**Technische Universität München**
Fakultät für Chemie
Lehrstuhl für Computergestützte Biokatalyse

# Large-scale coupling principles in the proton pumping of respiratory complex I

**Andrea Di Luca**

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines

**Doktors der Naturwissenschaften (Dr. rer. nat.)**

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Franz Hagn

Prüfer der Dissertation:

1. Prof. Dr. Ville Kaila
2. Priv.-Doz. Dr. Harald Oberhofer
3. Prof. Dr. Thorsten Friedrich

Die Dissertation wurde am 11.06.2019 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 03.07.2019 angenommen.

# Contents

# Abstract

Aerobic respiration powers the cell by using energy released from the reduction of $O_2$ to water. The process is catalyzed by the Electron Transport Chain (ETC), consisting of a series of enzymes that are shared across all domains of life. Complex I (NADH:ubiquinone oxidoreductase) is the largest and least understood enzyme of the ETC, and it transfers electrons from nicotinamide adenine dinucleotide (NADH) to ubiquinone (Q). The released free energy is used to generate an electrochemical proton gradient across a membrane, which in turn fuels the synthesis of adenosine triphosphate (ATP) and active transport.

In this doctoral thesis, we investigate different aspects of complex I related to proton pumping, Q binding, and large-scale dynamics. Our results show how the signal for the proton pumping can be transferred across long distances, and how the Q motion in its binding pocket can be linked to complex I function. Additionally, we show how accessory subunits modulate dynamics of complex I, and that important regulatory transitions are linked to global motions of the enzyme. Together, these results provide an enhanced molecular understanding of the mechanism of complex I.

# Zusammenfassung

Aerobe Atmung versorgt die Zelle mit Energie, die bei der Reduktion von molekularem Sauerstoff zu Wasser frei wird. Die Elektronentransportkette katalysiert diesen Prozess und besteht aus einer Reihe von Enzymen, die in jedem lebenden Organismus vorkommen. Komplex I (NADH:Coenzym Q Oxidoreductase) ist das größte und am wenigsten verstandene Enzym der Elektronentransportkette. Es transferiert Elektronen von Nicotinamidadenindinukleotid (NADH) zu Ubichinon (Q). Die dabei freiwerdende Energie wird zum Aufbau eines elektrochemischen Protonengradientens an einer Membran genutzt, welches wiederum zur Synthese von Adenosintriphosphat und dem Betreiben aktiver Transportvorgänge dient.

In dieser Dissertation untersuchen wir verschiedene Aspekte von Komplex I, die mit der Protonenpumpfunktion zusammenhängen. Unsere Untersuchungen legen dar, wie die Signaltransduktion über lange Distanzen erfolgen kann und wie die Ubichinonbewegung in der Bindetasche zur Funktion von Komplex I beiträgt. Weiterhin zeigen wir, wie die zusätzlichen Untereinheiten die Dynamik von Komplex I modulieren, und dass wichtige regulatorische Übergänge an die globalen Bewegungen des Enzyms gekoppelt sind. Zusammengenommen, liefern diese Ergebnisse ein besseres, molekulares Verständnis des Mechanismus von Komplex I.

# List of Publications

I. **Di Luca A**, Gamiz-Hernandez AP, Kaila VRI (2017) Symmetry related proton transfer pathways in respiratory complex I. *Proc Natl Acad Sci USA* 114:E6314-E6321

II. Fedor JG, Jones AJY, **Di Luca A**, Kaila VRI, Hirst J (2017) Correlating kinetic and structural data on ubiquinone binding and reduction by respiratory complex I. *Proc Natl Acad Sci USA* 114:12737-12742

III. **Di Luca A***, Mühlbauer M*, Saura P, Kaila VRI (2018) How inter-subunit contacts in the membrane domain ofcomplex I affect proton transfer energetics. *Biochim Biophys Acta Bioenerg* 1859:734-741

IV. **Di Luca A**, Kaila VRI (2018) Global collective motions in the mammalian and bacterial respiratory complex I. *Biochim Biophys Acta Bioenerg* 1859:326-332

V. Warnau J, Sharma V, Gamiz-Hernandez AP, **Di Luca A**, Haapanen O, Vattulainen I, Wikström M, Hummer G, Kaila VRI (2018) Redox-coupled quinone dynamics in the respiratory complex I. *Proc Natl Acad Sci USA* 115:E8413-E8420

VI. Jussupow A*, **Di Luca A***, Kaila VRI (2019) How cardiolipin modulates the structure and dynamics of respiratory complex I. *Sci Adv* 5:eaav1850

*Authors contributed equally

## Other publications:

VII. **Di Luca A**, Kaila VRI. Molecular strain in the active/deactive transition of the mammalian respiratory complex I. (*In preparation*).

# Author Contributions

I. ADL, APGH, and VRIK designed research. ADL, APGH, and VRIK performed research. ADL, APGH and VRIK analyzed data. ADL and VRIK wrote the paper.

II. JGF and JH designed research. JGF, AJYJ, ADL, and VRIK performed research. JGF, ADL, VRIK, and JH analyzed data. JGF and JH wrote the paper.

III. ADL, MEM, and VRIK designed research. ADL, MEM, PS, and VRIK performed research. ADL, MEM, and PS analyzed data. ADL, MEM, PS, and VRIK wrote the paper.

IV. ADL and VRIK designed research. ADL performed research. ADL and VRIK analyzed data. ADL and VRIK wrote the paper.

V. JW, VS, and VRIK designed research. JW, VS, APGH, ADL, OH, and VRIK performed research. JW, VS, APGH, ADL, OH, IV, MW, GH, and VRIK analyzed data. JW, VS, GH, and VRIK wrote the paper.

VI. AJ, ADL, and VRIK designed research. AJ and ADL performed research. AJ, ADL, and VRIK analyzed data. AJ, ADL, and VRIK wrote the paper.

# Abbreviations

| | |
|---|---|
| ANM | Anisotropic Network Model |
| AOX | Alternative Oxidase |
| ADP | Adenosine Diphosphate |
| ATP | Adenosine Triphosphate |
| CDL | Cardiolipin |
| CG | Coarse Grained |
| Cyt *c* | Cytochrome *c* |
| DFT | Density Functional Theory |
| EM | Electron Microscopy |
| eT | Electron Transfer |
| ETC | Electron Transport Chain |
| FMN | Flavin mononucleotide |
| GB | Generalized Born |
| GGA | Generalized Gradient Approximation |
| GNM | Gaussian Network Model |
| HF | Hartree-Fock |
| ISC | Iron-Sulfur Center |
| MC | Monte Carlo |
| MD | Molecular Dynamics |
| MM | Molecular Mechanics |
| NADH | Nicotineamide Adenine Dinucleotide |
| NMA | Normal Mode Analysis |
| PB | Poisson-Boltzmann |
| PBC | Periodic Boundary Conditions |
| PCA | Principal Component Analysis |
| PCET | Proton-Coupled Electron Transfer |
| PDB | Protein Data Bank |
| POPC | 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine |
| POPE | 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine |
| PME | Particle Mesh Ewald |
| $Q/QH_2$ | Ubiquinone/Ubiquinol |
| QM | Quantum Mechanical, Quantum Mechanics |
| QM/MM | Quantum Mechanical/ Molecular Mechanics |
| REUS | Replica Exchange Umbrella Sampling |

| | |
|---|---|
| RMSF | Root Mean Square Fluctuation |
| ROS | Reactive Oxygen Species |
| TM | Transmembrane |
| VdW | Van der Waals |

# Acknowledgements

My special thanks are for Ville, for his constant presence and support. I gave you some headaches (as you said, I am sometimes *bohemian*...), but thank you for always having the time to hear me out.

I would like to thank all my colleagues that made the Ph.D. journey a wonderful experience. Thanks to Shreyas Supekar, Sophie Mader, Ana Gamiz-Hernandez, Qi Luo, Alexander Jussupow, Mona Baumgart, Max Mühlbauer, my officemate Michael Röpke, Daniel Riepl, Patricia Saura, Carl-Mikael Suomivuori, Tapio Salminen, Mikko Muuronen, Ina Bisha, Sven Klumpe, Romina Wild and please -the order does not matter-, you are way too many! I also thank Justin Fedor, Andrew Y. Jones, and Judy Hirst for the great work together.

I have the feeling that I wouldn't be writing this thesis today, if it wasn't for many people that I met in during my life. I would like to thank everyone, but I cannot list you all! I hope a big hug will suffice.

During the years I met some amazing people around Munich and Freising. I just want to say that these were some of the most beautiful years of my life, and you all played an important role in it.

Additionally, a little while after moving here, a special person joined me and made my life so much better. Thank you Valentina, I feel that without you things wouldn't have been this special.

Last, but not the least, those that have always been supporting for me from far away: I want to thank my family, my mom Stefania, my dad Livio, my brother Ilario. I am a bad son, brother and grandchild, but I want to tell you that I am always thinking about you all.

This work was financially supported by DFG, and the CPU time was provided by the SuperMuc at the Leibniz Rechenzentrum (grant: pr48de and pr27xu).

So long, and thanks for all the CPU time!

Andrea

*I can't promise I'll be the best,*
*but I'll be the best that I can be.*

# Chapter 1

# Introduction

The evolution of first organisms took place over a billion years ago [1, 2]. Several theories have attempted to explain the complex series of events that generated life as we know it today [3], but they all share the same basis: building of living systems required a high amount of energy [4]. All living organisms rely on extraction of energy from external sources to fuel internal processes. Starting from proto-enzymes, nature has developed convoluted ways to release, store, and utilize energy, while maintaining high efficiencies in every process. After development of photosynthesis in cyanobacteria and the later appearance of diatomic oxygen ($O_2$) in the atmosphere, also called the great oxygenation event [5], aerobic organisms flourished and respiratory machineries started to evolve.

Aerobic respiration extracts electrons from organic molecules that contain carbon in a low oxidation state, and transfers them to $O_2$ producing water [6]. Oxygen is a strong oxidant, and its reduction produces large quantities of energy that organisms can use to drive endergonic reactions. To achieve high efficiencies, electrons are transferred stepwise, forming and consuming various intermediates. The last steps of this energy-conversion machinery are performed by membrane-bound enzymes, the *Electron Transport Chain* (ETC, or respiratory chain), and commonly referred to as *Oxidative Phosphorylation*. The free energy released during substrate degradation is then used for active transport and to synthesize adenosine triphosphate (ATP), the final high-energy product that cells use to perform several life-sustaining processes [7, 8].

During the second half of the $20^{th}$ century, a tremendous amount of work performed by several scientists started to unravel the principles behind the function of the ETC [9]. One of the biggest open questions during that time was related to the production of ATP. Thinking about a 'classical' chemical coupling, *i.e.* exchange of high-energy bonds, scientist were looking for an intermediate substrate of the ETC that could fuel ATP synthase [10]. Due to a lack of evidence of such an intermediate, a new hypothesis was formulated, suggesting that energy in cells can also be stored as an electrochemical potential across a semi-permeable membrane [7, 11]. This hypothesis, developed in large parts by Peter Mitchell, explained principles of energy converting processes in living organisms. Mitchell was awarded the Nobel Prize in 1978 "for his contribution to the understanding of biological energy transfer through the formulation of the *chemiosmotic*

*theory*". The key idea behind the chemiosmotic theory is that energy required for the synthesis of ATP and active transport is stored in form of an electrochemical proton gradient, the *proton motive force* or *pmf*:

$$pmf = \Delta\psi - 2.303\frac{RT}{F}\Delta\text{pH} \tag{1.1}$$

where $\Delta\psi$ is the electrical potential difference, and $\Delta$pH is the difference of the chemical activity of protons between the two sides of the membrane. The *pmf* is generated and maintained by respiratory enzymes using the free-energy derived from catalysis of redox reactions. This energy-converting mechanism is ubiquitous across all domains of life, suggesting that its origin traces back to the early stage of life evolution [12].

## 1.1 The Electron Transport Chain

The general mitochondrial and prokaryotic ETC scheme (Figure 1.1) comprises five membrane-bound enzymes, which catalyze the redox reactions used to generate the *pmf* (Complex I-IV) and synthesize ATP (Complex V). The overall reaction involves a transfer of electrons from nicotinamide adenine dinucleotide (NADH) or succinate to $O_2$, and coupling of the released energy to the synthesis of ATP from ADP and inorganic phosphate ($P_i$) [13]. The intermediate species generated during this metabolic pathways have

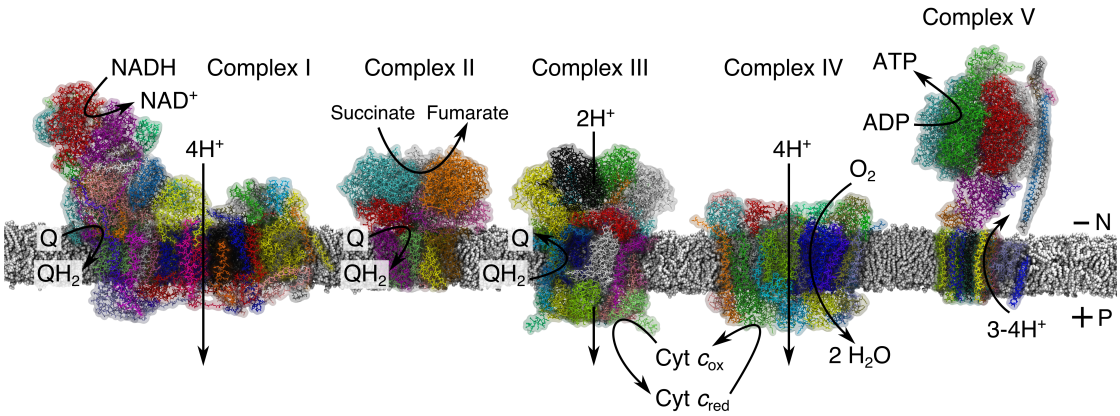


**Figure 1.1:** Enzymes of the respiratory chain. Complexes I-V (left to right) are embedded in the membrane. Protons are pumped from the negative (-, N) to the positive (+, P) side of the membrane against an electrochemical proton gradient. Complex V uses the *pmf* to drive ATP synthesis.

an increasing midpoint potential ($E_m$), ranging from -320 mV for NADH to +820 mV of $O_2$ at pH 7 [14].

The initial electron acceptors, complex I (NADH:ubiquinone oxidoreductase) and complex II (succinate:ubiquinone oxidoreductase), harvest high-energy electrons from NADH and succinate, respectively, which are the final products of other catabolic cellular processes. Complex I is generally considered the primary entry point for electrons in the

respiratory chain [15]. The reaction catalyzed by the enzyme can be written as:

$$\mathrm{NADH} + \mathrm{H^+} + \mathrm{Q} + 4\,\mathrm{H^+_N} \longrightarrow \mathrm{NAD^+} + \mathrm{QH_2} + 4\,\mathrm{H^+_P} \tag{1.2}$$

Complex I function and mechanism will be discussed in detail in Chapter 2. Ubiquinone (Q), or Coenzyme $Q_{10}$ (Figure 1.2), is a two-electron carrier, and part of the ETC. This membrane-soluble molecule is composed of a benzoquinone headgroup and a long hydrophobic tail of 10 isoprenoid units. Ubiquinone cofactors and its reduced form ubiquinol ($QH_2$) are ubiquitous in nature, and they are also synthetized as menaquinone or plastoquinone in other organisms [16, 17].

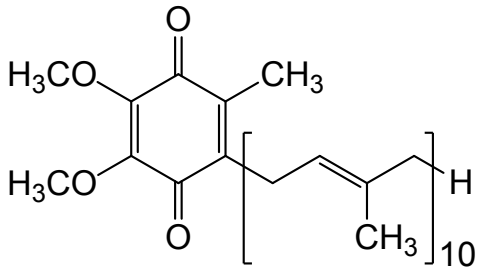


**Figure 1.2:** Chemical structure of ubiquinone ($Q_{10}$). The length of the isoprenoid tail can vary.

Complex II provides an alternative entry for electrons in the respiratory chain [18, 19], and catalyses the reaction:

$$\mathrm{Succinate} + \mathrm{Q} \longrightarrow \mathrm{Fumarate} + \mathrm{QH_2} \tag{1.3}$$

This enzyme is not electrogenic, and it only contributes to the reduction of the Q-pool. The oxidation of succinate to fumarate is also considered part of the citric acid cycle. The Q-pool reduced by complex I and II is re-oxidized by complex III (also cytochrome $bc_1$ or quinone:cytochrome *c* oxidoreductase). Formally, this enzyme uses the energy derived from oxidation of a $QH_2$ molecule to reduce two cytochrome *c* (cyt *c*) molecules, and release two protons across the membrane [20]. The net reaction is:

$$\mathrm{QH_2} + 2\,\mathrm{cyt}\ c_{\mathrm{ox}} \longrightarrow \mathrm{Q} + 2\,\mathrm{cyt}\ c_{\mathrm{red}} + 2\,\mathrm{H^+_P} \tag{1.4}$$

Cyt *c* is a small, *ca.* 10 kDa, water soluble protein that contains a covalently bound heme *c*, which acts as a one-electron carrier [21]. The reactions by which electrons are transferred from Q to cyt *c* are known as Q-cycle [22]. The complete reaction requires oxidation of two $QH_2$ and reduction of one Q, with electrons transferred across *b*- and *c*-type hemes and Rieske Iron-Sulfur Centers (ISC) [23]. Unlike complex I, the electrogenic protons are not pumped across the membrane: the transferred protons are part of a redox-loop and derived from the reduction/oxidation of Q and $QH_2$.

The last step of the electron transfer process is catalyzed by complex IV (cytochrome

*c* oxidase), which reduces the final electron acceptor $O_2$ [14, 24]:

$$4\,\text{cyt}\,c_{\text{red}} + O_2 + 8\,H_N^+ \longrightarrow 4\,\text{cyt}\,c_{\text{ox}} + 2\,H_2O + 4\,H_P^+ \tag{1.5}$$

The reaction is performed stepwise transferring one electron at a time via a bimetallic copper-center and an *a*-type heme [14, 25]. The total cycle of the enzyme pumps four protons per reduced $O_2$ molecule and also consumes four protons from the positive side of the membrane to form two water molecules.

