BBA - Bioenergetics 1860 (2019) 541–548

Contents lists available at ScienceDirect

BBA - Bioenergetics

journal homepage: www.elsevier.com/locate/bbabio

Balanced internal hydration discriminates substrate binding to respiratory complex I

Murilo Hoias Teixeira, Guilherme Menegon Arantes*

Department of Biochemistry, Instituto de Química, Universidade de São Paulo, Av. Prof. Lineu Prestes 748, 05508-900 São Paulo, SP, Brazil

ARTICLE INFO

Keywords:
Ubiquinone
Molecular recognition
Molecular dynamics simulations
Free energy profile

ABSTRACT

Molecular recognition of the amphiphilic electron carrier ubiquinone (Q) by respiratory complexes is a fundamental part of electron transfer chains in mitochondria and bacteria. The primary respiratory complex I binds Q in a long and narrow protein chamber to catalyse its reduction. But, the binding mechanism and the role of chamber hydration in substrate selectivity and stability are unclear. Here, large-scale atomistic molecular dynamics simulations and estimated free energy profiles are used to characterize in detail the binding mechanism to complex I of Q with short and with long isoprenoid tails. A highly stable binding site with two different poses near the chamber exit and a secondary reactive site near the N2 iron-sulfur cluster are found which may lead to an alternative Q redox chemistry and help to explain complex I reactivity. The binding energetics depends mainly on polar interactions of the Q-head and on the counterbalanced hydration of Q-tail isoprenoid units and hydrophobic residues inside the protein chamber. Selectivity upon variation of tail length arises by shifting the hydration balance. This internal hydration mechanism may have implications for binding of amphiphilic molecules to cavities in other membrane proteins.

1. Introduction

Respiratory complex I also known as NADH:ubiquinone oxidoreductase is the main entry protein of electron transfer chains in the inner membrane of mitochondria and in many bacteria. It catalyses oxidation of nicotine adenine dinucleotide (NADH) through a flavin mononucleotide and a chain of iron-sulfur (FeS) clusters to reduce ubiquinone (Q), an amphiphile composed by a *p*-benzoquinone ring (Q-head) attached to an isoprenoid chain (Q-tail), which diffuses along the membrane and carries electrons to subsequent respiratory complexes. Complex I is also a reversible proton pump and couples the redox process with generation of an electrochemical gradient across the membrane, thus contributing to ATP synthesis [1,2,3,4,5]. As an essential metabolic enzyme and a primary site for production of reactive oxygen species, malfunction of complex I has been linked to several common neuromuscular, degenerative and metabolic diseases, ischemia-reperfusion injury and aging [6,7].

Complex I is one of the largest asymmetrical membrane proteins known. The prokaryotic enzyme is usually composed by 14 core subunits (550 kDa mass) containing all the redox centers and proposed proton pumping channels. These subunits are sufficient for catalysis and highly conserved from bacteria to human enzymes. The mammalian complex I has 31 additional supernumerary subunits (total of 45 subunits and 980 kDa mass) involved in complex assembly, stability and specialized metabolic roles.

The first entire atomic structure of complex I was determined [8] for the eubacterium *Thermus thermophilus* enzyme by X-ray crystallography with a resolution of 3.3 Å. The L-shaped structure is composed by a membrane-bound arm where the proton channels are located, and a hydrophilic peripheral arm where all the redox cofactors and NADH binding site are found (Fig. 1A). Mitochondrial structures from the yeast *Yarrowia lipolytica* [9] and from several mammals (bovine [10], ovine [11], porcine [12] and mouse [13]) determined more recently by cryo-EM experiments revealed the external location of supernumerary subunits forming a protective shell around the core. The overall geometry of the core subunits is similar in all determined structures which suggests conservation of Q binding and catalytic mechanisms among species.

A 35 Å long and narrow Q binding chamber was identified on the interface of subunits Nqo4 (49 kDa in bovine nomenclature), Nqo6 (PSST) and Nqo8 (ND1) in the *T. thermophilus* structure (Fig. 1B). This remarkable protein cavity (Q-chamber) starts on top of a cleft between subunits Nqo4 and Nqo6 where the substrate Q-head may bind, receive electrons from the nearby N2 FeS cluster (center-to-center distance ~13 Å) and form hydrogen bonds through its carbonyl oxygens to the side chains of Nqo4 Tyr87 and His38, which are both invariant and

* Corresponding author.
E-mail address: garantes@iq.usp.br (G. Menegon Arantes).

https://doi.org/10.1016/j.bbabio.2019.05.004
Received 6 November 2018; Received in revised form 16 May 2019; Accepted 28 May 2019
Available online 04 June 2019
0005-2728/

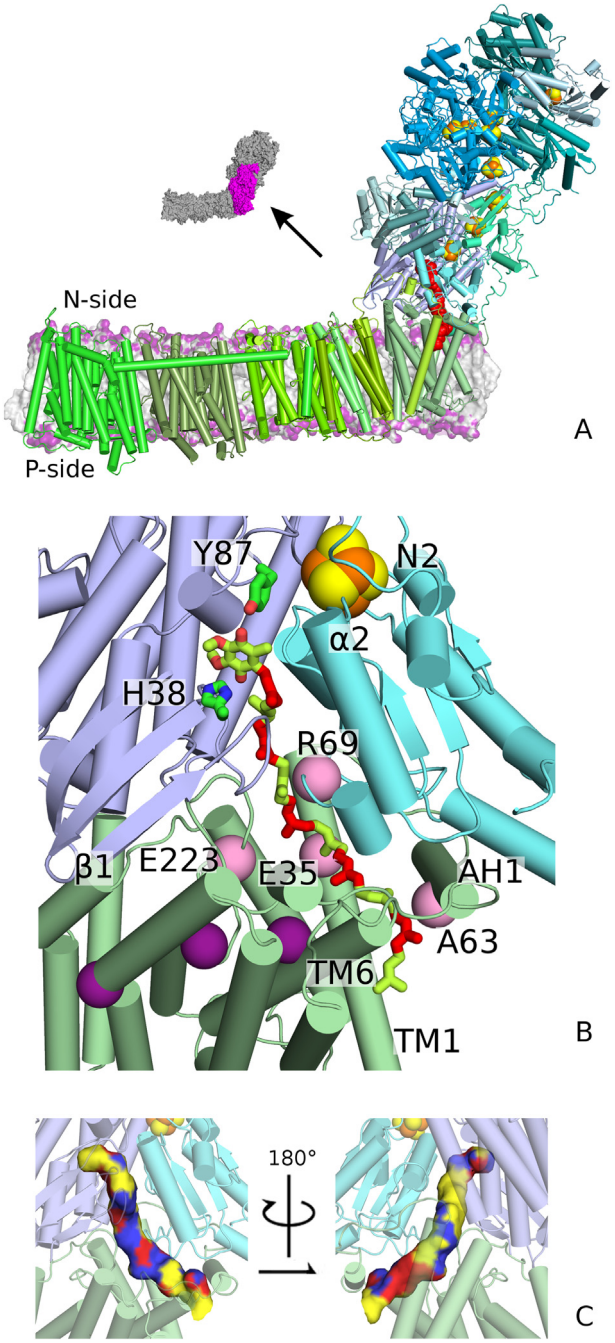

Fig. 1. Structure of the respiratory complex I and its Q binding chamber. (A) Overview in cartoon of the structural model from *T. thermophilus* embedded in a lipid membrane. FeS clusters are shown as orange and yellow spheres. Q_{10} is modeled in the reactive site in red. Inset shows the position of subunits Nqo4, Nqo6 and Nqo8. (B) Close view of subunits Nqo4 in light blue, Nqo6 in cyan and Nqo8 in pale green that form the Q-chamber with Q_{10} colored in green and red alternating for each isoprenoid unit. Elements of secondary structure and residues (green sticks and pink spheres) discussed in the text are indicated. Purple spheres show Nqo8 residues Glu163, Glu213 and Glu248 of the proposed E-channel [8]. (C) Molecular surface of bound Q colored by the first solvation layer: yellow is a hydrophobic residue, red is a hydrophilic residue and blue is water.

