a loose encounter state [32,42]—has a much narrower ET-competent area. We suggest that the size of the ET epitope can be used as a diagnostic tool to discriminate between highly-dynamic, multiple-orientation complexes (e.g., Cc-Cb5, Cc-SOX, Cc-Fcb2) and their more static, predominantly single-orientation counterparts (e.g., Cc-CcP, Cc-CCO, and Cc-CbC1) (Figure 7B). We hypothesize that differences in the widths of the ET-active areas reflect different evolutionary scenarios, contrasting the need to drive Cc binding to a well-defined, single orientation in Chc1 and CCO complexes (served to prevent wasteful interactions with other parts of these multisubunit assemblies and the negatively-charged mitochondrial membrane) and allowing rapid ET from multiple protein-protein orientations in the Cc-Cb5 complex.

Without the knowledge of the populations of the protein-protein orientations constituting the functional epitope, the occurrences of donor-acceptor distances and the corresponding ET rates derived in this work cannot be converted into the ET probability distributions, and only the upper-limit \(k_{ET}\) values can be calculated. It is conceivable that some of the highly ET-competent

---

**Figure 6. Cc-Cb5.** (A) The hypothetical structural model of Salemme [5]. Cc and Cb5 are in yellow and blue, respectively, with heme groups shown as red sticks. (B) The molecular surface of Cb5 colored by the electrostatic potential (see the scale bar). Protein orientation is the same as in (A). Cc in the Salemme model is shown as yellow cartoon (C) or molecular surface (D). See Videos S5 and S6 for an expanded view of (C) and (D), respectively. (E) Distributions of the intermolecular heme-heme distances (top) and ET rates (bottom). Red and black traces are the distributions for the binary and ternary complexes, respectively. Solid vertical lines indicate the heme-heme distance (top) and the corresponding ET rate (bottom) in the hypothetical model of Salemme [5]. The dashed line denotes the largest experimentally measured ET rate in the Cc-Cb5 complex [16].

doi:10.1371/journal.pcbi.1002807.g006
binding geometries are energetically unfavorable and, therefore, do not form during the interaction. On the other hand, many thermodynamically favorable orientations (be they ET-active or not) might never get populated on the protein-protein association path. Thus, accurate predictions of the ET rates from the functional epitopes would require estimating the relative thermodynamic stabilities of the ET-competent binding geometries and addressing the protein association kinetics on a microscopic level. As illustrated by early modeling efforts of the Cc-CcP complex [6], failure to consider kinetic properties of a system led to an over-optimization of predicted binding geometry (exhibiting high charge complementarity, large number of intermolecular hydrogen bonds, and snug steric fit of the protein sidechains), in stark contrast to the crystal structure featuring a single hydrogen bond and a loosely packed interface [37].

**Table 1.** ET properties of the Cc complexes modeled in this work.

<table>
<thead>
<tr>
<th>Cc partner</th>
<th>Redox centre</th>
<th>r, Å</th>
<th>$k_{ET}$ upper limit, $10^6$ s$^{-1}$</th>
<th>ET epitope size$^a$</th>
<th>$k_{ET}$, $10^6$ s$^{-1}$ measured$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CcP</td>
<td>W191</td>
<td>11.6</td>
<td>1.4 (0.36)</td>
<td>822</td>
<td>2 [41]</td>
</tr>
<tr>
<td></td>
<td>heme</td>
<td>11.1</td>
<td>2.4 (0.81)</td>
<td>1276</td>
<td></td>
</tr>
<tr>
<td>CbsC</td>
<td>heme</td>
<td>5.9</td>
<td>320 (113)</td>
<td>3306</td>
<td>0.55 [16]</td>
</tr>
<tr>
<td>CbsC</td>
<td>Cc, heme</td>
<td>6.2</td>
<td>0.15–2.7 (0.08–1.2)$^b$</td>
<td>144–342$^a$</td>
<td>0.014 [49]</td>
</tr>
<tr>
<td>CCO</td>
<td>W104</td>
<td>3.8</td>
<td>580 (131)</td>
<td>506</td>
<td>0.06 [71]</td>
</tr>
<tr>
<td></td>
<td>CuA</td>
<td>11.1</td>
<td>0.017 (0.016)</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Erv1</td>
<td>C130–C133</td>
<td>4.5</td>
<td>20–350 (5–65)$^b$</td>
<td>1341–2186$^a$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FAD</td>
<td>4.6</td>
<td>5.6–100 (2.4–33)$^b$</td>
<td>891–1622$^a$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ensemble</td>
<td>5.7</td>
<td>470 (212)</td>
<td>2232±431$^f$</td>
<td></td>
</tr>
<tr>
<td>SOX</td>
<td>X-ray</td>
<td>6.1</td>
<td>170 (69)</td>
<td>2268</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ensemble</td>
<td>5.5</td>
<td>390 (101)</td>
<td>4665±916$^f$</td>
<td></td>
</tr>
</tbody>
</table>

$^a$Shortest separation between redox cofactors; $^b$the values averaged over 10% of the most ET-active protein orientations with $k_{ET}>10^6$ s$^{-1}$ are given in parentheses; $^c$number of Cc CMS with $k_{ET}>10^6$ s$^{-1}$; $^d$experimentally measured (pseudo)first order intermolecular electron transfer rate constant (reference given in parentheses); $^e$the range of values reflects the uncertainty in the reorganization energy, $\lambda$ (see Text S2); $^f$the average and the standard deviation calculated over all members of the simulated ensemble.