Finally, complex V, or $F_0F_1$-ATP synthase, catalyzes the endergonic synthesis of ATP using the *pmf* produced by complex I, complex III and complex IV. The overall reaction is:

$$ADP + P_i + 3\text{-}4H_P^+ \longrightarrow ATP + H_2O + 3\text{-}4H_N^+ \tag{1.6}$$

where $P_i$ indicates inorganic phosphate. The synthesis occurs via a rotary-mechanism proposed by P. Boyer [26,27] and later confirmed by J. E. Walker and coworkers [28,29].

Over the course of the last 70 years, research on the ETC unraveled many of the functions of these membrane enzymes. Nonetheless, several questions still remain unanswered and new aspects keep emerging, as for example the role of supercomplexes [30–32], weakly associated units formed by respiratory enzymes. Structures of supercomplexes have recently been resolved to a near-atomic resolution [33–35], but their role is not yet fully understood. Similarly, complex I structure has only been resolved in 2013 [36], and a clear picture of its coupling mechanism is still lacking.

In some cases, energy-converting enzymes make use of exotic chemical species, as the manganese cluster in the photosynthetic enzymes [37, 38], or ferryl species in Complex IV [39]. Despite these exceptions, most of the processes occurring during function involve 'simple' electron and proton transfers or are linked to structural changes. It is important to notice that the extraordinary aspect of these complex machineries lies in the tight orchestration of all these reactions, which regulate the energetic balance in the cell through complex networks of signal transport. This coupling is internal to enzymes, but also extends beyond the single redox-coupled reactions, and regulates the entire ETC.

The high amount of energy generated from the ETC can be explained by thermodynamic considerations: the individual reactions do not involve large gaps in free energy transfers, leading to near-equilibrium processes. The reactions are close to reversibility, and therefore highly efficient.

Rationalization of the biochemical data in light of the thermodynamic principles is central to understanding how these enzymes work. As for any other system, laws of thermodynamics also regulate the function of cells. Explaining how, and why the processes take place on a molecular level is the key role of theoretical and computational chemistry.

## 1.2 The Role of Computational Chemistry

Theoretical studies allow molecular insight into the structure and function of systems of interest, and provide information that could be hard to obtain in 'wet lab' experiments. Theory can confirm experimental findings, but also gives feedback and suggests new experiments. The development of a theoretical framework to understand catalysis, *e.g.* in enzymes, not only provides the tools to make predictions, but also to design *de novo* catalytic systems [40].

Driven by increasing computational power, theoretical biochemistry and bioenergetics has advanced from early studies related mostly to thermodynamical considerations, and is now strongly based on computational approaches. In the early 1960s, the steadily improving computational resources led to the development of the first classical force fields for the treatment of organic molecules, discussed in the pioneering work of Lifson and Warshel [41]. Parallel to the classical methods, quantum chemistry had a rapid development, reaching accuracies high enough to describe reactions in active sites of enzymes. Among several methods, Density Functional Theory (DFT) has become very popular in the last 30 years, due to the optimal combination of accuracy and speed of calculation.

A plethora of methods is available nowadays to study biomolecules. Computational chemistry is a well-established tool, providing a unique approach to analyze single reaction events or protein dynamics. In Chapter 3 some details of computational approaches used in this thesis will be presented.

# Chapter 2

# The Respiratory Complex I

Complex I (NADH:ubiquinone oxidoreductase) is the largest and most challenging enzyme of the ETC. It is composed of up to 45 subunits in eukaryotes [42], with a molecular weight of *ca.* 1 MDa. Complex I is a redox-driven proton pump that catalyzes the eT between nicotinamide adenine dinucleotide (NADH) and quinone (Q), and uses the released free energy to pump four protons across the membrane (Figure 2.1) [15, 43, 44]. The overall reaction catalyzed by the enzyme is:

$$\mathrm{NADH} + \mathrm{H^+} + \mathrm{Q} + 4\,\mathrm{H^+_N} \longrightarrow \mathrm{NAD^+} + \mathrm{QH_2} + 4\,\mathrm{H^+_P} \tag{2.1}$$

where N and P denote the negative and positive side of the membrane, respectively. The catalytic core of the enzyme is composed of 14 subunits, which are conserved from bacteria to mammals [42]. The enzyme can be roughly divided into hydrophilic and membrane domains, where the eT and proton transport take place, respectively. An alternative classification that is based on evolutionary models divides complex I into the N, Q, and P modules [43, 45–47]. Throughout this thesis, which is mainly focused on the enzyme function, the distinction is made between hydrophilic and membrane domains, which also naturally separates the eT and proton transport processes. In this context, the Q binding pocket can be considered as an interface region that couples the two processes together. The eT and proton transport also occur across multiple subunits of the enzyme, creating a coupled machinery that extends for over 300 Å.

Despite over 70 years of research, complex I is still the least understood enzyme of the ETC. Although the different domains of the complex were separately crystallyzed [48, 49], and the overall organization of the complex was already known [50–52], the first near-atomic resolution crystal structure of the entire enzyme from the bacteria *Thermus thermophilus* was only resolved in 2013 by Baradaran *et al.* [36]. The study showed for the first time the atomistic structure of the entire Q binding pocket, central for the understanding the catalytic mechanism of the enzyme. A second near-atomic structure was obtained in 2015 by Zickermann and co-workers [53] from the yeast *Yarrowia lipolytica*, showing the organization of the additional supernumerary subunits around the catalytic core. The assignment of the novel subunits could not be completed due to the low resolution of the density map in the peripheral regions of the enzyme. In 2016, structures

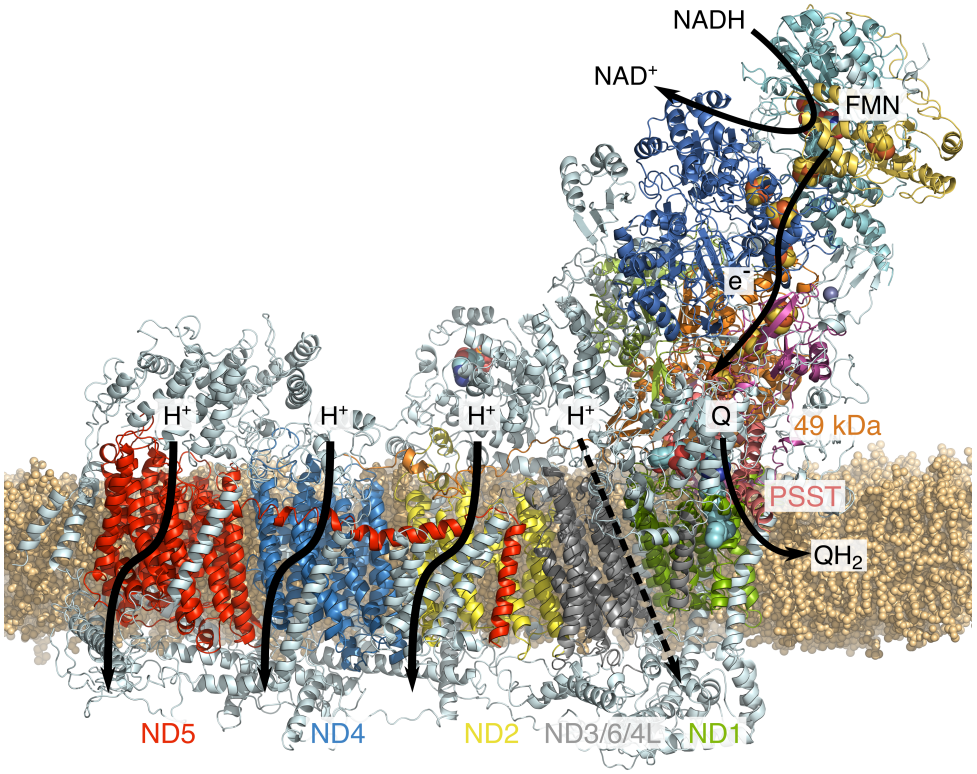


**Figure 2.1:** Entire structure of the respiratory complex I from *Mus musculus* (PDB ID:6G2J). Supernumerary subunits are colored in light blue. The electron and proton transport processes are indicated with black arrows.

from ovine (*Ovis aries*) and bovine (*Bos taurus*) complex I were obtained to a *ca.* 4 Å resolution [54,55], confirming and nearly completing the assignment of the 31 supernumerary subunits of mammals. Recently, structures from mouse (*Mus musculus*) [56, 57], porcine (*Sus scrofa*) [33,34], and a new structure of *Yarrowia lipolytica* [58], further improved the resolution, allowing deeper insight into the different states and function of the enzyme [59].

Due to its function and shape, it is natural to present the enzyme by following the 'charge motion' during turnover. The sections of this chapter will thus first describe the NADH binding site and the hydrophilic domain, where the reaction starts, followed by discussion about the Q binding site, the membrane domain, and supernumerary subunits.

## 2.1 Nomenclature

The nomenclature of the subunits of complex I differs among organisms. In this chapter, the description will follow the nomenclature of the bovine enzyme (unless specified). For consistency with the related publications, the results chapter will employ nomenclature of the respective isoform, and the corresponding subunits will be listed where needed for comparison. The nomenclature of the core subunits in different species is given in

Appendix A.

## 2.2 The Hydrophilic Domain

The hydrophilic domain of complex I comprises seven of the catalytic core subunits and is mainly devoted to the eT functions [15, 48, 60–64]. The electrons are transferred from NADH to Q along a chain of cofactors composed of a flavin mononucleotide (FMN) molecule and a series of iron-sulfur centers (ISC, Figure 2.2), which act as a connecting wire between the two distal reaction sites. The eT process through the ISC chain requires about 100 µs to occur [63]. The NADH binding site is located at the 'top-edge' of this region (Figure 2.1) and comprises the FMN molecule, that acts as the first electron acceptor of the chain [65]. The binding pocket is solvent-accessible and is formed by an unusual Rossmann-fold, which comprises only four $\beta$-strands instead of the more common six $\beta$-strands motif. In this cavity, the reaction is initiated by a proton-coupled electron transfer (PCET) from NADH to FMN. The reaction most likely involves formation of transient reduced FMN species, but the exact reaction steps are not clear, and only recently principles have started to emerge [66, 67]. The nearby ISC N1a is located off the main ISC pathway connecting FMN to Q (Figure 2.2), and its function is still under debate. Its distance from the NADH oxidation site suggests an active role in the eT mechanism, but $E_m$ values lower than -320 mV were indirectly derived from EPR signal in *Escherichia coli* [68], excluding the reduction by NADH.

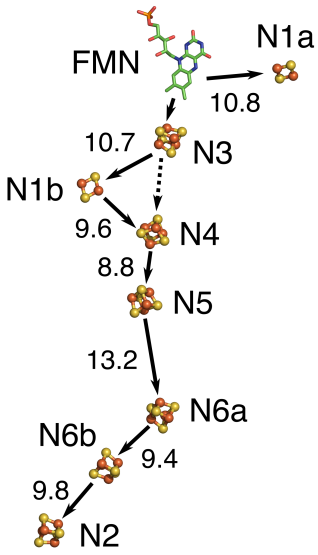


**Figure 2.2:** Chain of redox cofactors and their relative edge-to-edge distances in the hydrophilic domain of respiratory complex I (PDB ID:6G2J from *Mus musculus*)

The other ISC form a linear chain connecting FMN to Q, and their edge-to-edge distances (shortest distance between atoms) are lower than 14 Å (Figure 2.2), within the biological eT limit [69]. An exception is the ISC N7, only present in some organisms, *e.g.* *Thermus thermophilus*, and replaced by a zinc site in higher organisms such as mammals. This ISC is not involved in the eT and its role is likely to be related to the scaffold stability [48, 70]. The last ISC cluster, N2, is located *ca.* 12 Å from the quinone headgroup [36, 71], and has the lowest reduction potential of the ISC chain [72–74]. Although it was suggested to be a potential coupling site for the enzyme catalysis [75, 76], recent experiments [77] dispute this hypothesis.

The ISCs are one-electron carriers, while the reduction and oxidation of Q and NADH require two electrons. This implies that the complete Q reduction requires two eTs along the ISC chain. Hypotheses regarding the eT process suggest that, during catalysis, part of the ISC are reduced [15, 78], possibly lowering the time between subsequent reduction of Q.

The entire catalytic process is also reversible [79–81] and complex I can promote reverse eT from $QH_2$ to FMN using the *pmf*. Under these conditions, the enzyme produces a high amount of reactive oxygen species (ROS), which are harmful for the cell [80, 82, 83]. This process is important in cell signalling and during hypoxia conditions, and is therefore of extreme biomedical relevance [84–87].

The NADH oxidation mechanism, the structure of the ISC and their role in the charge transport process will be not discussed further. In this thesis, it is sufficient to assume that the two electrons are transferred stepwise to the Q binding site on a shorter time scale than the one required for the proton pumping. For a more in-depth discussion see for example ref. [15].

## 2.3 The Quinone Binding Pocket

The Q binding pocket is located at the interface between the hydrophilic and the membrane domain (Figure 2.1), in a cavity formed by subunits ND1, PSST and 49 kDa (Figure 2.3) [36]. This narrow channel is *ca.* 40 Å long, and accommodates nearly the entire Q molecule. Additionally, except for conformational changes related to enzyme function, the channel is overall similar among all complex I structures so far characterized. It is important to notice that, although several biochemical [71, 88–90] and structural studies [36, 52] suggest that the pocket comprises the Q binding site, the substrate has not yet been structurally resolved in a bound state.

The Q polar headgroup binds at the top of the pocket, *ca.* 30 Å above the membrane plane and between subunits PSST and 49 kDa (Figure 2.3). Biochemical studies [91] suggest that the Q headgroup is positioned between conserved residues Tyr108 and the His59/Asp160 pair of 49 kDa subunit. Several other conserved residues might play a role in the binding, *e.g.* Thr156 in the same subunit, but their exact function remained elusive so far. The tyrosine and histidine residues have also been suggested to be the proton donors for the Q upon its reduction to quinol [74, 92]. Recently, structures resolved by different groups [54,55,57,58] observed different conformations of the $\beta$1-$\beta$2 loop of the 49 kDa subunit, which constitutes the surface of cavity and comprises His59. In absence of Q, three different situations were observed: the loop closes the binding site, moving His59 close to the Tyr108 (*Ovis aries* structure), the loop becomes flexible (non-solved, *Bos taurus* structure) or remains essentially unchanged (*Yarrowia lipolytica* structure). This effect might also be linked to the active-deactive transition in some organisms (see Section 2.7) [56], and has not yet been fully clarified.

In addition to the primary Q binding site, EPR studies [93, 94] suggested the existence of a second Q binding site. Recently, computational work also part of this thesis (publication V) provided additional insight into the nature of this putative site in the binding cavity. Further discussion will follow in Chapter 5. From the binding cavity, a series of electrostatic/conformational changes transmit the proton pumping signal to the membrane domain [92, 95] (Section 2.4).

The unusually long Q-cavity raised several questions about the coupling mechanism of complex I, *e.g.* why the binding pocket is buried so deep in the enzyme. The entrance of the cavity is located below the membrane surface, and Q could easily accept electrons

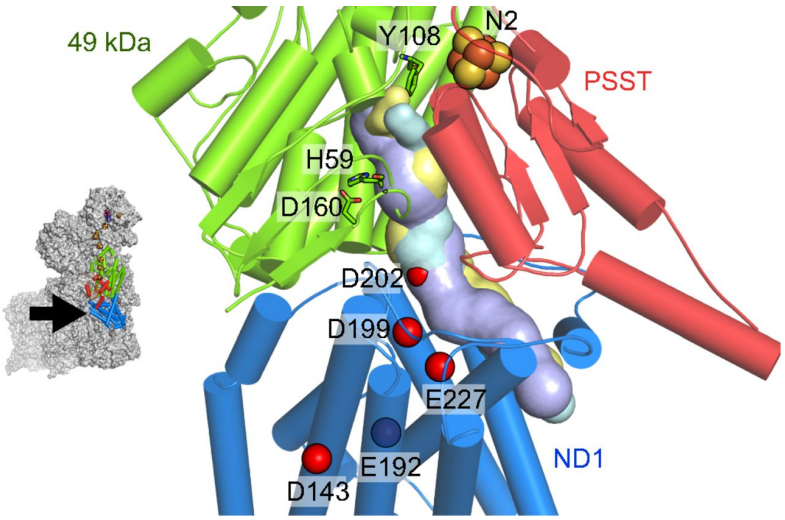


**Figure 2.3:** The quinone binding channel in the mammalian complex I (PDB ID: 5LC5). The subunits forming the quinone binding pocket are represented in cartoon. Surface of the cavity is obtained from bovine (cyan, *Bos tarurus*, [54]), porcine (yellow, *Sus scrofa*, [33]) and bacterial (purple, *Thermus thermophilus*, [36]) structures after proper alignment. The quinone polar headgroup is suggested to bind between Y108 and H59/D160. Other negatively charged residues in the pocket are represented as red/blue spheres. The figure is adapted from publication II.

in a binding site closer to the protein surface. Instead, the enzyme evolved maintaining this long cavity that can nearly fit the entire hydrophobic tail. This question is also directly related to the nature of the coupling between the two domains in complex I: how is the energy transferred during Q turnover? Although the topic is still under debate, recent studies [73] have made significant advances, combining structural, biochemical and computational efforts.

## 2.4 The Membrane Domain

The transmembrane domain of complex I is composed in its minimal assembly by seven linearly arranged subunits (ND1/3/6/4L/2/4/5 and Nqo8/7/10/11/14/13/12 in *Thermus thermophilus*) (Figure 2.4). This domain utilizes the energy derived from Q reduction to pump four protons across the membrane [96, 97] (current consensus value, although a value of three pumped protons has been proposed based on thermodynamics considerations [98]). Subunit ND1 forms the link with the hydrophilic domain, and is also the most conserved of the whole complex [36]. Subunits ND2, ND4, and ND5 are homologous to each other and to multiresistance and pH adaptation (Mrp) $Na^+$-$H^+$ antiporters [99, 100], and are involved in the proton pumping process [95, 101]. These subunits are also referred to as *antiporter-like* subunits. The location of the fourth proton channel is not yet clear, although it is believed to be located in the region between ND4L and ND1 (Nqo11/Nqo8 in *Thermus thermophilus*) [36, 53, 102].