required for full enzymatic activity. X-ray data show the inhibitor piericidin A and the substrate analogue decylubiquinone, both containing short hydrophobic tails, bind in the Q-chamber at this position [8,9]. The chamber continues through an amphipathic region, roughly composed by a charged surface facing the membrane arm and by a large hydrophobic patch on the opposite side (Fig. 1B and C). It ends exposed to the lipid membrane in an exit formed by Nqo8 helices TM1, TM6 and amphipathic helix AH1. The Q-chamber geometry and the free volume available to accommodate Q inside are similar among all determined active structures [8,9,11,10,13].

Although complex I has been co-crystallized with small Q analogues [8,9], it is still unclear how natural substrates with long isoprenoid chains such as ubiquinone-10 (Q_{10}) will bind and how many stable binding sites or poses may be present inside the long Q-chamber. Will each putative binding site lead to different Q redox chemistry? Is binding dominated by Q-head or by Q-tail interactions? Does water penetrate in the Q-chamber and play a role during binding? Answers will help to understand the mechanism of substrate binding in the amphipathic Q-chamber and to design more potent and selective inhibitors.

Molecular simulation is a valuable tool to investigate binding pathways and energetics of ligands complexed to protein tunnels or cavities [14,15,16,17,18,19]. For instance, unbinding of a small ligand from the prototypical T4 lysozyme engineered cavity has been shown to proceed through multiple competitive tunnels in the protein [19]. Pathways of plastoquinone exchange along different channels in the photosystem II complex have been determined by coarse-grain models [20]. For the respiratory complex I, molecular simulations have been performed [21,22,23] and a recent study [24] explored how the Q redox state modulates binding in the Q-chamber and triggers the coupled proton pumping.

Here we present large scale molecular dynamics simulations and estimate free energy profiles for binding of Q with long (Q_{10}) and short (ubiquinone-2, Q_2) isoprenoid chains to the hydrated chamber inside complex I from *T. thermophilus*. Notably, our model is based on the entire complex I structure [8] with corrected coordinates for Nqo6 loop $\alpha 2 - \beta 2$ and on a calibrated force field for Q [25], as described in the next section. Results show Q is stable in two separated binding and reactive sites, respectively near the Q-chamber exit and cluster N2. The binding site (BS) is broad with at least two different stable poses. The binding mechanism depends on an interesting but previously uncharacterized interplay of Q-head interactions and hydration of both Q-tail and internal chamber residues. We discuss the implications for substrate selectivity and reactivity catalysed by complex I and conclude this binding mechanism may be employed by amphiphilic molecules binding to other membrane proteins.

2. Materials & methods

2.1. Set-up of protein model and molecular dynamics simulations

The X-ray structure of the complete respiratory complex I from *T. thermophilus* (PDB ID: 4HEA) [8] was used to build the protein model. All core subunits (Nqo1-14) and two accessory subunits Nqo15-16 found in species related to *T. Thermophilus* were included. Coordinates for internal segments from subunit Nqo3 (residues 56-72 and 144-147) missing from the PDB file were constructed *de novo* using MODELLER (version 9.15) [26] with default settings. The Nqo3 residues are distant from the Q-chamber by more than 50 Å. Nqo6 residues 65–69, also missing from the PDB file, are located in the Q-chamber and in direct contact with bound Q. These residues are part of loop $\alpha 2 - \beta 2$ found in a different conformation in the corresponding PSST domain of the resolved active mammalian structures [11,10,13] (Fig. S1). Thus, we decided to rebuilt the entire loop (segment between residues D55-P72 of Nqo6) with MODELLER but using the equivalent segment (A82-P98 of subunit PSST) from the ovine enzyme (PDB ID: 5LNK) [11] as a structural template. The other two flexible loops in the Q-chamber (Nqo4 $\beta 1 - \beta 2$ and Nqo8 TM5–TM6) should be in appropriate position for Q binding as the *T. thermophilus* structure was co-crystalized with a substrate analogue. The rebuilt coordinates are in excellent agreement with a refined model for the *T. Thermophilus* structure (Leonid Sazanov

– IST Austria, personal communication) in which coordinates for Nqo6 loop $\alpha 2 - \beta 2$, including residues 65–69, were rebuilt directly from re-refinement of the original electron density [8].

Protonation states of side-chains were adjusted to neutral pH (positive charge for K and R, negative for D and E, and neutral for all other residues), except for Nqo4 His38 which was protonated. This is the expected His38 reactive state for complex I bound to Q, as suggested by site-direct mutagenesis [27] and simulations [21]. All FeS centers were modeled in their reduced form. Although the FeS cluster ensemble may be partially oxidized, cluster N2 is clearly reduced [28]. The exact redox state of other FeS centers is less important for Q binding because of their greater distance to the Q-chamber (> 20 Å).

Ubiquinone was modeled in the oxidized form with 10 (Q_{10}) or 2 (Q_2) isoprenoid units. Q_2 was first docked in the reactive position near Nqo4 Tyr87 and His38 using AutoDock 4 [29]. The remaining isoprenoid units for Q_{10} were manually placed along the Q-chamber [8] in an extended conformation. The last two isoprenoid units in Q_{10} pass through the least narrow Q-chamber exit formed between Nqo8 helices TM1, TM6 and AH1 and are exposed to the lipid membrane. All protein atoms remained in their X-ray structure as the Q-chamber has enough free volume to accommodate Q without protein movements.

The protein complex was embedded [30] in a solvated POPC (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine) membrane with 1036 lipid molecules, 192,951 water molecules, and 650 Na^+ and 607 Cl^- ions to neutralize the total system charge and keep a 0.1 M salt concentration, resulting in a total of 794,102 atoms. This model was relaxed during two molecular dynamics simulations of 100 ns each, first with all protein heavy atoms tethered to their initial position by harmonic restraints, then with all atoms free to move. This procedure resulted in the initial structural model used for the free energy simulations.

Modeling of the Q binding process requires a balanced description of Q interactions with the hydrated protein interior and the lipid membrane. This was accomplished by using a calibrated force-field for Q, previously shown to give very good agreement between simulated and experimental water-to-membrane partition coefficients and free energies [25,71]. Interactions of protein, lipids and ions were described with the all-atom CHARMM36 force-field [31,32]. Water was represented by standard TIP3P [33]. FeS centers were described using the Chang & Kim [34] parameters with corrections proposed by McCullagh & Voth [35]. These corrections are important to keep the cuboidal structure in Fe_4S_4 clusters.