Used in the past to investigate Cc interactions with CcP [8,65] and CbsC [9,54], Brownian dynamics (BD) simulations afford an effective weighting of various ET-competent orientations and allow extraction of bimolecular rate constants, $k_{ET}$, from association trajectories. When used with the ET reaction criteria (such as distances between the cofactors or the coupling strengths of ET pathways), the BD runs provide estimates for the intermolecular ET rates. Further, combination of BD and molecular dynamics (MD) allows to address the roles of structural fluctuations and the intervening solvent in the intermolecular ET process [63]. While BD and MD/MD are clearly methods of choice for accurate ET rate calculations, as a rule they require a complex force-field, adequate definition of the simulation conditions (e.g. proper partial charge and protonation state assignments for protein sidechains, boundary conditions, etc), a certain level of expertise, and considerable amount of time. Interestingly, the functional epitopes of Cc-CcP and Cc-CbsC complexes delineated with our coarse-grained approach, relying solely on the steric properties of the interacting molecules, are in excellent agreement with those found in BD simulations [8,9,65]. Thus, it appears that the simple rigid-body conformational search employed in this work is sufficient for an adequate description of the ET-competent areas in Cc complexes.

As can be seen from the $k_{ET}$ distributions, the studied Cc complexes contain a large number of binding geometries with the ET rates faster than the experimentally measured values. Even though not all of these ET-competent orientations will be populated in a given protein complex (see above), it seems likely that several highly-active conformations will dominate the intermolecular ET. This finding suggests that the overall ET rates in the Cc interactome are not limited by the ET event itself, but rather controlled by alterations in protein dynamics and/or conformational changes, occurring in protein complexes on slower time scales. Such binding-induced and/or redox-dependent conformational changes, reported for free Cc [66] and its complexes with CbsC [11] and CCO [67], could account for the apparent discrepancies between the observed and predicted ET rates.

**Figure 7. Cc binding interface and ET properties in redox complexes.** (A) Combined binding interface of ferric yeast Cc (PDB entry 2YCC) [73] in complexes with CcP [74], CbsC [56] and CCO [67] as determined by solution NMR spectroscopy. Residues experiencing binding effects in any of the three, two of the three, and all three complexes are colored light blue, green, and dark blue, respectively. The heme group is indicated by the white rectangle. Several lysine residues affected by the binding are labeled. (B) Absolute counts of Cc CMS with $k_{ET}>10^6$ s$^{-1}$ in the ET complexes studied in this work. doi:10.1371/journal.pcbi.1002807.g007
Concluding Remarks

The computational approach used here can be readily extended to other ET systems. Specifically, our analysis does not require the structures of protein complexes; can be applied in combination with homology modeling to systems with unknown protein structure(s) (as exemplified by Cc-Erv1); allows ensemble treatment of flexible groups (as shown for Cc complexes with Fcb2 and SOX); permits incorporation of side-chain optimization protocols (as done in control Cc-Cb5 runs); and enables calculations of the maximal, ‘activationless’ ET rates for the systems with unknown free energy changes (ΔG) and reorganization energies (λ). We believe that this work will engender future studies of the intermolecular ET in other biological networks.

Supporting Information

Dataset S1  Script files.
(BZ2)

Figure S1  Distributions of Cc CMs around CcP in the ternary complexes colored by the heme-W191 (A) or heme-heme (B) distances and the corresponding ET rates (see the scale bar).

Figure S2  Distribution of Cc CMs around the subunit II of CCO in binary (A) and ternary (B) complexes colored by the heme-CuA distances and the corresponding ET rates (see the scale bar).

Figure S3  Cc-Cbc1.

Figure S4  Intermolecular ET in the Cc-Fcb2 complex with the simulated interdomain motion.

Figure S5  Cc-SOX.

Figure S6  Functional epitope of Cc in the complex with Cb5.

Figure S7  Functional epitope of Cc in the complex with CcP.

Figure S8  Cc-Erv1.

Figure S9  The workflow in the computational protocol used for the Cc-CcP complex.

Table S1  Molecular systems modeled in this study.
Text S1  Description of the computational protocol.  
(PDF)

Text S2  Details of the ET rate calculations.  
(PDF)

Text S3  Cc-Erv1 complex.  
(PDF)

Text S4  Supplementary references.  
(PDF)

Video S1  Distribution of Cc CMs around CcP colored by the heme-W191 distances and the corresponding ET rates. For details, see the legend to Figure 3 in the main text.  
(MOV)

Video S2  Distribution of Cc CMs around CcP colored by the heme-heme distances and the corresponding ET rates. For details, see the legend to Figure 3 in the main text.  
(MOV)

Video S3  Distribution of Cc CMs around CcP in ternary complexes colored by the heme-W191 distances and the corresponding ET rates. See the legend to Figure S1 for details.  
(MOV)

Video S4  Distribution of Cc CMs around CcP in ternary complexes colored by the heme-heme distances and the corresponding ET rates. See the legend to Figure S1 for details.  
(MOV)

Video S5  Distribution of Cc CMs around Cb5 in binary complexes colored by the heme-heme distances and the corresponding ET rates. For details, see the legend to Figure 6 in the main text.  
(MOV)

Acknowledgments

We acknowledge the access to the VUB/ULB Computing Centre and thank Wim Vranken for the help with loop modeling, Lieven Buts for the assistance with the computational resource management, and Peter Crowley, Yan Sterckx, and Sophie Vanwetswinkel for critical reading of the manuscript.

Author Contributions

Conceived and designed the experiments: ANV. Performed the experiments: ANV. Analyzed the data: ANV, NAJvN. Wrote the paper: ANV, NAJvN.

Electron Transfer Interactome of Cytochrome c

References