Starting from the Q binding pocket in the ND1 subunit, an array of charged and

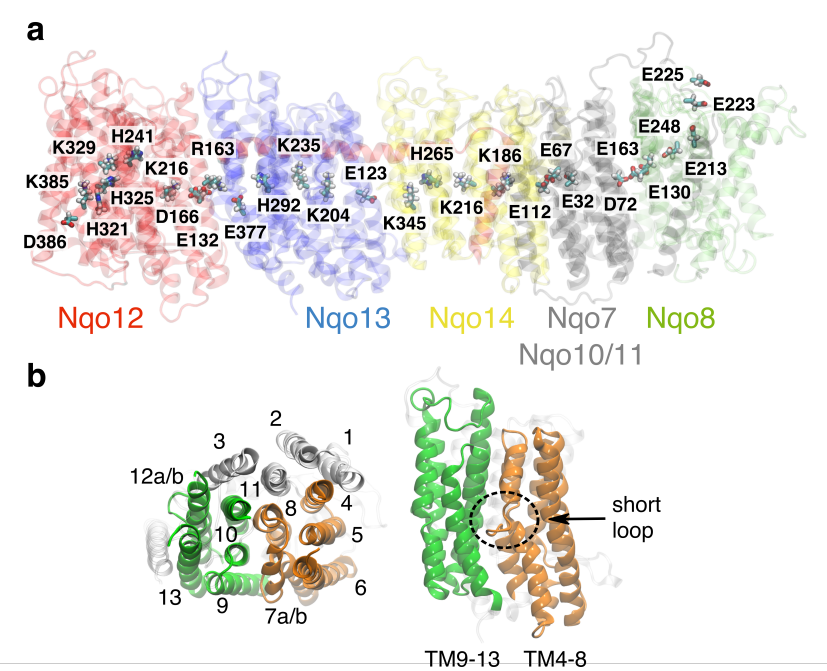


**Figure 2.4:** a) The membrane domain of respiratory complex I from *Thermus thermophilus*. Subunits are shown in red (Nqo12), blue (Nqo13), yellow (Nqo14), grey (Nqo7/10/11) and green (Nqo8). The polar array of charged amino acids extends from Nqo8 to Nqo12. b) TM helices of the antiporter-like subunits. The internally symmetric 5-helix bundles TM4-8 and TM9-13 are shown in orange and green, respectively. The figure is adapted from publication I.

polar residues spans the center of the entire membrane domain (Figure2.4). This long chain of amino acids is well conserved across multiple species and considered to play a key role in the pumping mechanism [103–110]. Interestingly, the array shares a similar pattern within the antiporter-like subunits. This repeated motif is composed of an ion pair that can create contacts to the previous subunit, followed by a central lysine residue, a terminal charged residue, either lysine or glutamate, and a histidine bridging the two charged side chains (Figure 2.4).

The antiporter-like subunits comprise an internal two-fold screw axis pseudo-symmetry (Figure 2.4) [49], with five transmembrane (TM)-helix elements repeated twice in each subunit (TM4-8, TM9-13). The repeating units contain a broken-helix connected by a short loop region, an element which has been shown to be related to the flexibility and formation of water channels for substrates or proton transport [111, 112]. An additional 5-helix bundle is also repeated in the ND1 subunit, forming part of the Q binding pocket. Here, one of the two broken helices segments is disordered and rich in glutamate residues, and is probably related to conformational changes during turnover [92]. Subunit ND5 comprises an additional long horizontal helix (HL) which spans half of the membrane domain creating a connection element between the three antiporter-like subunits [49, 52]. Although it was earlier suggested that this helix is involved in the proton

pumping mechanism, additional studies disproved this hypothesis [113, 114], and the element is now suggested to have a structural rather than a mechanistic role.

The three subunits ND3, ND6 and ND4L are smaller than the other core subunits of the membrane domain, and they link the Q reduction pocket to the antiporter-like subunits. Recently, a long loop region of ND3 was suggested to be involved in the proton pumping mechanism by cross-linking studies [115]. The same loop interacts with the 49 kDa subunit close to the Q-reduction site., and was also suggested to be involved to the A/D transition [54] (Section 2.7).

The free energy released during Q reduction is used for proton pumping. Since the Q reduction site and the most distal subunit are *ca.* 200 Å apart, the signal and energy for the proton pumping have to be vehicled through a series of conformational/electrostatic changes, starting from the redox site. Although many theories emerged during several years to explain this effect, the exact mechanism is not clear [52,98,116–124]. A possible mechanism is presented in publication I and will be discussed in Chapter 5.

## 2.5 The Effect of Membrane Composition

As for many transmembrane enzymes, the activity of complex I is also affected by the membrane composition [125]. In particular, cardiolipin (1,3-bis(sn-3'-phosphatidyl)-sn-glycerol, or diphosphatidylglycerol), has been shown to strongly affect the catalytic properties of respiratory enzymes in mitochondrial and bacterial membranes [126–133], where this lipid can constitute up to 20% of the total membrane composition. The role of cardiolipin and its association with electrogenic enzymes is linked to its unique structure. Cardiolipin comprises two phosphatidic acid groups linked by a glycerol backbone, thus carrying four acyl chains and up to two negative charges. In contrast to complex III, IV and V, where the effect of the lipid on catalysis has been in part understood by both experiments (see ref. [126]) and simulations [134–137], for complex I a mechanistic picture is still lacking. Complex I seems to tightly bind *ca.* 10 cardiolipin molecules together with other lipids [138]. Also, the activity of the enzyme is strongly dependent by the membrane composition, and in particular on the concentration of cardiolipin [128]. Recent higher resolution structures of mammalian enzymes [54, 55, 57] allowed to resolve lipids around the membrane domain including some cardiolipin molecules, but their role was not clear. A mechanistic interpretation of the effect of cardiolipin on the catalysis of complex I will be discussed later in Chapter 5 and has been presented in publication VI.

## 2.6 Supernumerary Subunits

In addition to the 14 core catalytic subunits, complex I comprises numerous supernumerary subunits (up to 31 in mammals) [139, 140]. The characterization of these subunits has been subject of extensive research over the years [141, 142], and the structural assignment was only recently completed and confirmed after publication of high-resolution structures of the complex [57,58]. The number of supernumerary (or accessory) subunits varies in different species, ranging from a 'basic' version of the enzyme in bacteria to the high complexity of mammals. Some of these subunits do not have orthologues in other species, many of them do not have homologs outside complex I [143], and all are unique,

with the exception of the acyl carrier protein SDAP in the mammalian enzyme [144].

Accessory subunits are located around both the membrane and hydrophilic domain. They form a 'cage-like' structure around the core subunits [54, 145], which are mostly unchanged from bacteria to mammals. Although the functions of most of these subunits are unclear, their role seems to be related to stabilization, regulation and assembly [54, 139, 146]. For example, deletion of the 13 kDa subunit (NUMM in *Yarrowia lipolytica*), which harbors the zinc site in the hydrophilic domain, has been shown to affect the final stage of complex I assembly [147]. Other subunits have been suggested to be important for regulatory functions. Based on studies on *Neurospora crassa* and *Bos taurus* [148, 149], subunit B13 was suggested to be involved in the A/D transition (Section 2.7). This subunit is bound to the hydrophilic domain and interacts predominantly with the 30 kDa subunit, and in higher eukaryotes with the 42 kDa subunit (Figure 2.5). The latter subunit, which is found only in some specific eukaryotes (including mammals), has been suggested to be loosely bound to the enzyme due to its partial loss upon purification [141]. Interestingly, the 42 kDa subunit harbours an ATP/ADP binding site that was also found to be phosphorylated in the bovine structure [150]. Subunit 39 kDa interacts directly with the PSST subunit (Figure 2.5) and it comprises an NADPH binding site, the role of which is unclear. The subunit is unlikely to be directly related to the Q reduction process due to the large spatial separation, but it was shown to have regulatory functions and to undergo conformational transition during the A/D transition [151].

## 2.7 The A/D Transition

Preparations of complex I of some species, including mammals [152], show an unusually long lag-phase (tenths of seconds) before reaching the native reaction rate of NADH oxidation. This phenomenon was observed over 50 years ago [153, 154], but only marginally understood, due to the lack of information on the enzyme structure and composition, and was referred to as the A/D (active/deactive,or inactive) transition [151, 155–157].

The D state is characterized by low turnover rates [158]. Early studies suggested that the transition does not affect the NADH oxidation rates, as well as the reduction of the ISC [159], hinting to a rearrangement of the quinone reduction site upon deactivation. The enzyme converts spontaneously to the D state at physiological temperatures and in absence of substrate [160]. The deactivation is characterized by a high activation energy (*ca.* 65 kcal $mol^{-1}$ in *Bos taurus* [158]) and is strongly dependent on the temperature, blocking the enzyme in the A state at 4°C [152]. The activation rate is also affected by other factors as divalent cations ($Ca^{2+}$ or $Mg^{2+}$) or pH [158, 161], which slow down the process. Cross-linking and chemical modification studies [162–165] showed critical conformational changes during the A/D transition in specific subunits of complex I, and suggested how these are linked to the regulation of the enzyme.

Recent mammalian enzyme structures [54, 57] gave new insight into the A/D transition, resolving distinct cryo-EM maps of the *Bos taurus* and *Mus musculus* enzymes. As the two structures showed well distinct features, they were assigned the A and D states of the enzyme. Their assignment was later confirmed by purifying and resolving a structure

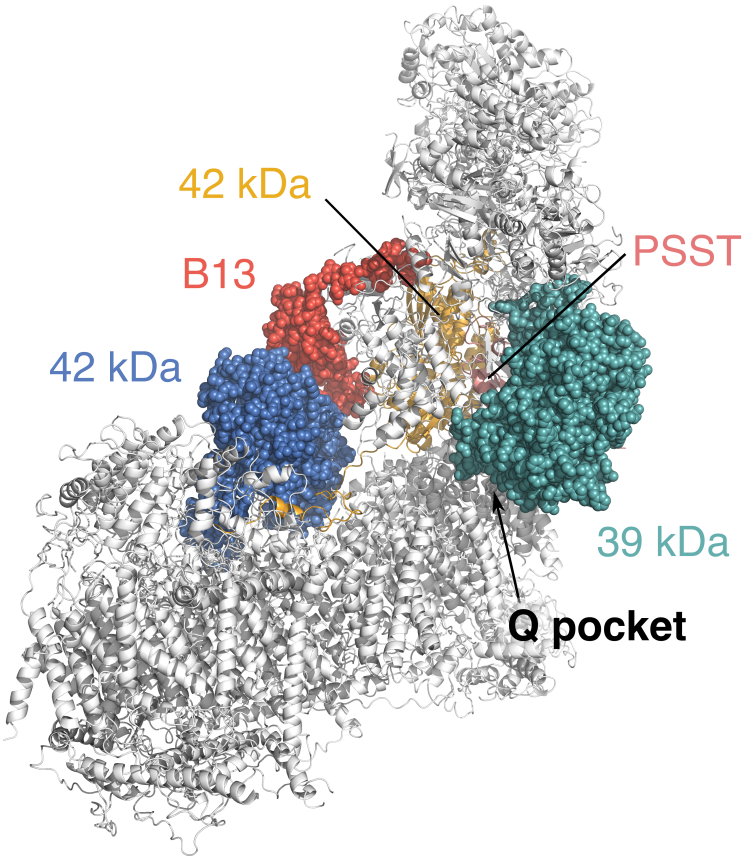


**Figure 2.5:** Complex I of *Mus musculus* (PDB ID:6G2J). Subunit 42 kDa (blue), B13 (red) and 39 kDa (green) have been suggested to be involved in the A/D transition.

of a completely deactive mammalian enzyme [56]. The two distinct structures differ by a twisting motion of the hydrophilic and membrane domains relative to each other, in which subunit ND1 act as a hinge. This global conformational change is reflected in the loss of density of loops in the hinge region of the complex, which have been suggested to become more exposed (ND3 subunit) or disordered (ND1-NDUFS2 subunit).

The idea of a large conformational transition upon deactivation was recently challenged by Parey *et al.* [58], that resolved both deactive and under-turnover structures of *Yarrowia lipolytica*. Although in this structure the resolution is lower (about 4.3 Å), they did not observe the same twisting transition between A and D states under turnover. They concluded that the A/D transition is not related to destructuring of critical loops of the enzyme, but rather to local conformational changes of the quinone binding pocket, and that the D form is not an off-pathway state of complex I but rather similar to a reaction intermediate. It has to be noted that *Yarrowia lipolytica* shows different kinetic parameters of the A/D transition, and could therefore differ from the mammalian process.

# Chapter 3

# Methods and Models

This chapter introduces the methods used in the publications presented in Chapter 5. Due to the extensive literature, only a brief discussion will be provided. The first two sections present an overview of the common methods and approximations used in classical and quantum mechanical frameworks. The third section is an introduction to molecular dynamics and other techniques which are based on the aforementioned schemes. The last section includes a list of the models and simulations used in the publications.

## 3.1 Classical Force Fields

Molecules can be described at several levels of accuracy. Large molecular systems cannot be described with accurate electronic structure calculations, and classical methods often provide a valid and computationally less expensive alternative. These potentials are generally called *force fields*, and they are based on classical mechanics to describe the motion of atoms in space. The atoms move in the external potential generated from the rest of the system [166], and all the energies are derived analytically. The potential energy function contains parameters derived from either experimentally measured properties or from quantum mechanical (QM) calculations. The fitting procedure is performed on a small set of molecules and the parameters are then applied to a wide variety of systems, based on their *transferability* for similar chemical structures.

Depending on the problem at hand, force fields can reach different levels of accuracy, ranging from an all-atom description to a more approximate coarse-grained (CG) representation of molecules. Although different particle types are unique for each force field, most of the models are based on simple 'ball and spring' models, with *ad hoc* potentials dependent on the physics described.

### 3.1.1 All-atom Force Fields and CHARMM

All-atom force fields are usually based on potentials of the form:

$$V = V_{bond} + V_{angle} + V_{dihe} + V_{impr} + V_{VdW} + V_{elec} \tag{3.1}$$

The energy of the system is approximated as a sum of potential energy terms, generally divided into bonded and non-bonded interactions. This way of describing molecular

connectivity and interactions is also referred to as Molecular Mechanics (MM). Each term of the Eq. 3.1 is a sum over interactions arising from the molecular structure and connectivity. The bonded potentials (Eq. 3.2) describe the intra-molecular interactions of bonded atoms by splitting the energy into bonded, angular, dihedral and improper dihedral terms. Both bond and angle terms are usually modeled as harmonic potentials. Dihedral terms describe rotation around bonds and are expressed as a sum of periodic functions over the rotation of the atoms. To enforce planarity when necessary, harmonic 'improper' potentials that describe the out-of-plane oscillation are also often included in force fields. These potentials take the following general form:

$$\begin{gathered} V_{bond} = \sum_{bonds} k_b(b - b_0)^2 \qquad V_{angle} = \sum_{angles} k_\theta(\theta - \theta_0)^2 \\ V_{dihedral} = \sum_{torsions} k_\phi \left[cos\left(n\phi + \delta\right) + 1\right] \quad V_{impr} = \sum_{impropers} k_\varphi(\varphi - \varphi_0)^2 \end{gathered} \tag{3.2}$$

where $b_0$, $\theta_0$ and $\varphi_0$ are the equilibrium values for bond and angle terms, $k_b$, $k_\theta$, $k_\phi$, and $k_\varphi$ are the force constants, $n$ and $\delta$ are the dihedral periodicity and phase, respectively.

The non-bonded potential terms describe the pairwise interaction of atoms that are three or more bonds apart. Exclusion rules usually remove non-bonded interactions of atoms less than two bonds apart, and the interaction of atoms spaced by three bonds are generally reduced by an empirical factor to avoid large forces during dynamics.

The non bonded interactions are usually divided into electrostatic and Van der Waals (VdW) terms:

$$\begin{aligned} V_{elec} &= \sum_{i,j \neq i} \frac{q_i q_j}{4\pi\epsilon_0 \epsilon r_{ij}} \\ V_{VdW} &= \sum_{i,j \neq i} \varepsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - 2 \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \end{aligned} \tag{3.3}$$

The nature of VdW interactions has been studied over several decades [167], and the most commonly used potential function to describe these forces is the 6-12 Lennard-Jones potential [168]. This potential divides the VdW interactions in attractive ($r^{-6}$) and repulsive forces ($r^{-12}$), and depends on two parameters: the collision radius $\sigma$ and the interaction energy at the same distance, $\varepsilon$ (Eq. 3.3). Since the interaction energies decay rapidly over distance, it is possible to evaluate the energy and forces by using only the neighbors of the central atom without loss of accuracy. This provides a basis for efficient calculations, as it strongly reduces the list of pairs for which the potential has to be evaluated. The cut-off distance, after which interactions are truncated, is generally around 12-15 Å for all-atom force fields.

The electrostatic interactions are usually treated as Coulomb interaction energies between point charges, which are assigned to each atom. The purpose of the atomic charge is to reproduce as accurately as possible the interaction with surrounding atoms, mim-

icking the electronic structure of the molecule. This point charge-partitioning scheme has proven to be satisfactory in many cases, but it is not ideal when charge distribution strongly depends of the molecular geometry, or where polarization effects are important [169]. To overcome these limitations, more accurate schemes based on charge fluctuations [170], Drude oscillator models [171], or multipole expansion have been derived [172]. Interestingly, although based on a more accurate physical model, the additional computational cost is not always reflected in a more accurate result [173], and the simple fixed point charge model is therefore still vastly employed. As the electrostatic potential decays as $1/r$ (Eq. 3.3), these long-range interactions can extend up to 30-40 Å, thus increasing the required cut-off distance. This long-range cut-off would in turn greatly increase the computational cost of the calculation, rapidly making it infeasible for large systems. To overcome this issue, and to avoid abrupt truncation of the electrostatic interactions, the Particle Mesh Ewald (PME) scheme is usually employed [174]. The PME algorithm divides calculation of interactions into short-range and long-range, calculated in real and reciprocal space, respectively. The treatment of the electrostatic interactions reduces the scaling from $N^2$ to $N \log N$, where $N$ is the total number of particles.