All molecular dynamics simulations were performed with GROMACS (versions 5.1.3 and 2016.3) [36] at constant temperature of 310 K, pressure of 1 atm and a time step of 2 fs. Long-range electrostatics was treated with the Particle Mesh Ewald method [37]. Further details are given in the Electronic Supplementary Information (ESI). Visualization and figure plotting were done using PyMol [38] and Matplotlib [39]. All simulation data and workflow scripts are available from the authors upon request.

2.2. Reaction coordinates and free-energy calculations

Two reaction coordinates were used to describe the Q binding process: the distance between the centers of mass (dCOM) of the six carbon atoms in the Q ring and ten C_α of residues in subunit Nqo4 surrounding the Q-head in the reactive site (Fig. S2); and a pathway collective variable (Path CV) [40] using a distance metric [16,17] as implemented in PLUMED 2.3.1 [41] between the Q-head and 55 C_α of residues in subunits Nqo4, Nqo6 and Nqo8 (Fig. S2) in direct contact with Q inside the chamber. The Path CV was evaluated with respect to 30 milestone configurations representing progressive binding of Q along the chamber (see ESI Methods). These two kinds of reaction coordinates have already been used successfully to describe ligand binding along protein tunnels or cavities [16,17,18,19]. Results are presented using the shifted coordinate sdCOM = 4 nm – dCOM, so that entrance of Q from the membrane into the Q-chamber proceeds from low to high values and is seen from left to right in the figures shown. This representation is more intuitive than using dCOM or Path CV, which run from high to low values during entrance of Q. All the conclusions drawn here are equivalent when projections over either dCOM (and its sdCOM representation) or the Path CV are used to describe the results (Fig. S3 and ESI).

Free energy profiles for Q binding were estimated with umbrella sampling (US) simulations [42]. Initial configurations for each umbrella window were generated from the initial structural models described above by pulling the terminal carbon in the isoprenoid tail along the membrane plane (XY direction) with a pulling velocity of 0.2 m/s. About 50 ns of steered molecular dynamics were enough to drive Q from the initially prepared reactive configuration (dCOM = 0.4 nm) towards protein dissociation to the membrane (dCOM = 4.1 nm). US windows separated by 0.1 nm were chosen to cover the full range of dCOM, which was restrained with a harmonic potential with force constant $k_{umb,dCOM} = 2000\ \mathrm{kJ\,mol^{-1}\,nm^{-2}}$. Additional windows were introduced to increase sampling and overlap of the dCOM and Path CV distributions, which were then restrained with force constants $k_{umb,dCOM} = 200$ and $k_{umb,Path} = 200\ \mathrm{kJ\,mol^{-1}\,nm^{-2}}$. A total of 55 and 46 windows were used to emulate binding for Q_{10} and Q_2, respectively (Table S1). Each coordinate window was sampled for 200 ns for Q_{10} and 150 ns for Q_2, with a sample collected every 20 ps. Total aggregate simulation time was over 28 μ s, including preliminary runs shown in the ESI. Potentials of mean-force were obtained from the 2D reaction coordinate distribution with WHAM [42,43] and the statistical uncertainty was estimated as 95% confidence intervals by bootstrap analysis with 50 resampling steps [44]. The initial 50 ns of each window were discarded to allow equilibration of orthogonal degrees of freedom. The 1D profiles shown are projections over the minimum 2D free energy pathway. All calculated properties such as number of contacts, hydration, COM distances, etc, were collected over the aggregated windows for each US simulation, binned along the reaction coordinate and averaged for each bin.

3. Results

3.1. Q transits through a hydrated chamber with a highly stable binding site

Q binding is described here by the reaction coordinate sdCOM (see Methods and Fig. S2). Entrance of Q from the lipid membrane into complex I corresponds to sdCOM running from low to high values.

Fig. 2 shows the free energy profile for Q_{10} binding estimated from umbrella sampling with molecular dynamics simulations. Before entrance of Q, the chamber interior in complex I is filled with tens of water molecules (Fig. S4). To initiate the binding process, a Q_{10} molecule approaching from the membrane pool will exchange contacts with lipids for exposed hydrophobic residues in Nqo8 helices TM1, TM6 and AH1 (denominated CE position, inset I of Fig. 2A. See also Figs. 2E and S5B). The hydrophobic and flexible isoprenoid Q-tail often folds on itself (Fig. S6A) and the average Q-head distance to the membrane center is only 0.2 nm higher than the equilibrium distance when Q is free in the lipid (Fig. 2B). This should minimize the free energy cost associated with formation of the initial protein-Q encounter complex [25]. The Q-head passes through the hydrophobic chamber exit at sdCOM = 0.4 –0.6 nm (Fig. S6A) with almost no energy cost as there is enough area for Q transit without significant protein deformation (Fig. S6B).

A broad and highly stable binding site is found between 1.0 < sdCOM < 1.5 nm (BS position, inset II in Fig. 2A) with two isoenergetic binding poses: at sdCOM = 1.1 nm the Q-head is highly hydrated and, at sdCOM = 1.4 nm Q oxygens form hydrogen bonds with conserved Nqo8 Arg36 and Lys65 (Fig. 3A). This is the global minimum of the free energy profile, suggesting that complex I will be frequently loaded with a Q_{10} molecule at this site. Hydrophobic residues internal