A good description of the physics underlying molecular geometries is linked to the functional form of the potentials and their adequacy to describe the molecular problem. In addition to that, the set of parameters used in the functions is crucial to determine the correctness of the simulation outcome. Different force fields use both different functional forms and parametrization schemes, and thus transferability between the force fields is not ensured. In force fields, the atoms are usually grouped as 'types', *i.e.* atoms which are expected to show similar properties even in different functional groups. An example is the distinction between aliphatic $sp^3$, $sp^2$ and $sp$ carbon atoms. A correct differentiation of the atom types ensures transferability of the parameters from the molecular structure used for parametrization to a more general context, thus making the force field more reliable.

To date, there are several MM force fields freely available, each suited for a specific context. Among the most used for macromolecules, we can find CHARMM [175, 176], AMBER [177] and GROMOS [178]. Due to their relatively low computational cost, force fields are usually used in combination with molecular dynamics (MD) algorithms, which are described in Section 3.3. Classical MD simulations were used in publications I-VI.

### 3.1.2 Coarse Grained Models and MARTINI

Simulations of macromolecules at an atomistic level provides a detailed description of the system of interest. In combination with MD techniques, classical force fields are nowadays used to describe dynamic of the system to the µs-ms timescale [179]. Such large timescales are often computationally demanding for large systems, and thus all-atom simulations are not well-suited for the purpose.

In case we are not interested in a full atomistic description of our system, coarse-graining (CG) techniques provide a valid alternative to reduce the number of degrees of

freedom of our system, hence decreasing the computational cost. CG methods simplify the description of the system under study and reduce the number of particles mapping $n$ atoms to a single bead (where $n$ depends on the force field used). The energy surface in the phase space of these simulations will be 'smoother', and a longer integration timestep can be used for MD, resulting in turn in longer timescales explored. Cut-off schemes and empirical corrections to achieve a correct structural behavior can be applied to improve both scaling and results of the models.

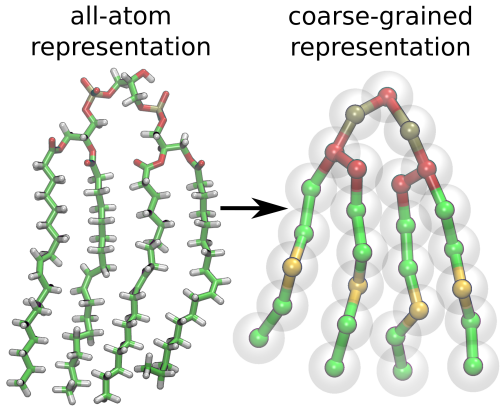


**Figure 3.1:** Typical coarse graining mapping of a lipid molecule (cardiolipin) in the MARTINI force field. The figure is adapted from publication VI.

Over the last ten years, the MARTINI force field [180, 181] became vastly used. It was originally developed as a force field for simulations of lipid bilayers, but since its extension for proteins [181], it has been widely employed to study large systems and long timescale events. MARTINI uses a 1-to-4 heavy atom mapping with some special rules for ring systems. The beads are classified in four large classes: charged, polar, non-polar and apolar. Each group comprises sub-groups with different force field parameters. The particles interact via a classical 6-12 Lennard-Jones potential (Eq. 3.3) with parameters $\varepsilon$ and $\sigma$ dependent on the bead type. Additionally, the charged particles also interact via a screened Coulomb potential ($\epsilon = 15$), and additional bond and angle harmonic potentials (similar to Eq. 3.2) ensure to keep the bound particles together. Additional elastic network schemes have been employed to improve the description of secondary and tertiary structure elements [182]. A typical mapping for a lipid molecule is shown in Figure 3.1. CGMD simulations were employed in simulation VI.

### 3.1.3 Elastic Network Models

Protein structure encodes information about their dynamics and function [183]. As more protein structures became available, computational studies focused their attention on the tight connection between the structure of proteins and their normal modes. Normal mode analysis (NMA) is a tool to retrieve information about collective structural vibration of molecules. The normal modes of enzymes can be important *e.g.* for catalysis or

substrate binding/dissociation [184]. Also, these motions are conserved among classes of proteins [185]. Calculation of normal modes requires prior minimization of the structure to a local minimum. Although it is possible to study global protein motion with all-atom models, the minimization, in this case, is somewhat challenging, since it might distort the original structure. As originally proposed by Tirion [186], a valid alternative to study normal modes is to model each amino acid as one point mass (usually using the position of $C_\alpha$ atoms), connected by harmonic springs within a certain cut-off radius. The total energy can be thus written as:

$$E = \sum_{a}^{N} \sum_{b=a+1}^{N} \frac{C}{2} f_{cut}(r_{ab}) \left(r_{ab} - r_{ab}^{0}\right)^2 \tag{3.4}$$

where $N$ is the total number of point masses, $r_{ab}^{0}$ is the structural (experimental) distance, $C$ represents the strength of the oscillator and $f_{cut}(r_{ab})$ is:

$$f_{cut}(r_{ab}) = \begin{cases} 0, & \text{if } r > r_{cut} \\ 1, & \text{otherwise} \end{cases} \tag{3.5}$$

Although this approximation might seem somewhat inaccurate in the treatment of the pairwise potential between beads, it is perfectly suited and exceptionally robust for the study of global 'protein mechanics' [187]. The basic assumption of these models is that low-energy motions of large molecules are independent of the atomic details of the structure, and mostly defined by their native contact topology. This approach has also the advantage of avoiding the initial minimization step, since by definition the network is in its minimum energy conformation. Although fine-tuning of the force constants and modulation of the cut-off scheme can provide a better agreement with experimental data, the results obtained by using identical spring constants and a 'hard' cut-off are generally considered to be quite robust and accurate [188].

The two most widely used schemes for these type of studies are Gaussian Network Models (GNM) [189, 190] and Anisotropic Network Models (ANM) [188, 191]. The difference between the two models arises from the energy function, which depends only on the distance of the nodes in the case of ANM, and also on their relative orientation in GNM. The parametrization of the network models to reproduce experimental B-factors of proteins resulted in a cut-off of 7 and 15 Å, for GNM and ANM respectively, using a force constant of 1 kcal $\text{mol}^{-1}$ [188].

Approaches combining MD and network models, but also a combination of multiple networks have also been proposed [192, 193] and successfully applied in the study of protein conformational changes, *e.g.* for ATP synthase [194]. ANM were used in publication IV.

## 3.2 Quantum Chemistry

In parallel to the more approximate classical models, quantum mechanical (QM) methods have also become more applicable to biological systems, both due to the increasing

computational power and more accurate methods to describe electronic structure. We discuss here, however, only the Density Functional Theory (DFT). This methodology is commonly employed in QM/MM calculations, due to the optimal balance between accuracy and computational speed.

### 3.2.1 Density Functional Theory

DFT is based on the Hohenberg-Kohn theorem [195], which demonstrates that the ground state electronic energy and other properties are uniquely defined by the electron density of the molecule. Thus, the energy is a *functional* of the electron density $\rho$:

$$E = E\left[\rho\left(\mathbf{r}\right)\right] \tag{3.6}$$

The exact form of the energy functional is not known. To circumvent the problem, Kohn and Sham [196] used a fictitious system with the same electron density $\rho$, but non-interacting electrons, and wrote the energy functional as:

$$E\left[\rho\right] = T_S\left[\rho\right] + E_{ne}\left[\rho\right] + J\left[\rho\right] + E_{xc}\left[\rho\right] \tag{3.7}$$

where $T_S$ is the exact kinetic energy of the ideal system, $E_{ne}$ is the coulombic interaction between nuclei and electrons, $J$ is the average repulsion of the electronic density, and $E_{xc}$ is the exchange-correlation term, which contains the unknown part of the energy functional. A central aspect of DFT improvement is the development of better exchange-correlation functionals, which cannot be determined exactly.

Similarly to Hartree-Fock (HF) theory, the introduction of the atomic orbitals to describe the density function leads to a set of equations, which has to be solved self-consistently, due to the dependence of the energy functionals on the density itself. The electronic problem is thus solved iteratively starting from an initial guess until a predefined precision is achieved. Two frequently used functionals used in chemistry are the hybrid functional B3LYP [197–199], and the generalized gradient approximation (GGA) functional BP86 [200, 201] due to their ability to reproduce chemical properties with good accuracy. Dispersion effects are usually not well reproduced by most DFT functionals, and to recover the correct dispersion interaction energies, a common approach is to introduce an additional energy term appropriately parametrized to reproduce the missing interaction. One of the most widely used approaches is Grimme's D3 dispersion correction [202].

### 3.2.2 QM/MM

A computational treatment of large biomolecules such as enzymes is intractable at QM level, due to a large number of computations that would be needed to solve the quantum mechanical problem. At the same time, classical methods do not describe electronic structures, and thus chemical reactivity cannot be directly modeled with classical approches. Hybrid QM/MM methods solve this problem by describing the relevant (reactive) part of the system of interest with QM methods, and the remaining surroundings with classical force field methods. The method was originally formulated by Warshel and Levitt, who applied it to study the stabilization of a carbonium ion in the lysozyme

reaction [203].

The partition of the interactions between classical and quantum regions of the system can be generally classified in subtractive and additive scheme:

$$\begin{aligned} E^{sub}_{QM/MM} &= E_{MM}(QM+MM) - E_{MM}(QM) + E_{QM}(QM) \\ E^{add}_{QM/MM} &= E_{QM/MM}(QM/MM) + E_{MM}(MM) + E_{QM}(QM) \end{aligned} \tag{3.8}$$

In the subtractive scheme, the total MM energy is calculated and the correction is then added by removing the MM region corresponding to the QM atoms and subsequently adding the QM energy. In the additive scheme, the total energy is the sum of three energy terms corresponding to the MM region, QM region and the explicit coupling between them ($E_{QM/MM}(QM/MM)$ in Eqn. 3.8). Specific details of the energy function depend also on the implementation of the algorithm. In general, it is desirable to introduce polarization of the QM region by the MM environment, since it could play an important role in catalysis.

The choice of the QM method to use in the simulation is generally driven by the specific problem at hand, but should in general not be computationally too expensive, since simulations are often used in combination with molecular dynamics, where each self-consistent field and gradient evaluation step is repeated for hundreds, or thousands of times. DFT is one of the most commonly used techniques in modeling of complex enzymes. [204].

Treatment of the QM/MM boundaries is often challenging for QM/MM models. Influence of the MM atoms on the QM region can strongly affect the computed properties. A common issue is that the boundaries of the QM regions suffer from 'overpolarization' effects. Due to their proximity to the QM atoms, the classical particles create a strong field that drags electron density away from the QM region and artificially polarizes it. To avoid this effect, point charges can be smeared out or selectively turned off, but the solutions are generally system specific and should be optimized for each case [205].

Proper treatment of the boundaries is especially relevant when the they cross covalent bonds. In these cases, it is necessary to derive a scheme to break down the forces acting on the interface atoms. Although many approaches have been proposed [206], one of the most common is the use of a link-atom between QM and MM atoms [205, 207, 208]. A fictitious hydrogen atom is added on the bond between QM and MM region to saturate the QM structure. Forces on this atom are distributed on the MM and QM atoms to properly achieve energy conservation. Although approximate, this scheme has been proven to be quite satisfactory in the description of the QM/MM systems. QM and QM/MM methods were applied in publications I and III.

## 3.3 Molecular Dynamics

For biomolecular systems, MD simulations are usually performed using either force fields (classical MD), *ab-initio* (QM) or mixed QM/MM potentials. Independent of the used type of potential, dynamics is obtained by integrating the Newton's second law of motion.

The force **F** acting on the particle $i$ defines acceleration via:

$$\mathbf{F}_i = -\nabla V_i = m_i \mathbf{a}_i = m_i \frac{\partial^2 \mathbf{r}_i}{\partial t^2} \tag{3.9}$$

where $V_i$ is the potential acting on the particle $i$, and $m_i$ is its mass. The potential $V_i$ is the sum of all the interactions of the particle with its environment. Even for classical potentials, the motion of a large number of particles used in an MD simulation is so complex that it is impossible to treat it analytically. For this reason, dynamic is usually discretized in timesteps ($\delta t$), and the particle motion is described via finite difference methods, where the forces and energies are calculated at each step. There are numerous ways to propagate the equations of motion, each with advantages and drawbacks. To derive integration schemes, position ($\mathbf{r}$) and velocity ($\mathbf{v}$) at time $t + \delta t$ are expanded as a Taylor series in $t$:

$$\begin{aligned} \mathbf{r}(t+\delta t) &= \mathbf{r}(t) + \delta t \mathbf{v}(t) + \frac{1}{2}\delta t^2 \mathbf{a}(t) + \dots \\ \mathbf{v}(t+\delta t) &= \mathbf{v}(t) + \delta t \mathbf{a}(t) + \dots \end{aligned} \tag{3.10}$$

One example is the *Verlet algorithm* [209], which is obtained by expanding $\mathbf{r}(t + \delta t)$ and $\mathbf{r}(t - \delta t)$:

$$\begin{aligned} \mathbf{r}(t+\delta t) &= \mathbf{r}(t) + \delta t \mathbf{v}(t) + \frac{1}{2}\delta t^2 \mathbf{a}(t) + \dots \\ \mathbf{r}(t-\delta t) &= \mathbf{r}(t) - \delta t \mathbf{v}(t) + \frac{1}{2}\delta t^2 \mathbf{a}(t) - \dots \end{aligned} \tag{3.11}$$

and adding these two equations one obtains:

$$\mathbf{r}(t+\delta t) = 2\mathbf{r}(t) - \mathbf{r}(t-\delta t) + \delta t^2 \mathbf{a}(t) + \mathcal{O}(\delta t^4) \tag{3.12}$$

The algorithm uses $\mathbf{r}(t)$, $\mathbf{r}(t - \delta t)$, and $\mathbf{a}(t)$ to calculate the position at the time $t + \delta t$. Velocities do not appear explicitly in the calculation and can be calculated by:

$$\mathbf{v}(t) = \frac{[\mathbf{r}(t+\delta t) - \mathbf{r}(t-\delta t)]}{2\delta t} \tag{3.13}$$

Also other integration schemes, such as the velocity Verlet [210], or the Leap-frog algorithm [211], are widely used. The choice of the timestep is important to ensure the best compromise between simulation stability and exploration of the phase space. It is important to choose the largest timestep possible and increase the total simulation time. At the same time, choosing a large timestep could result in overlapping atoms and large forces during dynamics, making the simulation unstable. Depending on the type of system under investigation, the timestep can be tuned to balance these two aspects. A good rule of thumb is to use a timestep about 10 times smaller than the fastest motion in the simulation. In the case of all-atom simulations of biomolecules, this turns out to be roughly 1 fs, since the X-H bonds vibrations usually occur in $10^{-14}$ s. Since we are often not

interested in the simulation of vibrations of hydrogen atoms, they are commonly frozen out using the SHAKE [212] or LINCS [213] algorithm. These algorithms usually allow to double the timestep used in all-atoms simulation without significant loss in accuracy. Other solutions are based on the reduction of the number of particles using united-atoms approaches, *e.g.* the GROMOS force field [214], which removes the fast-moving light atoms, similar to a first step of coarse graining. Simulations using MARTINI (Section 3.1.2) can use timesteps of 20 fs and still result in stable simulations.

In molecular dynamics simulations, Langevin dynamics is often employed to model correct solvent dynamics and to control temperature and pressure of the system. A friction coefficient of the form of:

$$\mathbf{F}_{frict} = -\gamma_i m_i \mathbf{v}_i \tag{3.14}$$

is added to the forces acting on the atoms, together with a normally distributed random force with zero mean. These additional terms ensure a stochastic behaviour of the simulation, and help in damping the temperature oscillations. Typical values for the friction coefficients are about 1-10 $ps^{-1}$.

Adequate treatment of system boundaries is important to reproduce properties of bulk phase. To properly model the surrounding solvent, it might be necessary to simulate several thousands of particles in order to accurately model few internal molecules. Although this is not a problem with today's computing power, it is important to reduce the cost necessary to evaluate a property as much as possible. To this end, *Periodic Boundary Conditions* (PBC) are a commonly employed to minimize boundary effects. PBC are applied by modelling the simulation box as a unit cell of an infinite periodic lattice, with a replica of itself in each periodic cell. This results in a 'wrapping' of the simulation box, so that molecules close to the edge will interact with particles on the other side of the box. There are few possible shapes for unit cells, but the simplest, orthorhombic shape is the most widely used. With PBC, the system will experience a long-range order which may affect the results of the simulation. Moreover, the periodic box should be large enough to satisfy the 'minimal image convention', which ensures that each particle of the box interacts at maximum only once with each other particle of the box. The usual cut-off scheme employed in force fields truncates the direct interaction, thus meaning that the simulation box has to be at least twice as large as the cut-off radius. In addition, large biomolecules should not interact, or to the least extent possible, with copies of themselves in the other images. To reduce the interaction, a rule of thumb is to space the periodic images of the molecules of at least 30 Å, where even long range electrostatic energies decay to less than 1 kcal $mol^{-1}$.

MD simulations can be performed in different *ensembles*. For biomolecular simulations, we are often interested in performing MD in either constant temperature (*NVT*) or constant pressure and temperature (*NPT*) ensembles. The temperature of the system is directly linked to the total kinetic energy via:

$$\langle K \rangle = \frac{3}{2} N k_{\mathrm{B}} T \tag{3.15}$$

and hence, to the velocity of the particles. Velocity rescaling was one of the first ways to simulate dynamics at constant temperature [215], but further developments adopted the coupling of the particles to a thermostat, which acts as an energy reservoir and controls the temperature of the particles, or the addition of Langevin terms. Over the years, thermostats as the Nose-Hoover [216] or Berendsen [217] became widely used. For *NPT* simulations, the pressure can be controlled similarly to the temperature. The instantaneous pressure can be calculated via the virial coefficient, through the particles' position and forces (potential). Control of the pressure is usually achieved via modulation of the volume of the periodic cell, and barostats act in a similar way to thermostats [218].

## 3.4 Free Energy Methods

The free energy is a central quantity in thermodynamics since it carries the information about the likelihood of states. Although the foundations for calculating free energies were developed already in the 1930s [219], computing power limited their application to biochemical systems. In the 1950s, Zwanzig [220] revolutionized the field by applying a perturbation approach to the free energy calculation, creating the basis for what are now known as free energy perturbation (FEP) methods [221].

The free energy difference between two states, 1 and 0, can be calculated as:

$$\Delta G = -\frac{1}{\beta} \ln \frac{Q_1}{Q_0} \tag{3.16}$$

where $\beta = 1/(\mathrm{k_B T})$ and $Q_i$ is the partition function of state $i$. The partition function is proportional to:

$$Q_i \propto \iint \exp\left[-\beta H_i(\mathbf{x}, \mathbf{p})\right] d\mathbf{x} d\mathbf{p} \tag{3.17}$$

where $H_i$ is the hamiltonian of system $i$, $\mathbf{x}$ are the position of the particles, and $\mathbf{p}$ are their momenta. If it is possible to express the hamiltonian of the final system (1) as:

$$H_1(\mathbf{x}, \mathbf{p}) = H_0(\mathbf{x}, \mathbf{p}) + \Delta H(\mathbf{x}, \mathbf{p}) \tag{3.18}$$

where $H_0$ is the hamiltonian of state 0 and $\Delta H$ the difference in potential between state 1 and 0, we can rewrite equation 3.16 in terms of 3.17 and 3.18 and obtain:

$$\Delta G = -\frac{1}{\beta} \ln \frac{\iint \exp\left[-\beta \Delta H(\mathbf{x}, \mathbf{p})\right] \exp\left[-\beta H_0(\mathbf{x}, \mathbf{p})\right] d\mathbf{x} d\mathbf{p}}{\iint \exp\left[-\beta H_0(\mathbf{x}, \mathbf{p})\right] d\mathbf{x} d\mathbf{p}} \tag{3.19}$$

or, in simpler notation:

$$\Delta G = -\frac{1}{\beta} \ln \langle \exp\left[-\beta \Delta H(\mathbf{x}, \mathbf{p})\right] \rangle_0 \tag{3.20}$$

where the $\langle \dots \rangle_0$ denotes an average over the state 0. This equation demonstrates that it is possible to calculate free energy differences between two states of the system just

by simulating one of them, at least in theory. The key aspect in determination of an accurate free energy difference is the sampling of relevant states of the target system while simulating the reference one. In fact, if the statistical sampling of the low energy regions of the target state is not accurate enough, the application of eq. 3.20 will result in a poor estimate. Molecular dynamics simulations generally fail in this goal, since they are likely explore the region around the reference state according to a Boltzmann distribution, and only more rarely diffuse to different regions of the phase space.

To overcome this issue, and to obtain a better free energy estimate, the idea of 'stratification' has been of central importance. With stratification, the sampling of the entire phase space regions of interest is obtained by dividing the reaction path (collective variable) in 'windows', and enforcing the system to stay in specific phase space regions by addition of restraints. This idea is exploited by umbrella sampling techniques.

### 3.4.1 Replica Exchange Umbrella Sampling

Replica Exchange Umbrella Sampling (REUS), and in general umbrella sampling, uses the idea of stratification of the reaction coordinate to estimate free energy difference. The addition of the replica exchange algorithm ensures a better sampling and a faster convergence [222–224].

We are usually interested in the free energy change along a reaction coordinate, $\xi$, that describe a chemical of physical process. Sampling of the entire phase space based on the initial condition is in theory possible, but might require unreachable simulation timescales. The idea is then to run multiple simulations adding different biasing potentials, in order to modify the sampling probability of the states, and enforce the sampling of high energy (low probability) regions. The Hamiltonian of window $i$ can be written as:

$$H(\lambda_i) = H_0 + V_i(\xi, \lambda_i) \tag{3.21}$$

where $\lambda_i$ is a parameter that describes the position of the windows along $\xi$.The positioning of the windows, their 'width' and the strength of the biasing potential have to be decided a priori, but methods to improve the initial choice during the simulation have also been developed [225]. The recovery of the full free energy profile requires reweighting of the probability distributions obtained along $\xi$ by methods such as Weighted Histogram Analysis Method (WHAM, see section 3.4.2). The replica exchange methods (called also bias-exchange, or parallel tempering in case of exchanging temperature) aim to improve the statistical sampling of the simulations by making (in principle) all the simulations accessible to all the biasing potentials [223, 226]. By running the simulations in parallel, a trial move to exchange the biasing potential between two simulations is attempted every $n$ steps. By doing so, the simulations are able to 'diffuse' between different windows, which improves the statistical accuracy of the free energy estimation without any additional cost. The probability of exchange is given by a Metropolis criterion, using the $\Delta E$ of exchange as a probability measure. This exchange rule can be seen as an extent of the 'overlap' between the sampling regions of the two windows.

### 3.4.2 Weighted Histogram Analysis Method

Histogram reweighting methods to construct free energy profiles are based on the recovery of the 'true' probability distribution of sampled windows. This is achieved by removing the effect of the bias added to enforce sampling of high energy regions, *e.g.* during umbrella sampling simulations. In practice, the probability distributions are described as histograms, and therefore these methods are called histogram reweighting methods. One of the most popular approaches is the Weighted Histogram Analysis Method (WHAM).

WHAM is based on the combination of multiple histograms from different simulations with different biases, weighting each probability distribution by a factor, $w_i$.

$$p(\xi) = \sum_i w_i p(\xi, \lambda_i) \tag{3.22}$$

The weighting factors are normalized and optimized to ensure the lowest statistical error of the total probability distribution. This histogram-based free energy reconstruction, proposed initially in 1989 by Ferrenberg *et al.* [227] and extended by Kumar *et al.* [228], is one the most widely used for biomolecular processes and often used in combination with umbrella sampling methods [229]. Histogram-free reweighting methods, such as MBAR, have also been suggested [230]. REUS and WHAM were applied in publication I.

### 3.4.3 MM-GBSA/MM-PBSA

Methods as FEP or umbrella sampling can be computationally demanding, or even infeasible to apply for calculations of binding free energy ($\Delta G_{bind}$) of substrates to enzyme active sites. In this case, more approximate ways to estimate binding free energies are desirable. MM-GBSA/PBSA (Generalized Born/Poisson Boltzmann and Surface Area) are a class of methods developed for estimation of binding free energy by direct simulations of the compounds of interest [231, 232]. The free energy is estimated as a sum of multiple factors [233]:

$$G = E_{bond} + E_{elec} + G_{pol} + E_{VdW} + G_{apol} - TS \tag{3.23}$$

where $E_{bond}$,$E_{elec}$ and $E_{VdW}$ are the same terms of Eq. 3.1, $G_{apol}$ refers to the free energy of formation of the interface surface, and $G_{pol}$ to the solvation energy of the molecule. The latter term is estimated via PB or GB calculation [234, 235]. The entropy term can be estimated via NMA [233], and other strategies have been suggested [231], but since it is rather expensive and often not well-converged, it is generally omitted.

If we consider the binding process:

$$\mathrm{P + L} \rightleftharpoons \mathrm{PL} \tag{3.24}$$

The binding free energy can be evaluated as:

$$\Delta G_{bind} = \langle G_{PL} \rangle_{PL} - \langle G_P \rangle_P - \langle G_L \rangle_L \tag{3.25}$$

where $G_{PL}$ is the free energy of the protein-ligand complex, and $G_P$, $G_L$ the free energy

of protein and ligand, respectively. The averaging is performed on the structure obtained from MD simulation of the PL, P and L systems. Often, simulation of the three different systems leads to slower convergence of the results, due to the large fluctuations in energy. To improve convergence at the cost of accuracy, a common approach is to simulate only the PL complex and extract energies of P and L from the structures of the complex:

$$\Delta G_{bind} = \langle G_{PL} - G_P - G_L \rangle_{PL} \tag{3.26}$$

In addition to reducing the computational cost by roughly one third, this approximation reduces the energy fluctuations, since $E_{bond}$ in Eq. 3.25 cancels out, and $E_{elec}$ and $E_{VdW}$ only refers to protein-ligand interactions. Indeed, the underlying approximation is that both enzyme and substrate do not undergo large conformational changes upon binding, and their bound- and apo-state geometries are similar. MM/GBSA method was applied in publication II.

## 3.5 Principal Component Analysis

Principal component analysis (PCA) is a multivariate statistical technique used to reduce the complexity of variables distribution [236]. PCA transforms the data by extracting the highest variance directions in the variable space and using them as new collective coordinates for the description of the dataset. In the context of protein simulations, the analysis helps in extracting underlying important collective motions from noisy finite-temperature dynamics [237]. The analysis uses the correlation matrix $\mathbf{C}_{ij}$ defined as:

$$\mathbf{C}_{ij} = \langle (q_i - \langle q_i \rangle)(q_j - \langle q_j \rangle) \rangle \tag{3.27}$$

where $q_{1..N}$ are the values of the variables used for analysis. In the analysis of protein dynamics, cartesian coordinates of $C_\alpha$ atoms are often used, but internal coordinates have been suggested to be a better choice for the description of collective motions [238]. The structures used for the calculation are pre-aligned to minimize the variance with respect to the average position.

PCA is performed by diagonalizing the $\mathbf{C}_{ij}$ matrix to retrieve the eigenvalues/eigenvectors of the decomposition. The eigenvectors describe the internal protein dynamics, ordered by eigenvalue, or amount of variance recovered, and usually the lowest (*ca.* 20) eigenvectors are enough to recover a large part of the entire protein motion. This is in principle a large reduction of variables, but also an indication that the principal components can describe important molecular motions that are functionally important. In this respect, the PCA is similar to a NMA, and recovers from a direct ensemble of structures similar information about essential protein motions.

PCA is just the simplest of this class of methods devoted to protein dynamics analysis. More complex analysis are based on similar ideas, *e.g.* kernel PCA [239] or tICA analysis [240]. PCA was applied in paper III, IV, VI.

## 3.6 Continuum Electrostatic Models and p$K_a$ evaluation

Evaluation of the p$K_a$ values of amino acids is of central importance to predict protein function and they can be strongly affected by local environment. p$K_a$ shifts can be of extreme relevance for electrostatic-driven conformational changes, or during reactions which involve exchange of protons. The difference of pH across the membrane is a way to store free energy, and proton uptake/release is also achieved by tuning pKa values of protein residues [241]. To date, several methods to calculate p$K_a$ values of residues are avalaible, based on heuristic, QM, or electrostatic [242–244] calculations.

The p$K_a$ evaluation performed in this thesis are based on the electrostatic approach [245]. The p$K_a$ shift moving from a solvent-exposed side chain to the protein interior is performed in two steps: evaluation of the intrinsic p$K_a$, and iterative calculation of the pairwise interaction energies of titratable groups. In the first step, the intrinsic p$K_a$ is evaluated through the formula:

$$\mathrm{p}K_a^{\ prot} = \mathrm{p}K_a^{\ bulk} + \frac{\Delta G^{A^-}_{solv}}{2.303RT} - \frac{\Delta G^{AH}_{solv}}{2.303RT} \tag{3.28}$$

where the p$K_a^{\ bulk}$ is the p$K_a$ in the bulk solvent (typically water), and $\Delta G^{A^-}_{solv}$ and $\Delta G^{AH}_{solv}$ the change in free energy upon transfer from solvent to protein of the deprotonated and protonated species, respectively. While the first term can be experimentally measured, the latter two are estimated via Poisson-Boltzmann equation [234]. The free energy change can be estimated from two contributions: the Born energy, and the interactions of the background charges. Since buried residues can mutually affect their p$K_a$, protonation probabilities are also weighted for the interaction energies of the amino acids [246]. The free energy of a protonation free energy of residue $\mu$ can be written as: :

$$G_\mu = \left(x_\mu - x^0_\mu\right) RT \ln 10(\mathrm{pH} - \mathrm{p}K_{a_\mu}^{\ prot}) + \frac{1}{2}\sum_\nu^N W_{\mu\nu}\left(x_\mu + z^o_\mu\right)\left(x_\nu + z^o_\nu\right) \tag{3.29}$$

where $N$ is the total number of titratable groups, $x_\mu$ is 1 or 0 depending on the protonation state and $z^o_\mu$ is the formal charge of the group. To achieve a proper estimate of the mutual interaction energies, it is possible to perform a Monte Carlo evaluation [245] of the $2^N$ possible states ($N$ is the number of titratable residues), and find the most probable protonation state of the system. Continuum electrostatic calculations were applied in publication I, V, VI.

## 3.7 Models

The energy-conversion mechanism of respiratory complex I presented in this thesis has been studied with a variety of computational methods, which are briefly described in this section. A complete explanation of the systems setup and methods used can be found in the original articles.

### 3.7.1 Complex I Structure

Structures of complex I used in this thesis are the *Thermus thermophilus* (PDB ID:4HEA) [36] and *Bos taurus* in the active and deactive states (PDB ID:5LC5, 5LDW) [54]. Placement of the Q was performed initially identifying the cavity with the HOLE software [247], and placing the head group between His59 and Tyr108 of 49 kDa subunit (*Bos taurus* numbering).

### 3.7.2 Force Fields and Parameters

The protein was modeled using the CHARMM27 [175, 248] (publication I, III, IV, V) or CHARMM36 [249] (publication II, VI) force fields. Additional parameters for the cofactors (FMN, Q, ISC) were obtained by in-house parametrization with DFT at the B3LYP-D3/def2-SVP/TZVP level [250] performed with TURBOMOLE [251]. In publication VI, the protein was simulated using the MARTINI 2.2 force field [180, 252].

### 3.7.3 Molecular Dynamics

In publication I, III, and V, complex I was simulated in a lipid-water-ion surrounding, mimicking the native biological environment. We embedded the protein in a POPC membrane, solvated with TIP3P water [253] and added ions (NaCl) to a 100 mM concentration. CHARMM-GUI [254] was used in the preparation of the system in publication VI. The simulations were performed using a 2 fs timestep and in combination with the ShakeH algorithm. The simulations were performed in an *NPT* ensemble with $T$= 310 K and $P$= 1 atm with Langevin dynamics. The long-range electrostatic interactions were treated with the PME method [174]. All the simulations were performed with CHARMM [255], NAMD2 [256] or GROMACS [257, 258] simulation packages. All the systems were subject to initial minimization and equilibration. The system visualization and setup were performed with VMD [259] and PyMOL [260]. In publication VI, Caver [261] was used with a probe radius of 2.25 Å to identify the cavities that could work as proton transfer pathways.

### 3.7.4 Normal Mode Analysis and Principal Component Analysis

In publication IV, NMA was performed on *Bos taurus* and *Thermus thermophilus* structures, using $C_\alpha$ atoms as network nodes, a cut-off of 15 Å and force constant of 1 kcal mol$^{-1}$. PCA was performed using $C_\alpha$ atoms of an ensemble of structures obtained from publication I. The calculations were performed using ProDy [262]. In publication VI, PCA was performed using PyEMMA [263] on the backbone atoms.

### 3.7.5 Protonation States by Continuum Electrostatic Calculations

In publication I, III, and V protonation states of polar residues were probed using PB continuum electrostatic calculations, followed by MC sampling of the $2^N$ protonation states. The PB calculations were performed using Adaptive Poisson-Boltzmann Solver (APBS) [234] and Karlsberg+ [244, 245] to sample the protonation states. The protein

was described as explicit atoms embedded in a dielectric medium of $\varepsilon = 4$. The surface of the protein was calculated using a probe radius of 1.4 Å, and bulk water described with $\varepsilon = 80$ and an ionic strength of 100 mM.

### 3.7.6 QM/MM Molecular Dynamics

In publication I, we performed QM/MM simulations based on structure obtained from the classical MD trajectories. The QM/MM systems were built using Nqo12/13/14 subunits and surrounding lipids and water molecules (*ca.* 75,000 atoms). The QM region comprised the residues involved in the proton transfer reaction and nearby residues, and it was treated at DFT level using the BP86 functional [200, 201], def2-SVP basis set and the multipole accelerated resolution of identity (MARIJ) approximation [264]. The QM/MM interface was treated by adding link atoms between $C_\alpha$ and $C_\beta$ atoms for each amino acid. Simulations were performed at 310 K with a 1 fs timestep, using the CHARMM/TURBOMOLE interface [265].

### 3.7.7 Free Energy Profiles

Free energy profiles were obtained in publication I along the ion pair dissociation coordinate. The initial geometries were obtained from a Steered Molecular Dynamics simulation (SMD). The REUS was performed using 25 replicas of the system and adding a restraint of 50 kcal mol$^{-1}$ Å$^{-2}$ to the distance of Lys-204($N_\zeta$)-Glu123($C_\delta$). The distance was restrained between 3 and 12 Å. The simulation were extended for 14 ns each to ensure convergence of the free energy profiles. The reconstruction of the free energy profiles was performed using WHAM [266].

# Chapter 4

# Aims of the Study

Elucidation of the function of complex I is one of the most important topics in modern bioenergetics, and numerous biochemical, computational, and structural studies uncovered several aspects of the enzymatic reactions in recent years. Yet, the energy transduction mechanism and the network of signal transport between subunits is not well understood. The aims of this work were:

- To elucidate the location of the proton conducting pathways in the membrane domain.
- To understand the long-range coupling process between Q reduction and proton pumping.
- To investigate the role of global dynamics of the enzyme in the regulation of the catalysis.
- To clarify the role of lipids in the modulation of enzyme catalysis.

In addition to these aspects, modeling of the enzyme substrate in its binding pocket served as a basis to interpret experimental results regarding the effect of different quinone tail lengths in enzyme catalysis. This latter study was conducted in collaboration with Dr. Judy Hirst's team at the MRC Laboratory of Molecular Biology (Cambridge) and is presented in Section 5.2.

# Chapter 5

# Results and Discussion

The articles presented in this thesis cover many aspects of the function of complex I. The results section is organized as follows:

1. Proton pumping subunits and their long-range coupling.
2. Aspects of the quinone catalytic cycle and motion in the binding pocket.
3. Role of global motions in regulation of enzymatic activity.
4. Effect of the membrane composition on the regulation and activity of the enzyme.

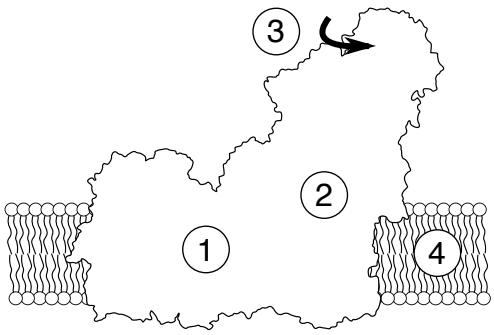


**Figure 5.1:** Schematic view of complex I embedded in a lipid membrane. The numbering in the structure of complex I is based on the different subsections in the results chapter.