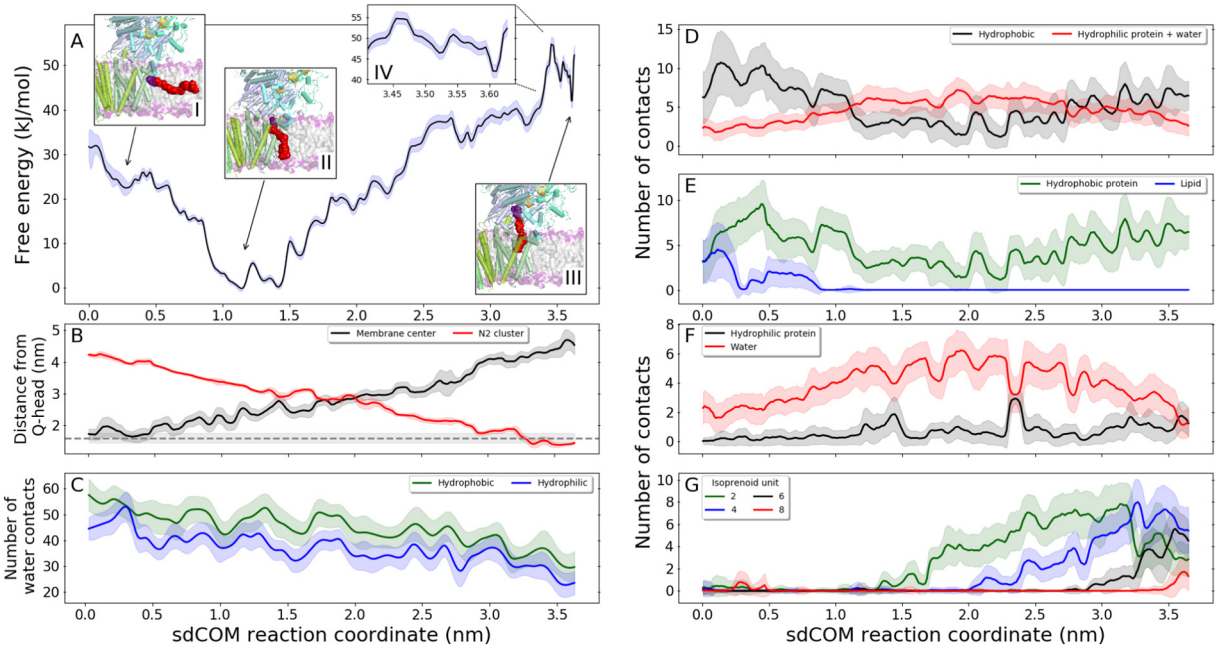

Fig. 2. Binding of Q_{10} along the chamber in respiratory complex I. (A) Free energy profile (black line) with statistical uncertainty (blue shadow). Insets I, II and III show the structure of complex I with Q_{10} (red, Q-head in purple) located in the chamber exit (CE), binding site (BS) and reactive site (RS), respectively. Inset IV zooms the profile in the RS region. (B) Distance of Q-head COM to the membrane center (black) and to the N2 cluster COM (red). Dashed line (gray) indicates the Q-head to membrane equilibrium distance when Q is free in the lipid [25], 1.60 ± 0.15 nm. (C) Number of water contacts to hydrophobic (green) and hydrophilic (blue) protein residues in the Q-chamber interior. Q oxygen contacts with: (D) hydrophobic (black, apolar protein + lipid hydrocarbon tail) and hydrophilic (red, polar and charged protein + water) groups; (E) hydrophobic protein (green) and lipid (blue, both polar head and hydrophobic tail) groups; (F) hydrophilic protein (black) and water (red) groups. (G) Water contacts with Q_{10} isoprenoid units 2 (green), 4 (blue), 6 (black) and 8 (red). The x-axis in all panels displays the sdCOM reaction coordinate. In panels B–G, lines indicate average properties and shadows indicate one standard deviation.

to the Q-chamber dehydrate due to entrance of Q and partial expulsion of water molecules from the chamber to the aqueous phase (Fig. 2C). Q-head interactions with hydrophobic groups decrease, hydrophilic contacts are formed and hydration increases from an average of 2 water contacts in sdCOM = 0.3 nm to almost 5 contacts in sdCOM = 1.2 nm (Fig. 2D and F).

As Q_{10} proceeds into the chamber, the free energy increases through a sequence of shallow local minima and reaches a plateau of ~37 kJ/mol between 2.6 < sdCOM < 3.4 nm (Fig. 2A). Considerable wetting of isoprenoid units 1 to 4 is observed in the same sdCOM range (Fig. 2G).

In the minimum found at 3.2 < sdCOM < 3.4 nm, a Q carbonyl oxygen can form a hydrogen bond with the protonated Nqo4 His38 side-chain (pre-RS position, Fig. 3B). Q is still far from Nqo4 Tyr87 and the other closest acidic residue, Nqo8 Glu225, is more than 8Å away. In order to form hydrogen bonds with both His38 and Tyr87, the barrier at sdCOM = 3.45 nm has to be transposed.

A metastable reactive site (RS position, inset III of Fig. 2A) is found at 3.5 < sdCOM < 3.65 nm where the lowest free energy is 42 kJ/mol (inset IV of Fig. 2A). The chamber is more hydrophobic in this region (Fig. 1C) and the Q-head dehydrates to form hydrogen bonds with both Tyr87 and His38 (Fig. 2D and F). The Q-head is 13 Å from the N2 cluster, in appropriate position to be reduced to the quinol form (Fig. 3B). Two other minima may be identified in the RS (using the Path CV reaction coordinate, Fig. 3C), with the Q-head performing only one hydrogen bond with either Tyr87 (Path CV = 4) or His38 (Path

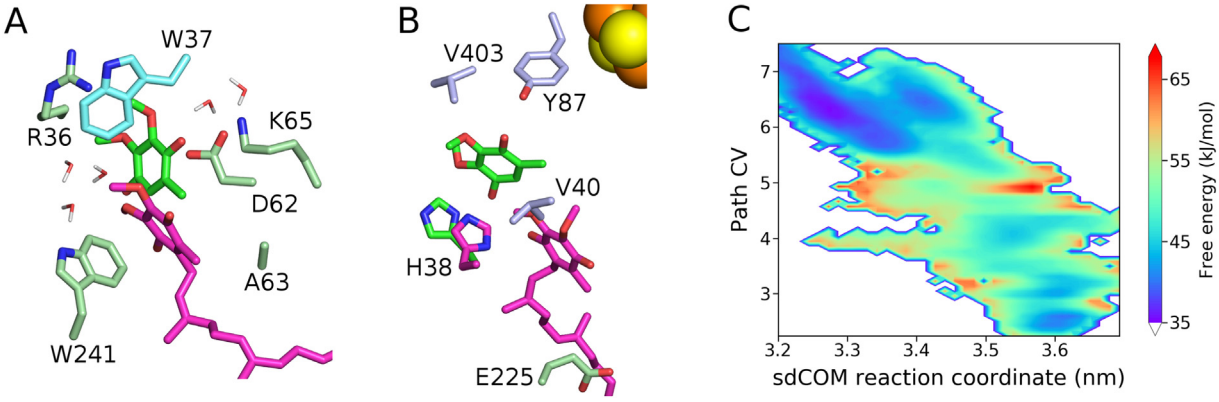

Fig. 3. Binding (BS) and reactive (RS) sites in complex I. (A) Protein environment in the BS with two stable Q_{10} positions in magenta (sdCOM = 1.1 nm) and green (sdCOM = 1.4 nm). Isoprenoid tail of the green Q was removed to ease visualization and water molecules indicate chamber hydration. (B) Environment of Q_{10} at the pre-RS position (sdCOM = 3.3 nm, Path CV = 6) in magenta and at the lowest minimum in the RS (sdCOM = 3.6 nm, Path CV = 2.5) in green, with the two respective Nqo4 His38 conformations shown in the same colors. Other residues are colored as their subunits in Fig. 1. (C) Two-dimensional free energy profile for Q_{10} binding in the RS region.