## 5.1 The Proton Pumping Mechanism

Structural and biochemical studies pointed out residues that could be involved in the proton pumping of complex I [49, 95, 101], but a clear picture of the mechanism and proton translocation pathways is still missing. Following the simulations in ref. [102], we focused in publication I and III on the proton pathways in the membrane domain of

complex I, and how a long-range energy transduction process can be achieved. These studies were performed prior to structures available for the entire mammalian enzyme, and we therefore used the enzyme from *Thermus thermophilus*.

### 5.1.1 Symmetric Proton Transfer Pathways

To probe the pumping mechanism in complex I, we first studied in publication I the channel topology of the membrane domain. To this extent, we performed MD simulations of the entire enzyme starting from the crystal structure (PDB ID:4HEA), and studied the formation of water wires connecting buried charged residues with bulk water. The charged residues were modelled in different protonation states, to study their effect on the formation of water channels. The simulations were extended for a total of 3.8 μs. To probe the identified proton transfer pathways, we performed additional QM/MM calculations, and used REUS simulations to investigate the free energy of dissociation of buried ion pairs.

During the dynamics, we observed opening of putative proton channels and formation of quasi-1D water wires in the antiporter-like subunits, similarly to many proton channels [267–269] (Figure 5.2A). The wires formed on a 0.2-0.4 μs timescale (Figure 5.2B), connecting the buried charged residues with the N- or P-side of the membrane in multiple sites. A similar proton channel topology was observed in the three antiporter-like subunits. The chains do not form continuous connections across the two sides of the membrane, but are broken by conserved histidine residues, part of the long central axis of charged amino acids. Interestingly, we observe that the proton pathways not only present similarities between the antiporter-like subunit, but also resemble their internal pseudo-symmetry (Figure 5.2C). Considering each subunit as formed by a pair of 5-TM helix bundle elements with roto-translational symmetry (see Chapter 2), the connectivity to the bulk is established on the side of the broken helix element for all six bundles (in subunit Nqo12-14). Upon opening of the proton channels, we observe changes in the orientation of the broken helix segments, which are also known to introduce flexibility in the protein structures [111]. The pathways observed during MD simulations are stable over several hundreds of ns, and are sensitive to protonation state of the charged residues. Starting from the N-side of the membrane, the water wires end at conserved lysine residues in the middle of the antiporter-like subunits. In all three subunits, deprotonation of the middle lysine leads to closure of the channel by disruption of the water wire. This effect suggests that the protonation state of buried conserved residues could control changes in the hydration level of the membrane domain subunits, as also suggested for other energy converting proteins [270].

The Nqo8 subunit comprises a similar 5-TM helix bundle element, in which half of the broken helix is unstructured and forms part of the quinone binding pocket. Interestingly, we observe a similar hydration pattern as in the other antiporter-like subunits, with hydration most prominent in the region of the unstructured broken helix. This region comprises several charged residues that could be of central importance for transmission of the pumping signal to the membrane subunits [92].

To further investigate the role of the identified proton transfer pathways, we per-

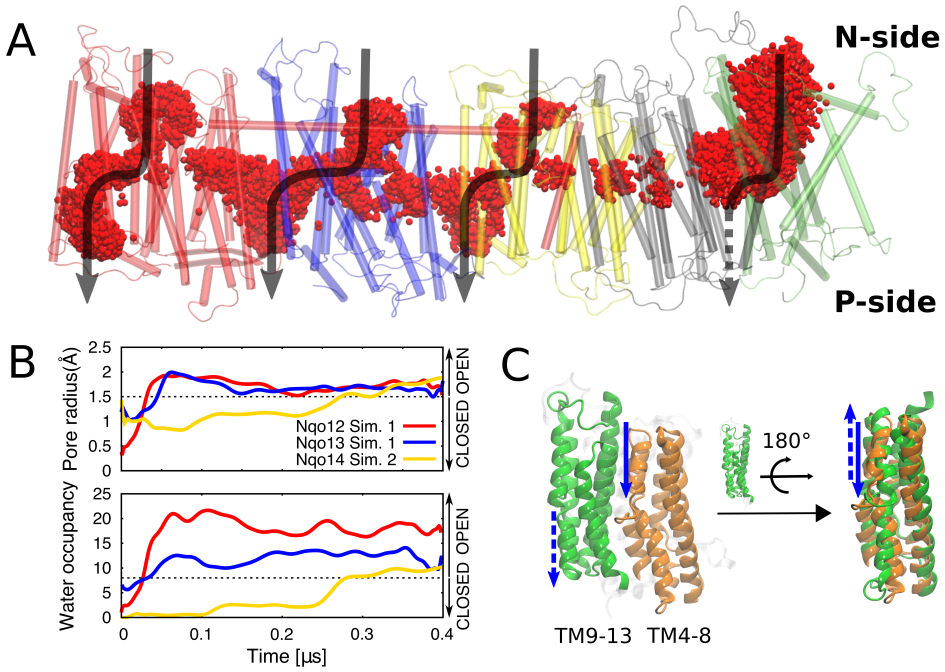


**Figure 5.2:** Topology of the proton transfer pathways in the membrane domain of complex I. A) Overlap of water molecules (in red) around key residues of the membrane domain subunits during 3 μs of MD. Pathways are shown with black arrows. B) Pore radius and water occupancy of the N-side channels of the three antiporter-like subunits Nqo12, Nqo13, Nqo14 (red, blue and yellow, respectively) during 400 ns of MD simulation. C) Symmetry of the N- and P- water channels within a single antiporter-like subunit resembles the internal symmetry of the 5-TM helical bundles (TM4-8 and TM9-13 in orange and green, respectively). The figure is adapted from publication I.

formed QM/MM simulations of the proton transfer reactions along the observed water molecules. We simulated two different proton transfer steps in the Nqo13 subunit: from the N-side bulk phase to the central conserved Lys235 residue, and a 'lateral' proton transfer step from the central lysine to the terminal Glu377 at the interface with the Nqo12 subunit. Both simulations show a rapid (ps) proton transfer, which takes place by a Grotthus-type process [271]. To test the dependence of the N-side bulk/Lys235 transfer on the hydration state of the channel, we performed QM/MM simulations starting from several hydration levels of the proton channel. We observe that the simulations that started from dry or partially hydrated states show less tendency to transfer the proton. This result suggests that proton transfer events could occur on a fast (ps) timescale, after the slow (μs) hydration event. In the proton transfer from Lys235 to Glu377, we observe the proton transfer via His292, which acts as a proton acceptor/donor. The simulations support the idea that the proton transfer is feasible in the membrane domain of complex I, and the lateral proton transfer step is relevant for the proton pumping process.

The picture emerging from the MD simulations involves three proton transfer reactions: from the N-side bulk water to the center of the antiporter-like subunits (central lysine), a second lateral proton transfer from the central lysine to the terminal charged residue, and a terminal transfer from here to the P-side of the membrane. Notably, this transfer process would position the ion pair in each antiporter-like subunit out from the proton translocation pathway. To better understand the role of the ion pair in the pumping process, we calculated free energy profiles of the Glu123/Lys204 ion pair dissociation in the Nqo13 subunit, with neutral or protonated Lys235. The simulations were started from hydrated structures obtained from MD simulations and performed using the Replica-Exchange Umbrella Sampling (REUS) method (see Chapter 3). The free energy profiles were reconstructed with WHAM. The results show a clear dependence of the dissociation profile on the state of the central Lys235 (Figure 5.3). The two profiles show a similar minimum in the associated ion pair state, whereas dissociation leads to a high increase in the free energy when Lys235 is protonated. The 'open' state has a minimum when Lys235 is deprotonated, with a barrier between open and close state of *ca.* 4 kcal mol$^{-1}$. In the open state, Glu123 interacts with the Lys345 of Nqo14 subunit, suggesting that this transition can modulate interactions between subunits.

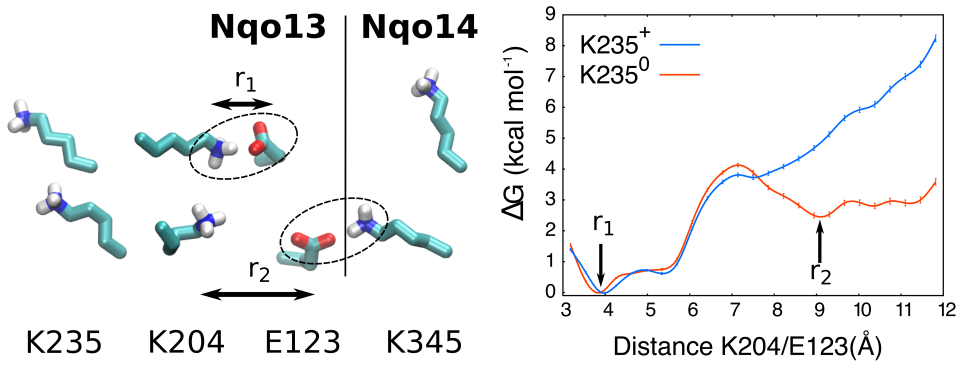


**Figure 5.3:** Free energy profiles of Lys204/Glu123 dissociation with Lys235 protonated (blue) and deprotonated (red) obtained from REUS simulations. The minimum at large distances is absent when Lys235 is protonated, possibly due to the electrostatic repulsion from Lys204. The figure is adapted from publication I.

The combined results of MD, QM/MM and REUS simulations suggest that the long-range proton transfer mechanism in complex I is achieved via a combination of hydration/dehydration of the proton channels, proton transfer reactions and open/close ion pair switching.

The channels, which form at symmetry-related positions, are sensitive to the protonation state of buried charged residues involved in the proton transfer reaction. Based on this, we proposed that the signal is transported in Nqo12/13/14 as follows:

- Initial quinone reduction that triggers the ion pair dissociation in Nqo14 subunit.

- The ion pair dissociation favors the lateral proton transfer toward the interface with the following subunit, and closure of the N-side pathway.
- Opening of the P-side channel, and triggering of the ion pair opening of the Nqo13 subunit.
- Repetition of the steps up to the Nqo12 subunit.

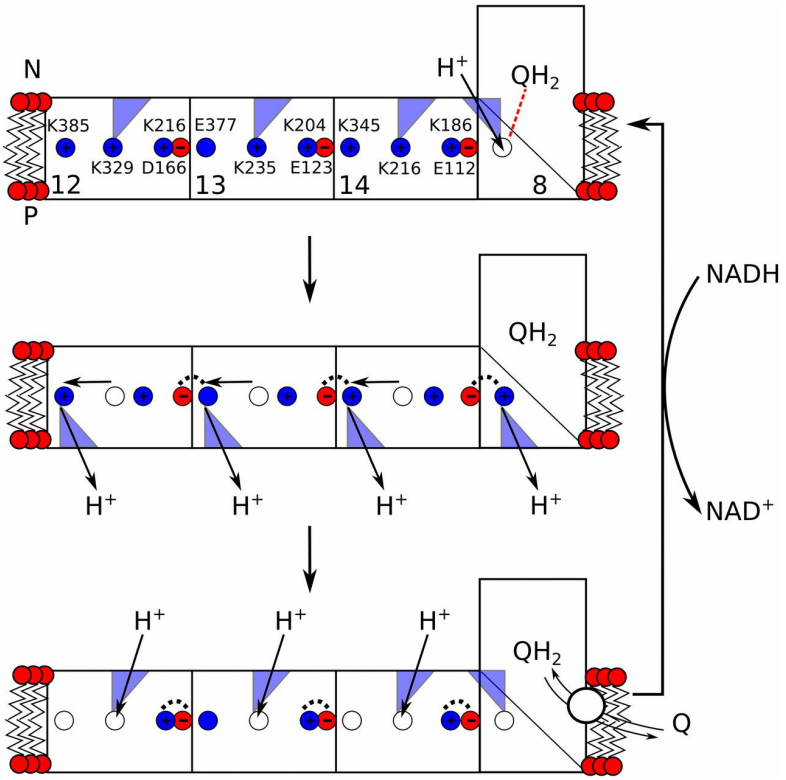


**Figure 5.4:** Proton pumping mechanism in complex I. The schematic picture outlines putative phases of the cycle. (*Top*) After Q reduction, a proton is uptaken in Nqo8 and triggers the proton pumping machinery. (*Middle*) The ion pairs opening and proton transfer events alternate and propagate the signal horizontally along the membrane domain. (*Bottom*) The enzyme restores the initial state opening the N-side channels and uptaking the proton from the bulk. The figure is adapted from publication I.

This mechanism, depicted in Figure 5.4, favours a sequential 'domino' propagation of strong local electrostatic interactions to build the large-scale transmission mechanism. It has later been suggested that release of protons to the P-side takes place during prorogation of backward wave from Nqo12 to Nqo8 [73].

### 5.1.2 Inter-Subunit Interfaces Modulate Signal Transport

In publication III, we further investigated the relationship between ion pair dynamics in the antiporter-like subunits and the proton transport mechanism. We used results from

MD simulations also used in publication I, and combined them with DFT models to understand the basic function of the electrostatic coupling mechanism.

The MD simulations suggest that the ion pairs in each antiporter-like subunit alternate between intra- and inter-subunit contacts at the interfaces Nqo12/Nqo13, Nqo13/-Nqo14 and Nqo14/Nqo11 (Figure 5.5), with the Nqo13/Nqo14 having the most evident switching behaviour. At this interface, we observe a switching dynamics of Glu123, which alternatively interacts with $Lys204_{13}$ and $Lys345_{14}$. The same residue pair is found in the open state in the crystal structure from *Thermus thermophilus*. The distance of the ion pair at the interface correlates with a change in the distance between $Lys204_{13}$ and $Lys235_{13}$, the latter of which could be the primary proton donor/acceptor in the Nqo13 subunit. The two residues also show a different average distance when simulated in their protonated state (*ca.* 14 Å), or when one of the residues is modeled in its neutral state and the other is protonated (*ca.* 10.5 Å).

At the interface between the Nqo11/Nqo14 subunits, two acidic residues ($Glu67_{11}$ and $Glu32_{11}$) are in the same region as the ion pair formed by $Glu112_{14}$ and $Lys186_{14}$. $Glu67_{11}$ undergoes a similar switching behaviour as $Glu123_{13}$, alternating interaction between $Lys186_{14}$ and $Glu32_{11}$. Similar results are also obtained for the Nqo12 subunit, where the simple ion pair motif is replaced by the two ion pairs Arg163-Glu132 and Lys216-Asp166. We find that $Arg163_{12}$ can also interact with $Glu377_{13}$, adopting an additional third intermediate conformation, which could arise from the simultaneous interaction with multiple acidic residues. The interaction of $Asp166_{12}$-$Arg163_{12}$ and $Glu377_{13}$-$Arg163_{12}$ are also anticorrelated.

We further performed PCA analysis of the MD trajectories, to identify communication channels and possible coupling effects during wet/dry transitions of the channels. Results show that the coupling between subunit is sensitive to both the hydration and protonation states, suggesting that water molecules might have an active role in modulation of the signal transport mechanism. To further probe the idea behind the 'electrostatic cascade' in antiporter-like subunits of complex I, we built DFT models consisting of the sidechain Lys-His-Lys and Lys-His-Glu and connecting water molecules, resembling the lateral proton transfer step postulated in the mechanism. To simulate the 'signal', we added a lysine side chain in a position derived from the lysine of the triggering ion pair. Despite the simplicity of the model, the results suggest that the effect of the positive charge lowers the activation barrier and makes the reaction more exergonic, for both terminal Glu and Lys residues. These simple effect could thus modulate the proton transfer energetics in complex I, and create a tightly coupled long-range proton pumping.

## 5.2 Quinone Catalysis

The quinone pocket and nearby regions form important structural elements that are key for the long-range coupling process in complex I. The following studies (publication II, V) are related to the quinone binding and diffusion in the cavity.

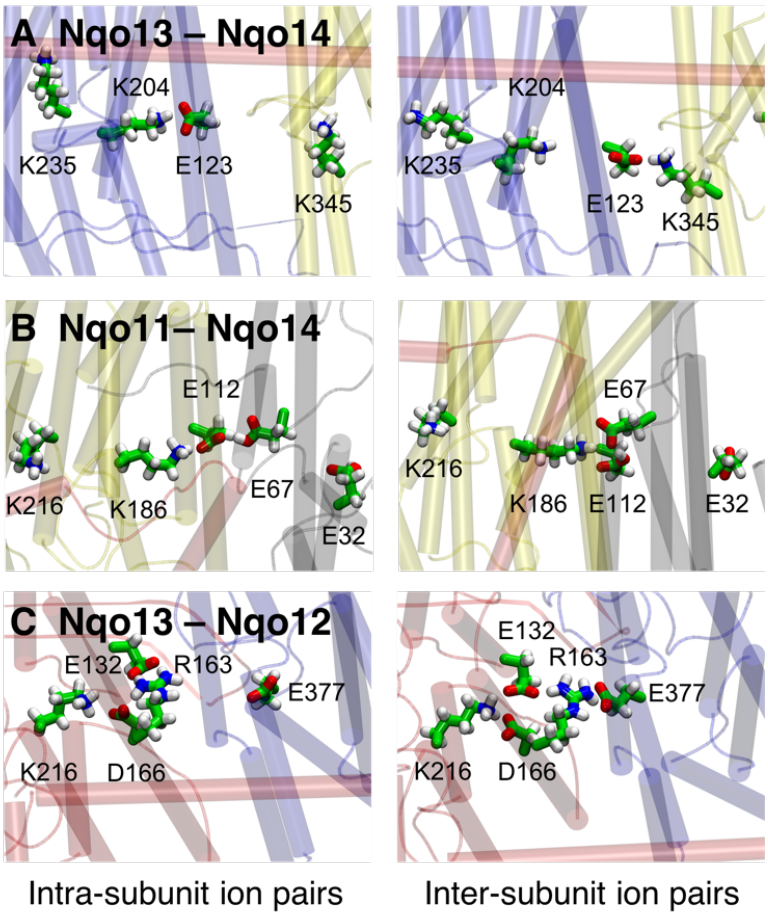


**Figure 5.5:** Interfaces between antiporter-like subunits in the membrane domain of respiratory complex I. A) Nqo13/Nqo14 , B) Nqo11/Nqo14 and C) Nqo13/Nqo12 interfaces. The conserved ion pairs form both intra- (left) and inter- (right) subunit contacts. Structures taken from snapshots of MD simulations. The figure is adapted from publication III.

### 5.2.1 Redox-Dependent Quinone Dynamics

In publication V, we show that the quinone motion in the binding pocket is modulated by its redox state. We also observe that during its diffusion along the channel, quinone finds additional binding regions in the lower part of the binding cavity.

The study was conducted using the crystal structure of complex I from *Thermus thermophilus* [36], and performing calculations with different quinone substrates with 1, 6, and 10 isoprenoid units in the tail. We performed MD simulation of the quinone in the oxidized (Q) and reduced ($QH_2$) form, and found that the latter moves spontaneously away from the primary binding site. It was also previously shown that dynamics of $QH_2$ triggers conformational changes of conserved residues in the binding pocket [92]. We further employed umbrella sampling and a one-dimensional diffusion model, both for Q and $QH_2$, to explore the energetics of the diffusion processes. To ensure convergence,

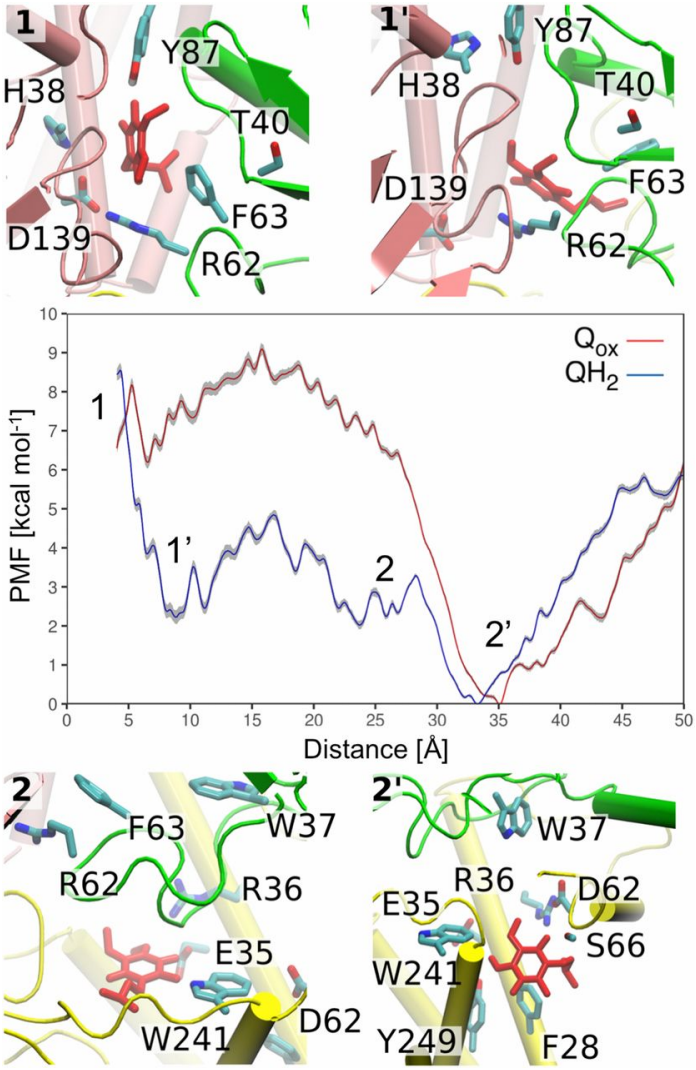


**Figure 5.6:** Free energy profiles of Q and $QH_2$ motion along the binding pocket obtained from umbrella sampling simulations. Structures of the identified binding sites are shown in 1, 1' (top) and 2, 2' (bottom). Residues surrounding the quinone (in red) are shown in sticks representation. The figure is adapted from publication V.

the umbrella sampling was performed using the short-tailed $Q_1$.

The free-energy profile as a function of the distance from Tyr87 (Figure 5.6) shows that Q has to overcome an activation barrier of roughly 8 kcal $mol^{-1}$ to move from the lower part of the cavity, about 30-35 Å from Tyr87, to its primary binding site at *ca.* 5 Å from Tyr87 (site 1). Site 1 (Figure 5.6) shows a relatively stable minimum with a barrier of about 3 kcal $mol^{-1}$ for the backward diffusion. In contrast, $QH_2$ has a low barrier to diffuse from the distal to the upper part of the cavity, but is unstable in the binding site 1 and moves to a second region slightly below (site 1', Figure 5.6). This site coincides with the refined density of a quinone inhibitor in the binding cavity of *Yarrowia lipolytica* [53],

and it might have a biological function in preventing reverse electron transfer processes.

Both Q and $QH_2$ show a second binding minimum of possible mechanistic relevance in the region around 30-35 Å. To investigate this hypothesis we performed additional MD simulations of the quinone in different position in the binding pocket, and obtained diffusion profiles from a 1-D Bayesian diffusion model as described in reference [272]. To generate the starting structure, we first performed steered MD simulations pulling the $Q/QH_2$ out of the channel. We started multiple independent simulations along the pulling coordinate and extracted information about local diffusion properties from the histogram analysis of the MD trajectories. The results are similar to the profiles shown in Figure 5.6, further supporting the presence of additional binding sites of Q and $QH_2$ in the lower part of the cavity.

In addition to the free energy calculation, we analyzed the coupling between the motion of Q and ion pairs in the hinge subunits of complex I. We find the largest difference in the interaction network between the simulations in which Q is in the first binding region (1/1'), and $QH_2$ is in the second binding region (2/2', Figure 5.6). Several ion pairs open up in the latter case, except for the $Arg36_8/Asp62_8$ ion pair. The observed changes are linked both to the redox potential of Q, which is up-shifted by *ca.* 200 mV, and protonation states of key residues as $Glu130_8$, at the end of the E-funnel.

The identified second binding region around 30 Å from the primary binding site has possible mechanistic implications. After the complete reduction of Q by the ISC chain, a coupled proton transfer reaction forms $QH_2$. This process reduces the binding free energy, and the substrate is released toward site 1'. Both the dissociation and further diffusion along the cavity are exergonic processes, which could provide the energy for the proton pumping in the membrane domain of complex I. Our calculations indicate that $QH_2$ at the second binding site could alter the protonation state of residues important for the enzyme function. These results suggest that a Q molecule could 'shuttle' between two binding sites in the pocket, undergoing repeated cycles of reduction/oxidation and activation of the proton pump. EPR data indirectly support the possibility of multiple binding sites [71, 93, 273], as well as a binding of multiple Q in the pocket.

### 5.2.2 Kinetics and Structure of Ubiquinone in the Binding Pocket

Although structural data suggests that the quinone enters/exits the binding pocket from a narrow channel in the ND1 subunit (Nqo8 in *Thermus thermophilus*) subunit, little is known about the biological timescales of binding/dissociation, and the rate limiting steps during enzyme turnover. In publication II, by studying kinetics and structure of quinones of different hydrophobic tail length, we showed that quinone diffusion in the cavity is not rate limiting for the catalysis, and additionally that $Q_{10}$ provides an optimal substrate for complex I. We could also show that charged amino acids in the kink region of the binding pocket are important for the quinone diffusion/binding process. This study was conducted in collaboration with the Judy Hirst Lab at MRC, Cambridge.

The experimental setup included the alternative oxidase (AOX), which ensured a largely oxidized quinone pool for all the substrates used in the study ($Q_1$, $Q_2$, $Q_4$, $Q_6$, $Q_8$, $Q_{10}$). Details of the setups are explained in the Methods section of publication II. By

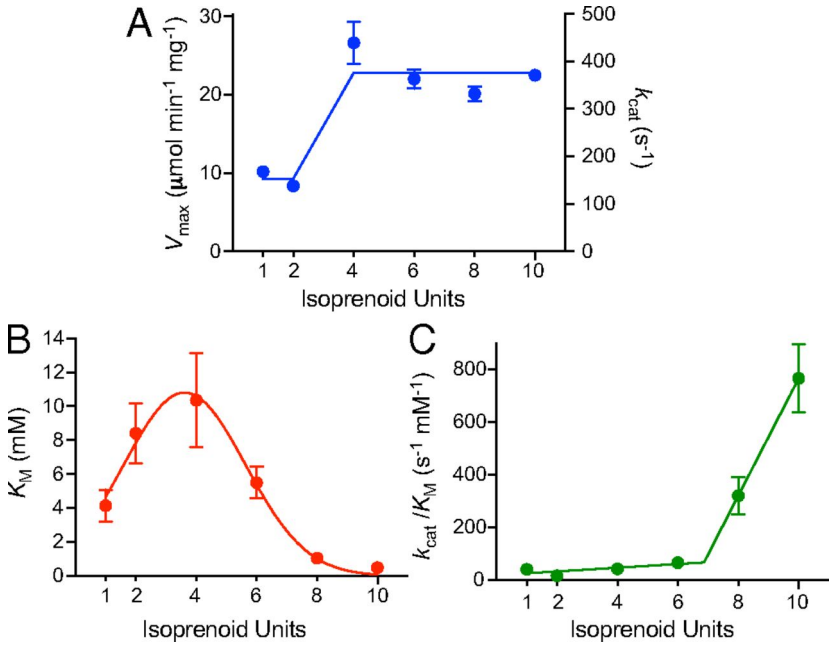


**Figure 5.7:** Comparison of the Michaelis-Menten parameters for different chain lengths. A) $V_{max}$ ($k_{cat}$), B) $K_M$ and C) $k_{cat}/K_M$ values for quinones with 1, 2, 4, 6, 8 and 10 isoprenoid units. The figure is adapted from publication II.

monitoring the NADH consumption, we obtained Michaelis-Menten curves for different Q lengths and compared their kinetic parameters (Figure 5.7). The results show strong and non-linear differences in the kinetic parameters of the quinone, depending on the tail length. $V_{max}$ has a two-state behaviour, while $K_M$ is well represented by a bell-shaped curve with the maximum around the 4-5 isoprenoid unit. The catalytic efficiency ($k_{cat}/K_M$) shows a net increase for $Q_8$ and $Q_{10}$, suggesting a clear preference of long-isoprenoid tails over short ones.

$V_{max}$ ($k_{cat}$) contains information about all the reaction steps after binding of the substrate. The constant value for $Q_4$-$Q_{10}$ indicates that the product dissociation, which is possibly dependent on the substrate length, is not rate-limiting for catalysis. In the case of the short-tailed $Q_1$ and $Q_2$, the change in $V_{max}$ indicates either that some rate-limiting steps is further slowed down, or completely changed. $k_{cat}/K_M$ increases with the number of isoprenoid units, consistent with the increasing transit rate along the cavity with the tail length.

In order to investigate possible causes for such a behaviour, we modelled the different Q substrates, and performed MD simulations of the subunits forming the binding pocket (Figure 5.8). Side chains of the surrounding residues do not show conformational changes linked to the Q-tail length except Phe224$_{ND1}$, which switches to an open conformation in case of $Q_8$ and $Q_{10}$. RMSF (Root Mean Square Fluctuation) of the isoprenoid units for the different substrates shows that $Q_{6-8-10}$ are more constrained in the cavity than $Q_{1-2-4}$ (Figure 5.8). In addition to that, the region around isoprenoid units 4-7 shows a strong kink of the quinone tail, with an angle of *ca.* 100°. The same area is rich in

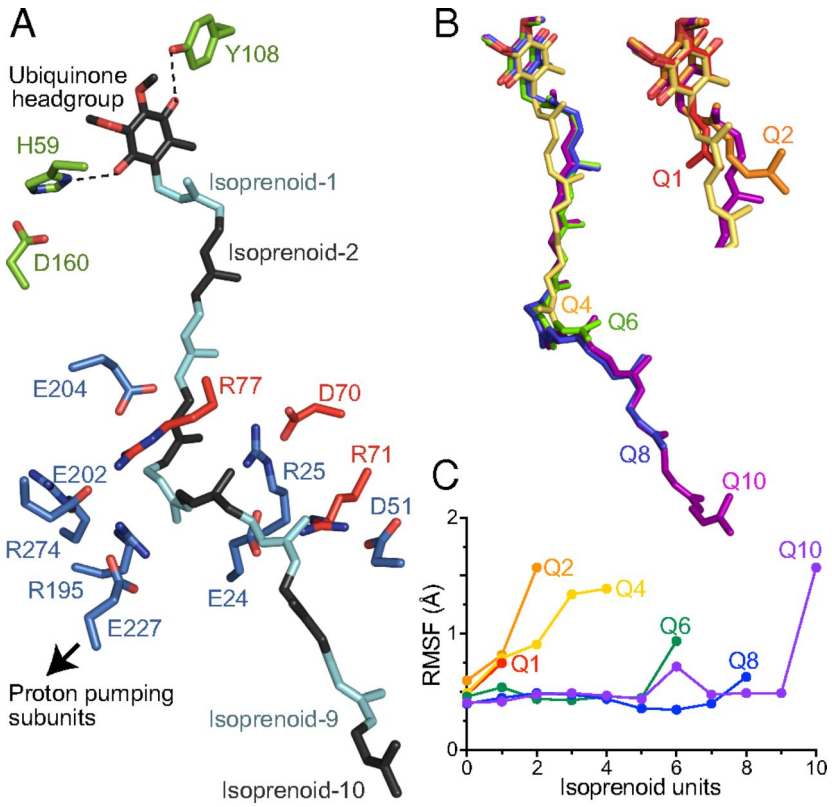


**Figure 5.8:** A) Modelled structure of $Q_{10}$ in the binding pocket and surrounding polar residues. The different isoprenoid units are highlighted in cyan and black. The residues are colored in red (PSST), blue (ND1) and green (49 kDa). B) Comparison of the relaxed tails of different quinone substrates. C) RMSF of the different isoprenoid units of the simulated subustrates. $Q_{1-2-4}$ show a higher mobility in the cavity. The figure is adapted from publication II.

conserved charged and polar amino acids, in particular arginine residues (Figure 5.8). The arginine residues interact both with the isoprenoid tail, via $\pi$-stacking interactions, and with glutamate/aspartate residues in the binding pocket, forming stable electrostatic interactions, possibly related to the proton pumping function.

The experimental evidence suggests a binding/dissociation dependent on the substrate tail length. We proposed three explanations to this effect:

- The higher mobility of residues $Q_1$, $Q_2$ and $Q_4$ hinders their dissociation, and the kink acts as a anchor for the tail.
- For shorter quinones, a second molecule could enter the pocket and block its dissociation.
- The quinone substrates distribute differently in the membrane, and their polar headgroup could be located at different depths. This effect would disfavour some substrates from entering the channel, and alter the binding kinetics.

The highly charged region around isoprenoid units 4-7 could also affect the binding of the substrate by decreasing its affinity, which could be important for the product release. Similarly, increased polarity of the pocket could attract water molecules, which would need to be displaced/released in each cycle, altering the energetics of the quinone motion. The effect of this polar region thus would contrast the two hydrophobic region at the top and bottom of the cavity, fine-tuning the mobility of the substrate needed for efficient catalysis.

## 5.3 Global Dynamics of Complex I

Recent Cryo-EM studies conducted on the mammalian complex I [54, 55] resolved different structures that were suggested to be related to biologically active (A) and deactive (D) states (see Chapter 2). Using both network models and atomistic molecular dynamics simulations, we analyzed (publication IV) the global motions of the enzyme to investigate their relationship with the A/D transition. We showed that the two lowest-frequency normal modes of complex I resemble the differences observed between the experimentally resolved structures. Additionally, we highlighted the similarity between the mammalian and bacterial enzyme, and showed the regulatory effect of the supernumerary subunits on the identified motions.

We used anisotropic network models (ANM) to calculate normal modes based on the experimental structure of *Bos tarurus* (PDB ID:5LC5), and compared them with the experimental difference between the Cryo-EM structures. The A and D states (class 2 and 1, respectively, in the original publication [54]) have been assigned based on structural features such as disordered loops of subunits ND1, 49 kDa and ND3 subunits in the active site, and on the knowledge of biochemical characterization of the states [151]. This hypothesis has later been confirmed by Cryo-EM studies [56]. A third structure, class 3, has been described as 'partially broken' due to loss of density of the lateral horizontal helix of ND5 subunit.

The results show that the first two normal modes, comprising twisting and bending motion, strongly resemble the two experimentally identified motions between class $1 \rightarrow 2$ and class $1 \rightarrow 3$, respectively (Figure 5.9). Both twisting and bending modes consist of a relative displacement of the hydrophilic and membrane domain with respect to each other, with the quinone binding region acting as a hinge for the motion. The twisting (Figure 5.9B) rotates the two domains in opposite direction along their longest axis, while bending brings the two domains closer to each other in a scissor-like motion (Figure 5.9A).

Although class 3 has been described as partially unfolded, the similarity between the bending motion, the experimental transition, and the motion described for the similar ovine enzyme [55] suggests that this motion could also be catalytically relevant. A similar result is obtained by PC analysis of the all atomistic MD simulations of the bacterial enzyme (*Thermus thermophilus*), showing similar bending and twisting motions, and suggesting an overall conserved global dynamics.