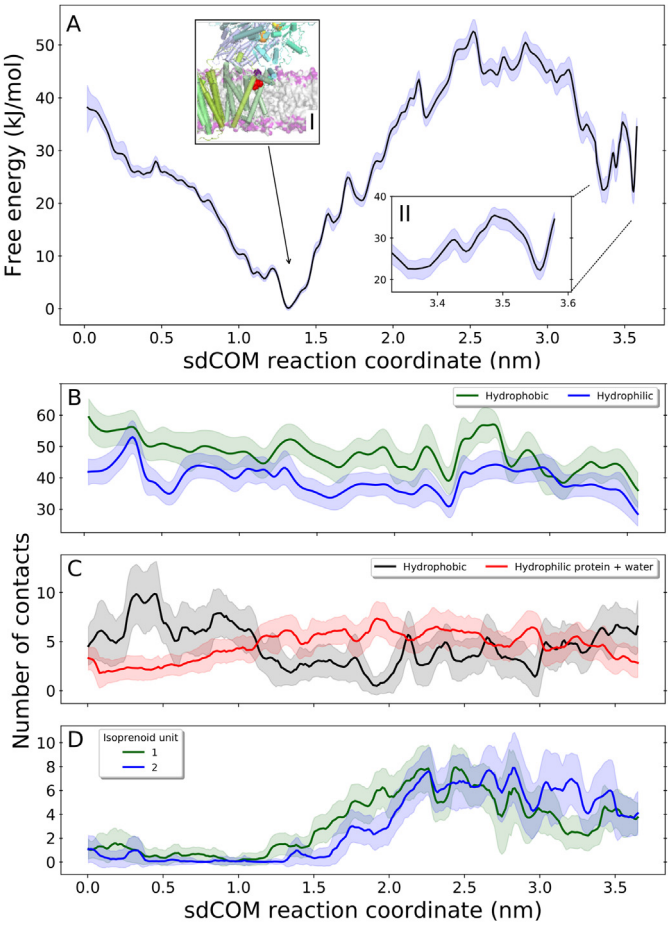

Fig. 4. Binding of Q_2 along the chamber in complex I. (A) Free energy profile (black line) with statistical uncertainty (blue shadow). Inset I shows the structure of complex I with Q_2 (red, Q-head in purple) at the binding site and inset II zooms in the RS region. (B) Water contacts to hydrophobic (green) and hydrophilic (blue) protein residues internal to the Q-chamber. (C) Q oxygen contacts with hydrophobic (black) and hydrophilic (red) groups; (D) Water contacts with Q_2 isoprenoid units 1 (green) and 2 (blue).

CV = 3), as found in a previous simulation [22].

There is a network of ionic side chain contacts near the Q-chamber (Fig. S5A). While part of these contacts remain stable during the full binding process (for instance, Nqo8 Arg294 and Lys65. See Fig. S7 and the section "Possible role of Q binding in activation of the E-channel" in ESI Supplementary Results), some ionic contacts are perturbed either by hydrogen bonding with the Q-head (Nqo8 Arg36 and Nqo6 Arg69) or indirectly (Nqo8 Arg216) via a combination of electrostatic interactions. In the proposed E-channel (Fig. 1B), Nqo8 Glu213 approaches Glu163 at 2.2 < sdCOM < 3.5 nm in response to coordination with Arg216, which also depends on contacts with Nqo8 Glu248 and the highly flexible Glu223 (Fig. S7). These correlated motions might link Q transit with activation of the E-channel and proton pumping in the membrane arm.

3.2. A shorter isoprenoid tail increases stability of Q in the reactive site

Although natural complex I substrates have 6 to 10 isoprenoid units, experiments are often carried out with more soluble short-tail analogues [45,46,47,1,27]. The effect of Q-tail length on binding was investigated by simulating Q_2 as shown in Fig. 4. The free energy profile is similar to Q_{10} binding up to Q entrance in the BS (sdCOM < 1.5 nm, Fig. 4A). The only difference is the relative smaller stability of the hydrated BS pose for Q_2 (sdCOM = 1.1 nm, similar to a binding site identified for Q_1 in another simulation study [24]). Hydration of the Q_2-tail and the chamber interior, and contacts performed by the Q_2-head (Fig. 4 B–D) are also equivalent to those performed by Q_{10} up to this point of the binding process.

As Q_2 proceeds into the chamber, the free energy increases steeply and reaches a rough plateau of ~45 kJ/mol between 2.3 < sdCOM < 3.1 nm (Fig. S5C). Unfavorable hydration of the isoprenoid tail and of hydrophobic residues inside the Q-chamber also increase in this profile region (Fig. 4B and D).

At sdCOM ~ 3.2 nm the free energy decreases simultaneously to dehydration of the Q_2-tail and of chamber residues near the more hydrophobic RS (Figs. 4D and S6D). During the Q_2 binding process, water contacts with hydrophobic residues inside the chamber decrease for sdCOM ≤ 1.7 nm when the Q-tail is occupying the BS, then increase until sdCOM ~ 2.7 nm due to re-hydration of residues near the BS and the chamber exit, and decrease again as the Q-head approaches the RS (Figs. 4B and S6E).

The barrier at sdCOM ~ 3.4–3.5 nm separating the RS is independent of Q-tail length (13 kJ/mol for Q_2 and 15 kJ/mol for Q_{10}, Figs. 4A and 2A) and corresponds to dissociation of the hydrogen bond formed between the twisted His38 side chain and a Q carbonyl oxygen. As the Q-head enters the RS, His38 flips back and reforms contacts with the Q-head and Nqo4 D139 (Figs. 3B and S6E). At the RS, Q_2 has only one clear local free energy minimum (Fig. S3E) which is 20 kJ/mol more stable than the Q_{10} minimum at the RS.

4. Discussion

4.1. Binding stability depends on Q-head interactions and on counterbalanced hydration of Q-tail and chamber residues

The free energy profiles estimated here for Q binding into complex I may be rationalized by the interplay of Q-head interactions and hydration of both Q-tail and apolar residues inside the amphipathic Q-chamber.

For initial binding of either Q_2 or Q_{10}, the Q-head exchanges unfavorable hydrophobic interactions with lipids and the apolar protein exterior for stabilizing hydrophilic interactions inside the BS. Hydration of the isoprenoid tail is minimal and entrance of Q expels to the aqueous phase part of the water molecules in contact with apolar residues near the chamber exit. These interactions and the considerable driving force (stabilization of 20–25 kJ/mol) for initial Q binding are independent of tail length.

However, internal hydration varies with the length of the Q-tail as binding proceeds towards the RS (Fig. S4). For Q_{10}, hydrophobic residues in the chamber are continuously dehydrated due to internal water expulsion by the long Q-tail and the free energy increases mostly due to hydration of isoprenoid units. For Q_2, both isoprenoid tail *and* apolar residues near the chamber exit are hydrated, leading to an even higher free energy. As Q_2 approaches the more hydrophobic RS, the free energy decreases due to partial dehydration of the short isoprenoid tail. For Q_{10}, the free energy remains relatively stable while approaching the RS because dehydration of primary isoprenoid units (1–2) is compensated by hydration of units (5–6) in the middle of the Q-tail. This balanced internal hydration is a particular case of hydrophobic effect that may be employed for selective recognition of amphiphilic ligands [48,49].

The amphipathic nature of the Q-chamber (Fig. 1C) is conserved among species (Fig. S5D) suggesting the balance of interactions described here is essential for binding of the amphiphilic Q molecule. In fact, the role of long isoprenoid tails in dewetting the Q-chamber may have exerted part of the evolutionary pressure that led to complex I natural substrates with 6 or more isoprenoid units.

The extent of internal hydration in the Q-chamber may be questioned [50] since the highest resolution of available complex I structures (3.3 Å [8,13]) is not enough to determine the position of water

molecules. But, water exchange from inside the Q-chamber is frequently observed and takes a few ns or less during our simulations (Fig. S8), in line with the hydration of internal cavities observed for other stable proteins [48,51]. The high water content found inside the Q-chamber here and in previous simulations of complex I [21,23] may also play a structural role for stabilization of the network of ionic residues present near the BS by shielding part of the unfavorable dielectric effect of charges buried in a protein [52]. In mitochondrial complex I, external supernumerary units [9,11,10] may decrease water penetration in the Q-chamber, and hence lead to a relatively less stable minimum for Q binding in the BS and lower free energies for Q reaching the RS.

4.2. Mechanistic interpretations in comparison to previous experiments and future proposals

The high affinity BS found near the Q-chamber exit will be often loaded with a substrate molecule, even if Q is depleted from the membrane pool. Unwanted or reverse redox reactions involving Q bound in the BS will be minimized given the large distance to FeS centers [53].

A sequential and separated BS in the Q-chamber [54,3] is supported experimentally by the slow-relaxing EPR signal (SQ_{Ns}) attributed to semiquinone radical interaction with cluster N2 separated by ~ 30 Å (Fig. 2B) [55,56,57], and by auto-inhibition of complex I reductase activity observed in high concentrations of short-chain Q substrates [45,46,58].

However, neither Q with a short or long tail have been observed bound to the BS in reported complex I structures. The intensive treatment with detergent used for protein purification will wash out endogenous Q. The BS has a broad free energy basin, where at least two Q-head poses were found (Fig. 3A). This will blur Q densities and difficult the resolution of a unique binding position. Dehydration of the water rich BS during crystallization may also destabilize Q binding at this site, in favor of the less hydrated RS. The relative free energy difference between both sites is quite low for short Q-tail and may depend on medium composition or even be reversed in crystals [8,9].

Structural determination carried out using lipid phases with high content of Q_{10} or long-tail analogues may be able to observe Q bound to the BS. During enzyme turnover, Q binding to the BS is expected to involve a fast pre-equilibrium in relation to slower formation of the RS state. This could be probed by pre-steady state experiments of burst kinetics, analyzing the relaxation time in different substrate concentrations [59].

The pre-RS pose (sdCOM = 3.3 nm) found here with Q protonation only by His38 for both Q_{10} and Q_2 suggests a secondary site for Q reduction and possible production of $QH^{(\cdot)}$ semiquinone radicals [56,57]. This could be an alternative reduction mechanism when re-protonation of Tyr87 during turnover has been impaired or when Tyr87 is mutated (Table S5) [60]. For two-electron reduction, the high basicity of Q^{2-} or even QH^- suggests that the second proton to form the quinol could be borrowed from a water molecule present in the chamber. The pre-RS site corresponds to the position occupied by the inhibitor 4-quinazolinylamine co-crystalized with the *Y. lipolytica* complex I structure [9]. This site distance to the N2 cluster is 13–15 Å (Fig. 2B), also in line with the fast-relaxing EPR signal (SQ_{Nf}) attributed to semiquinone formation [55,56,57].

The experimental turnover rate of complex I catalysis is similar for reduction of Q with 10 to 4 isoprenoid units (average $k_{cat} = 380 \pm 39\,s^{-1}$), but decreases significantly for Q with shorter tail ($k_{cat} = 138 \pm 7\,s^{-1}$ for Q_2 reduction) [50]. This rate drop-off may be interpreted by assuming the intrinsic barrier for redox conversion of oxidized Q to quinol (QH_2) bound at the RS is independent of Q-tail length and by noting that free energies calculated here show similar barrier heights for Q_2 and Q_{10} but higher stability for Q_2 (favored by 20 kJ/mol when bound at or near the RS). Thus, the rate-limiting step in the slower Q_2 turnover should change in relation to Q with long tails and take place after formation of bound quinol, probably product release. Due to high stability of Q bound to the BS and possible fast Q entrance inside complex I (high k_{on}), a second short-tail substrate molecule may easily enter the chamber after the first one, and block dissociation once the first is reduced. This would not occur for Q with a long tail which occupies the BS and is also in line with the auto-inhibition of complex I activity observed in high concentrations of Q with short tails [45,46,58].

Another molecular dynamics simulation study of Q binding into the chamber in complex I was recently published [24]. A free energy profile obtained with umbrella sampling was shown for Q_1 binding and a stable BS was also found. The profile is qualitatively similar to the Q_2 profile presented here, but with lower free energies for Q moving towards the RS. Q_{10} binding was also studied, with an approximate diffusion model. The resulting free energy profile is almost flat and completely different from the Q_{10} profile presented here. These differences may be due to the unpublished force-field, truncated complex I structural model (for US simulations) and significantly shorter (10 to 20 times less than here) simulation times used, and precluded analysis of chamber hydration along the binding process [24].

5. Conclusions

Free energy profiles obtained with large-scale molecular dynamics simulations for binding of amphiphilic molecules Q_2 and Q_{10} into complex I were presented here. A secondary reactive site was found near cluster N2 with Q-head protonation by the rotated His38 sidechain. This site may lead to an alternative Q redox chemistry. A binding site near the chamber exit was also found with two binding poses stabilized by hydrophilic Q-head contacts. The profiles also help to interpret previous complex I observables, propose novel experiments to confirm the stability of Q at the secondary binding site, and suggest that mechanistic inferences of complex I binding and activity of natural Q substrates should be made with caution when based on experiments with short tail analogues.

Interestingly, the energetics of Q transit inside the chamber towards the reactive site is determined by the balanced hydration of Q-tail and of apolar residues within the Q-chamber. The Q-tail length modulates substrate selectivity by shifting this balance: a long tail will expel more water molecules from the Q-chamber but also bear more isoprenoid groups hydrated by the remaining internal water, whereas a short tail will leave more water inside the chamber and in contact with hydrophobic residues. This mechanism may be employed by amphiphilic molecules binding to internal hydrated cavities in other membrane proteins, such as lipid flippases and scramblases [61] or fatty acid-binding proteins [62].

Transparency document

The Transparency document associated with this article can be found, in online version.

Acknowledgments

We thank Vanesa V. Galassi (USP) for assistance with the initial model construction, Sandro R. Marana (USP) for fruitful discussions, Leonid Sazanov (IST, Austria) for showing us the revised *T. Thermophilus* structure before publication and several comments from participants of the GRC Bioenergetics Meeting in 2017 when an initial version of this work was presented. Funding from FAPESP (projects 14/21900-2 and 16/24096-5) and computational resources from the SDumont cluster in the National Laboratory for Scientific Computing (LNCC/MCTI) are gratefully acknowledged.

Supplementary data

Additional description of the computational methods, results & discussion, including seven figures, five tables, one animated video and additional references ([63,64,65,66,67,68,69,70]) are available online. Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbabio.2019.05.004.

References

[1] U. Brandt, Energy converting NADH: quinone oxidoreductase (complex I), Annu. Rev. Biochem. 75 (2006) 69–92.
[2] J. Hirst, Mitochondrial complex I, Annu. Rev. Biochem. 82 (2013) 551–575.
[3] M. Verkhovskaya, D.A. Bloch, Energy-converting respiratory Complex I: on the way to the molecular mechanism of the proton pump, Int. J. Biochem. Cell Biol. 45 (2013) 491–511.
[4] L.A. Sazanov, A giant molecular proton pump: structure and mechanism of respiratory complex I, Nat. Rev. Mol. Cell Biol. 16 (2015) 375–388.
[5] M. Wikström, V. Sharma, V.R.I. Kaila, J.P. Hosler, G. Hummer, New perspectives on proton pumping in cellular respiration, Chem. Rev. 115 (2015) 2196–2221.
[6] J. Pätsi, P. Maliniemi, S. Pakanen, R. Hinttala, J. Uusimaa, K. Majamaa, T. Nyström, M. Kervinen, I.E. Hassinen, LHON/MELAS overlap mutation in ND1 subunit of mitochondrial complex I affects ubiquinone binding as revealed by modeling in *Escherichia coli* NDH-1, Biochim. Biophys. Acta 1817 (2012) 312–318.
[7] S. Cadenas, Mitochondrial uncoupling, ROS generation and cardioprotection, Biochim. Biophys. Acta 1859 (2018) 940–950.
[8] R. Baradaran, J.M. Berrisford, G.S. Minhas, L.A. Sazanov, Crystal structure of the entire respiratory complex I, Nature 494 (2013) 443–448.
[9] V. Zickermann, C. Wirth, H. Nasiri, K. Siegmund, H. Schwalbe, C. Hunte, U. Brandt, Mechanistic insight from the crystal structure of mitochondrial complex I, Science 347 (2015) 44–49.
[10] J. Zhu, K.R. Vinothkumar, J. Hirst, Structure of mammalian respiratory complex I, Nature 536 (2016) 354–358.
[11] K. Fiedorczuk, J.A. Letts, G. Degliesposti, K. Kaszuba, M. Skehel, L.A. Sazanov, Atomic structure of the entire mammalian mitochondrial complex I, Nature 538 (2016) 406–410.
[12] M. Wu, J. Gu, R. Guo, Y. Huang, M. Yang, Structure of mammalian respiratory supercomplex $I_1III_2IV_1$, Cell 167 (2016) 1598–1609.
[13] A.-N.A. Agip, J.N. Blaza, H.R. Bridges, C. Viscomi, S. Rawson, S.P. Muench, J. Hirst, Cryo-EM structures of complex I from mouse heart mitochondria in two biochemically defined states, Nat. Struct. Mol. Biol. 25 (2018) 548–556.
[14] J.C. Gumbart, I. Teo, B. Roux, K. Schulten, Reconciling the roles of kinetic and thermodynamic factors in membrane-protein insertion, J. Am. Chem. Soc. 135 (2013) 2291–2297.
[15] F. Bai, Y. Xu, J. Chen, Q. Liu, J. Gu, X. Wang, J. Ma, H. Li, J.N. Onuchic, H. Jiang, Free energy landscape for the binding process of Huperzine A to acetylcholinesterase, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 4273–4278.
[16] I. Bisha, A. Rodriguez, A. Laio, A. Magistrato, Metadynamics simulations reveal a Na+ independent exiting path of galactose for the inward-facing conformation of vSGLT, PLoS Comput. Biol. 10 (2014) 1–8.
[17] M. Paloncýová, V. Navrátilová, K. Berka, A. Laio, M. Otyepka, Role of enzyme flexibility in ligand access and egress to active site: bias-exchange metadynamics study of 1,3,7-trimethyluric acid in cytochrome P450 3A4, J. Chem. Theory Comput. 12 (2016) 2101–2109.
[18] H. Torabifard, G.A. Cisneros, Computational investigation of O_2 diffusion through an intra-molecular tunnel in AlkB. Influence of polarization on O_2 transport, Chem. Sci. 8 (2017) 6230–6238.
[19] A. Nunes-Alves, D.M. Zuckerman, G.M. Arantes, Escape of a small molecule from inside T4 lysozyme by multiple pathways, Biophys. J. 114 (2018) 1058–1066.
[20] F.J. Van Eerden, M.N. Melo, P.W. Frederix, X. Periole, S.J. Marrink, Exchange pathways of plastoquinone and plastoquinol in the photosystem II complex, Nat. Commun. 8 (2017) 1–8.
[21] V. Sharma, G. Belevich, A.P. Gamiz-Hernandez, T. Róg, I. Vattulainen, M.L. Verkhovskaya, M. Wikström, G. Hummer, V.R.I. Kaila, Redox-induced activation of the proton pump in the respiratory complex I, Proc. Natl. Acad. Sci. U. S. A. 112 (2015) 11571–11576.
[22] A.P. Gamiz-Hernandez, A. Jussupow, M.P. Johansson, V.R. Kaila, Terminal electron-proton transfer dynamics in the quinone reduction of respiratory complex I, J. Am. Chem. Soc. 139 (2017) 16282–16288.
[23] A. Di Luca, A.P. Gamiz-Hernandez, V.R.I. Kaila, Symmetry-related proton transfer pathways in respiratory complex I, Proc. Natl. Acad. Sci. U. S. A. 114 (2017) 6314–6321.
[24] J. Warnau, V. Sharma, A.P. Gamiz-Hernandez, A.D. Luca, O. Haapanen, I. Vattulainen, M. Wikstrom, G. Hummer, V.R.I. Kaila, Redox-coupled quinone dynamics in the respiratory complex I, Proc. Natl. Acad. Sci. U. S. A. 115 (2018) E8413–E8420.
[25] V.V. Galassi, G.M. Arantes, Partition, orientation and mobility of ubiquinones in a lipid bilayer, Biochim. Biophys. Acta 1847 (2015) 1345.
[26] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779–815.
[27] H. Angerer, H.R. Nasiri, V. Niedergesäß, S. Kerscher, H. Schwalbe, U. Brandt, Tracing the tail of ubiquinone in mitochondrial complex I, Biochim. Biophys. Acta 1817 (2012) 1776–1784.
[28] H.R. Bridges, E. Bill, J. Hirst, Mossbauer spectroscopy on respiratory complex I: the iron-sulfur cluster ensemble in the NADH-reduced enzyme is partially oxidized, Biochemistry 51 (2012) 149–158.
[29] R. Huey, G.M. Morris, A.J. Olson, D.S. Goodsell, A semiempirical free energy force field with charge-based desolvation, J. Comput. Chem. 28 (2007) 1145–1152.
[30] M. Javanainen, Universal method for embedding proteins into complex lipid bilayers for molecular dynamics simulations, J. Chem. Theory Comput. 10 (2014) 2577–2582.
[31] J.B. Klauda, R.M. Venable, J.A. Freites, J.W. O'Connor, D.J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A.D.M. Jr, R.W. Pastor, Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types, J. Phys. Chem. B 114 (23) (2010) 7830–7843.
[32] J. Huang, A.D. MacKerell Jr, CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data, J. Comput. Chem. 34 (2013) 2135–2145.
[33] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935.
[34] C.H. Chang, K. Kim, Density functional theory calculation of bonding and charge parameters for molecular dynamics studies on [FeFe] hydrogenases, J. Chem. Theory Comput. 5 (2009) 1137–1145.
[35] M. McCullagh, G.A. Voth, Unraveling the role of the protein environment for [FeFe]-hydrogenase: a new application of coarse-graining, J. Phys. Chem. B 117 (2013) 4062–4071.
[36] M.J. Abraham, T. Murtola, R. Schulz, S. Pall, J.C. Smith, B. Hess, E. Lindahl, GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1-2 (2015) 19–25.
[37] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[38] The PyMOL Molecular Graphics System, Version 1.8, Schrödinger, LLC, 2015.
[39] J.D. Hunter, Matplotlib: a 2D graphics environment, Comput. Sci. Eng. 9 (2007) 90–95.
[40] D. Branduardi, F.L. Gervasio, M. Parrinello, From A to B in free energy space, J. Chem. Phys. 126 (2007) 054103.
[41] G.A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, G. Bussi, Plumed 2: new feathers for an old bird, Comput. Phys. Commun. 185 (2014) 604–613.
[42] B. Roux, The calculation of the potential of mean force using computer simulations, Comput. Phys. Commun. 91 (1995) 275–282.
[43] M.J. Field, The pDynamo Program for molecular simulations using hybrid quantum chemical and molecular mechanical potentials, J. Chem. Theory Comput. 4 (2008) 1151–1161.
[44] R.W. Johnson, An introduction to the bootstrap, Teach. Stat. 23 (2001) 49–54.
[45] R. Fato, E. Estornell, S. Di Bernardo, F. Pallotti, G.P. Castelli, G. Lenaz, Steady-state kinetics of the reduction of coenzyme Q analogs by complex I (NADH:ubiquinone oxidoreductase) in bovine heart mitochondria and submitochondrial particles, Biochemistry 35 (1996) 2705–2716.
[46] V. Zickermann, B. Barquera, M. Wikström, M. Finel, Analysis of the pathogenic human mitochondrial mutation ND1/3460, and mutations of strictly conserved residues in its vicinity, using the bacterium Paracoccus denitrificans, Biochemistry 37 (1998) 11792–11796.
[47] K. Sakamoto, H. Miyoshi, M. Ohshima, K. Kuwabara, K. Kano, T. Akagi, T. Mogi, H. Iwamura, Role of the isoprenyl tail of ubiquinone in reaction with respiratory enzymes: studies with bovine heart mitochondrial complex I and *Escherichia coli* bo-type ubiquinol oxidase, Biochemistry 37 (1998) 15106–15113.
[48] P.W. Snyder, J. Mecinovic, D.T. Moustakas, S.W. Thomas III, M. Harder, E.T. Mack, M.R. Lockett, A. Héroux, W. Sherman, G.M. Whitesides, Mechanism of the hydrophobic effect in the biomolecular recognition of arylsulfonamides by carbonic anhydrase, Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 17889–17894.
[49] H. Yao, H. Ke, X. Zhang, S.-J. Pan, M.-S. Li, L.-P. Yang, G. Schreckenbach, W. Jiang, Molecular recognition of hydrophilic molecules in water by combining the hydrophobic effect with hydrogen bonding, J. Am. Chem. Soc. 140 (2018) 13466–13477.
[50] J.G. Fedor, A.J.Y. Jones, A.D. Luca, V.R.I. Kaila, J. Hirst, Correlating kinetic and structural data on ubiquinone binding and reduction by respiratory complex I, Proc. Natl. Acad. Sci. U. S. A. 114 (2017) 12737–12742.
[51] L. Zhang, J. Hermans, Hydrophilicity of cavities in proteins, Proteins Struct. Funct. Genet. 24 (1996) 433–438.
[52] D.G. Isom, C.A. Castañeda, B.R. Cannon, P.D. Velu, B. García-Moreno E. , Charges in the hydrophobic interior of proteins, Proc. Natl. Acad. Sci. U. S. A. 107 (37) (2010) 16096–16100.
[53] C.C. Moser, J.M. Keske, K. Warncke, R.S. Farid, P.L. Dutton, Nature of biological electron transfer, Nature 355 (1992) 796–802.
[54] T. Ohnishi, Iron-sulfur clusters & semiquinones in Complex I, Biochim. Biophys. Acta 1364 (1998) 186–206.
[55] T. Ohnishi, V.D. Sled, T. Yano, T. Yagi, D.S. Burbaev, A.D. Vinogradov, Structure-function studies of iron-sulfur clusters and semiquinones in the NADH-Q oxidoreductase segment of the respiratory chain, Biochim. Biophys. Acta 1365 (1998) 301–308.
[56] T. Yano, W.R. Dunham, T. Ohnishi, Characterization of the $\Delta\mu_{H^+}$-sensitive ubisemiquinone species (SQ_{Nf}) and the interaction with cluster N2: new insight into the energy-coupled electron transfer in complex I, Biochemistry 44 (2005) 1744–1754.
[57] J. Hirst, M.M. Roessler, Energy conversion, redox catalysis and generation of reactive oxygen species by respiratory complex I, Biochim. Biophys. Acta 1857 (2016) 872–883.

[58] S. Kurki, V. Zickermann, M. Kervinen, I. Hassinen, M. Finel, Mutagenesis of three conserved Glu residues in a bacterial homologue of the ND1 subunit of complex I affects ubiquinone reduction kinetics but not inhibition by dicyclohexylcarbodiimide, Biochemistry 39 (2000) 13496–13502.
[59] A. Fersht, Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding, 1st ed., W. H. Freeman, New York, 1999.
[60] M.A. Tocilescu, U. Fendel, K. Zwicker, S. Drose, S. Kerscher, U. Brandt, The role of a conserved tyrosine in the 49-kDa subunit of complex I for ubiquinone binding and reduction, Biochim. Biophys. Acta 1797 (2010) 625–632.
[61] T.G. Pomorski, A.K. Menon, Lipid somersaults: uncovering the mechanisms of protein-mediated lipid flipping, Prog. Lipid Res. 64 (2016) 69–84.
[62] M. Furuhashi, G.S. Hotamisligil, Fatty acid-binding proteins: role in metabolic diseases and potential as drug targets, Nat. Rev. Drug Discov. 7 (2008) 489–503.
[63] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, J. Chem. Phys. 126 (2007) 014101.
[64] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.
[65] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: a linear constraint solver for molecular simulations, J. Comput. Chem. 18 (1997) 1463–1472.
[66] H.-J. Woo, B. Roux, Calculation of absolute protein-ligand binding free energy from computer simulations, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 6825–6830.
[67] J.N. Blaza, K.R. Vinothkumar, J. Hirst, Structure of the deactive state of mammalian respiratory complex I, Structure 26 (2018) 312–319.
[68] H. Ashkenazy, S. Abadi, E. Martz, O. Chay, I. Mayrose, T. Pupko, N. Ben-Tal, ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules, Nucl. Acids Res. 53 (2016) 199–206.
[69] L. van Hove, Correlations in space and time and born approximation scattering in systems of interacting particles, Phys. Rev. 95 (1) (1954) 249–3720.
[70] M. Allen, D. Tildesley, Computer Simulation of Liquids, 1st ed., Oxford University Press, New York, 1987.
[71] M.H. Teixeira, G.M. Arantes, Effects of lipid composition on membrane distribution and permeability of natural quinones, RSC Adv. 9 (2019) 16892–16899.