Motions of the bacterial and mammalian complex I are similar, but not identical. By selectively removing the supernumerary subunits of the mammalian enzyme, we analyzed which of them has the largest effect on the difference of the global motion of the

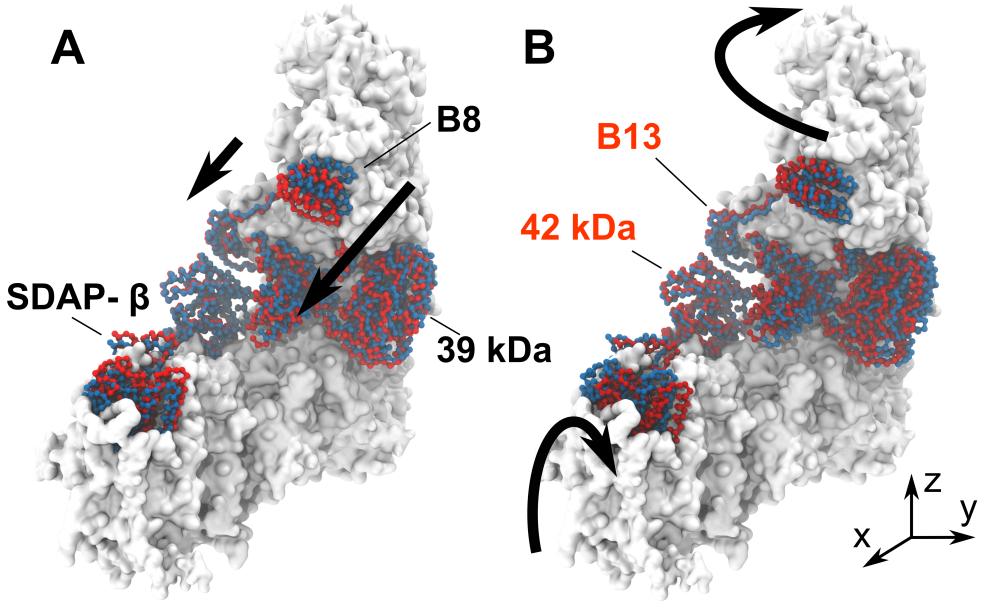


**Figure 5.9:** A) Bending and B) Twisting low-frequency motions of mammalian complex I. Displacement of the supernumerary subunits is shown in red and blue. Subunits B13 and 42 kDa regulates the enzyme dynamics. The figure is adapted from publication IV.

two enzymes. We found that the 42 kDa and B13 subunits (Figure 5.9) strongly affect the bending and twisting mdoes, making them identical to the motions of the bacterial enzyme upon removal. This result suggests that these two subunits, which are absent in the bacterial enzyme, could regulate the twisting motion and transform it in the A/D transition in the mammalian enzyme. Interestingly, *Yarrowia lipolytica*, which shows a low activation barrier of the A/D transition, also lacks the 42 kDa subunit. The results shown here were further corroborated from the more recent *Mus musculus* cryo-EM structure [57], which showed a relative displacement of the two supernumerary subunits between A and D states.

## 5.4 Effect of the Membrane Bilayer on Catalysis Regulation

As presented in Chapter 2, turnover rates of complex I strongly depend on the membrane composition. In particular, cardiolipin (CDL) has the strongest effect on complex I activity, and the purified enzyme retains *ca.* 10 CDL molecules [138]. To study how CDL influences the enzyme function, we performed coarse-grained (CG) MD simulations of the respiratory complex I from *Thermus thermophilus* embedded in a multicomponent (POPE, POPC, CDL) membrane, with and without CDL, and analyzed the effect on quinone dynamics and global motions. To further investigate the atomistic details

of lipids-protein interactions, we also performed atomistic MD simulations of the hinge region of complex I.

The CG simulations show that lipids have a different interaction time with the membrane domain of complex I: while interactions of POPC and POPE are transient and last for about 100 ns, CDL strongly binds to the protein, with average retention times up to 1.2 µs around the Nqo8 subunit. The same subunit forms a preferential binding site for CDL, interacting on average with 10-15 CDL molecules. Although the timescale of the simulation is different, we observe a similar trend in the atomistic simulations of complex I hinge region. CDL interacts with the positively charged surface of Nqo8 and neighbouring subunits, with their negatively charged head groups forming interactions with arginine side chains. An unusual amount of phenylanine residues on the surface of the Nqo8 subunit as compared to the rest of the enzyme suggests also that they might play a role in enhancing CDL binding as for other respiratory enzymes [134]. We also observe binding sites of CDL that are located at the interface between antiporter-like subunits, close to the exit of the proton pathways observed in previous simulations (publication I). Similar binding sites were recently resolved in the cryo-EM structure of the mammalian enzyme of *Mus musculus* and *Ovis aries* [55, 57].

To analyze the effect of CDL binding on the dynamics of complex I, we performed PC analysis of the CG trajectories. The two PCs with highest variance are twisting and bending modes, where Nqo8 acts as a hinge, similarly to what has been observed previously (publication IV). A third PC is strongly modulated by CDL, and consists of a coupled bending-twisting motion that displaces half of the hydrophilic and membrane domain. This motion can be linked to the presence of a CDL molecule at the interface between Nqo13 and Nqo14 subunits.

To probe the role of CDL in shaping the quinone dynamics, we analyzed the correlation between the quinone motion and global dynamics with and without CDL (Figure 5.10). We observed that the quinone finds a stable binding position at the top of the cavity, slightly below this top-part binding position, and in a lower region of the pocket. These sites correlate well with minima observed in previous atomistic simulations (publication V). We observe a correlation between the bending of complex I and quinone motion from the lower side of the cavity to the top (Figure 5.10A). At the same time, motion of the quinone along the strongly bent region in the cavity causes a transition along the twisting coordinate. When the quinone is bound to the lower side of the cavity, complex I samples conformations similar to the *apo* state of the enzyme, suggesting that the initial quinone binding does not alter the enzyme conformation. Conversely, these correlations are lost in absence of CDL, although the enzyme undergoes the same twisting and bending motions, and the quinone explores all the different binding sites in the cavity. This could indicate that, in absence of CDL, the two events are decoupled, suggesting that CDL affects complex I turnover by coordinating its global motion.

CDL also enhanced stability of cavities that lead to the quinone binding site from the bulk water (Figure 5.10C) in the simulations. One cavity leads to Tyr87 of the Nqo4 subunit, which has been indicated as one of the proton donor to the reduced quinone molecule. A second cavity connects the N-side bulk phase to the lower part of the quinone

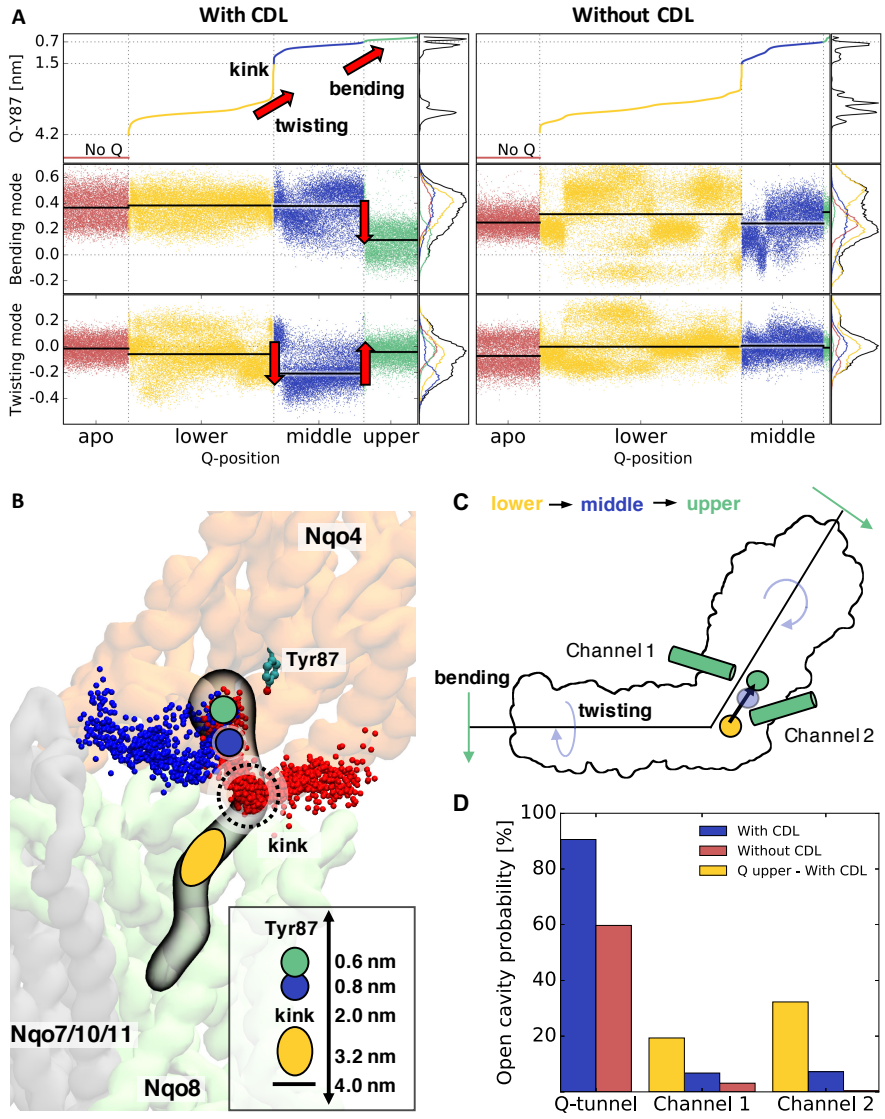


**Figure 5.10:** Modulation of complex I dynamics by CDL. A) Projection of complex I PC1 (bending) and PC2 (twisting) modes sorted by quinone (Q) position along the channel. Motion of Q is linked to global conformational changes described by PC1 and PC2. B) Binding sites of quinone: site 1 (*ca.* 0.6 nm from Tyr87, green), site 1' (*ca.* 0.8 nm from Tyr87, blue) and site 2 (*ca.* 3.5 nm from Tyr87, yellow). Channel 1 (blue) leads to the site 1, while channel 2 leads to the kink region. C) Schematic representation of the coupled quinone motions/conformational transition during dynamics. D) Channel stability dependent on absence/presence of CDL (red/blue) and quinone in site 1 (yellow). The figure is adapted from publication VI.

pocket (Figure 5.10B). Such cavities could work as proton channels to restore the protonation state of the catalytic residues in the binding pocket.

# Chapter 6

# Conclusions

In this thesis, molecular dynamics simulations were employed to investigate different aspects of the proton pumping mechanism in respiratory complex I. The work focused on three aspects: 1) long-range proton-pumping activation (publications I,III), 2) quinone dynamics in its binding pocket (publications II,V,VI), and 3) coupling of the function and global motion of the enzyme (publication IV).

Based on the simulation data, we suggested that complex I activates the proton-pumping machinery by creating a cascade of signals from the first antiporter-like subunit. This signal propagates via ion pair dissociation and proton transfer events, and involves interaction between residues of neighbouring subunits. The proton transfer channels open up on a μs timescale, and connect buried conserved charged residues to the two sides of the membrane.

We identified the binding mode of quinone, and how its diffusion to a putative second binding site might be relevant for activating the proton pumping machinery. It was found that the hydrophobic tail of the quinone could also have a role in the regulation of binding energetics and help its diffusion along the cavity.

Analysis of the global dynamics and normal modes of the enzyme showed two prominent bending and twisting motions. These two conformational transitions seem to play a role both in the mammalian active/deactive transition, and in quinone dynamics in the cavity. Comparison between bacterial and mammalian global motions highlights the effect of supernumerary subunits in the modulation of the A/D transition.

The complete complex I mechanism is still far from being completely understood. The topics presented in this thesis contribute to some of the numerous unanswered questions in the coupling mechanism of complex I. In future research, the focus will be directed to the identification of the exact coupling sites/events, which also entails the elucidation of the energy conversion from the 'chemical' electron transfer to the 'electrochemical' proton motive force.

# Appendix A

**Table 1:** Nomenclature of the core subunits of complex I in *Thermus thermophilus*, *Bos taurus* and *Mus musculus*.

| *Thermus thermophilus* | *Bos taurus* | *Mus Musculus* |
|---|---|---|
| Nqo1 | 51 kDa | NDUFV1 |
| Nqo2 | 24 kDa | NDUFV2 |
| Nqo3 | 75 kDa | NDUFS1 |
| Nqo4 | 49 kDa | NDUFS2 |
| Nqo5 | 30 kDa | NDUFS3 |
| Nqo6 | PSST | NDUFS7 |
| Nqo7 | ND3 | ND3 |
| Nqo8 | ND1 | ND1 |
| Nqo9 | TYKY | NDUFS8 |
| Nqo10 | ND6 | ND6 |
| Nqo11 | ND4L | ND4L |
| Nqo12 | ND5 | ND5 |
| Nqo13 | ND4 | ND4 |
| Nqo14 | ND2 | ND2 |

# Bibliography


[1] Dodd MS, et al. (2017) Evidence for early life in Earth's oldest hydrothermal vent precipitates. *Nature* 543:60–64.

[2] Kasting JF (1993) Earth's early atmosphere. *Science* 259:920–926.

[3] Peretó J (2005) Controversies on the origin of life. *Int Microbiol* 8:23–31.

[4] Oparin A, Morgulis S (1953) *The origin of life*. (Dover, New York), 2nd ed. edition.

[5] Buick R (2008) When did oxygenic photosynthesis evolve? *Philos Trans R Soc Lond B Biol Sci* 363:2731–2743.

[6] Babcock GT (1999) How oxygen is activated and reduced in respiration. *Proc Natl Acad Sci USA* 96:12971–12973.

[7] Mitchell P (1961) Coupling of phosporilation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* 191:144–148.

[8] Yoshida M, Muneyuki E, Hisabori T (2001) ATP synthase - a marvellous rotary engine of the cell. *Nat Rev Mol Cell Biol* 2:669–677.

[9] Saraste M (1999) Oxidative phosphorylation at the fin de siècle. *Science* 283:1488–1493.

[10] Slater EC (1953) Mechanism of phosphorylation in the respiratory chain. *Nature* 172:975–978.

[11] Mitchell P (1966) Chemiosmotic coupling in oxidative and photosynthetic phosphorylation. *Biol Rev Camb Philos Soc* 41:445–502.

[12] Lane N, Allen JF, Martin W (2010) How did LUCA make a living? Chemiosmosis in the origin of life. *BioEssays* 32:271–280.

[13] Hosler JP, Ferguson-Miller S, Mills DA (2006) Energy transduction: proton transfer through the respiratory complexes. *Annu Rev Biochem* 75:165–187.

[14] Kaila VRI, Verkhovsky MI, Wikström M (2010) Proton-coupled electron transfer in cytochrome oxidase. *Chem Rev* 110:7062–7081.

[15] Hirst J (2013) Mitochondrial complex I. *Annu Rev Biochem* 82:551–575.

[16] Ingledew WJ, Poole RK (1984) The respiratory chains of *Escherichia coli*. *Microbiol Rev* 48:222–71.

[17] Prince RC, Leslie Dutton P, Malcolm Bruce J (1983) Electrochemistry of ubiquinones: menaquinones and plastoquinones in aprotic solvents. *FEBS Lett* 160:273–276.

[18] Cecchini G (2003) Function and structure of complex II of the respiratory chain. *Annu Rev Biochem* 72:77–109.

[19] Bezawork-Geleta A, Rohlena J, Dong L, Pacak K, Neuzil J (2017) Mitochondrial complex II: at the crossroads. *Trends Biochem Sci* 42:312–325.

[20] Crofts AR (2004) The cytochrome $bc_1$ complex: function in the context of structure. *Annu Rev Physiol* 66:689–733.

[21] Bertini I, Cavallaro G, Rosato A (2006) Cytochrome *c*: occurrence and functions. *Chem Rev* 106:90–115.

[22] Mitchell P (1976) Possible molecular mechanisms of the protonmotive function of cytochrome systems. *J Theor Biol* 62:327–367.

[23] Iwata S, Saynovits M, Link TA, Michel H (1996) Structure of a water soluble fragment of the 'Rieske' iron-sulfur protein of the bovine heart mitochondrial cytochrome $bc_1$ complex determined by MAD phasing at 1.5 Å resolution. *Structure* 4:567–579.

[24] Wikström M (1977) Proton pump coupled to cytochrome *c* oxidase in mitochondria. *Nature* 266:271–273.

[25] Michel H, Behr J, Harrenga A, Kannt A (1998) Cytochrome *c* oxidase: structure and spectroscopy. *Annu Rev Biophys Biomol Struct* 27:329–356.

[26] Boyle PD, Falcone AB, Harrison WH (1954) Reversal and mechanism of oxidative phosphorylation. *Nature* 174:401–402.

[27] Boyer PD (2002) A research journey with ATP synthase. *J Biol Chem* 277:39045–39061.

[28] Stock D, Leslie AGW, Walker JE (1999) Molecular architecture of the rotary motor in ATP synthase. *Science* 286:1700–1705.

[29] Walker JE (1998) ATP synthesis by rotary catalysis. *Angew Chem Int Ed* 37:2308–2319.

[30] Milenkovic D, Blaza JN, Larsson NG, Hirst J (2017) The enigma of the respiratory chain supercomplex. *Cell Metab* 25:765–776.

[31] Fedor JG, Hirst J (2018) Mitochondrial supercomplexes do not enhance catalysis by quinone channeling. *Cell Metab* 28:525–531.

[32] Hirst J (2018) Open questions: respiratory chain supercomplexes - why are they there and what do they do? *BMC Biol* 16:111.

[33] Wu M, Gu J, Guo R, Huang Y, Yang M (2016) Structure of mammalian respiratory supercomplex $I_1III_2IV_1$. *Cell* 167:1598–1609.e10.

[34] Gu J, et al. (2016) The architecture of the mammalian respirasome. *Nature* 537:639–643.

[35] Letts JA, Fiedorczuk K, Sazanov LA (2016) The architecture of respiratory supercomplexes. *Nature* 537:644–648.

[36] Baradaran R, Berrisford JM, Minhas GS, Sazanov LA (2013) Crystal structure of the entire respiratory complex I. *Nature* 494:443–448.

[37] Suga M, et al. (2015) Native structure of photosystem II at 1.95Å resolution viewed by femtosecond X-ray pulses. *Nature* 517:99–103.

[38] Ugur I, Rutherford AW, Kaila VRI (2016) Redox-coupled substrate water reorganization in the active site of Photosystem II - the role of calcium in substrate water delivery. *Biochim Biophys Acta Bioenerg* 1857:740–748.

[39] Wikström M, Krab K, Sharma V (2018) Oxygen activation and energy conservation by cytochrome *c* oxidase. *Chem Rev* 118:2469–2490.

[40] Huang PS, Boyken SE, Baker D (2016) The coming of age of *de novo* protein design. *Nature* 537:320–327.

[41] Lifson S, Warshel A (1968) Consistent force field for calculations of conformations, vibrational Spectra, and enthalpies of cycloalkane and n-alkane molecules. *J Chem Phys* 49:5116–5129.

[42] Hirst J, Carroll J, Fearnley IM, Shannon RJ, Walker JE (2003) The nuclear encoded subunits of complex I from bovine heart mitochondria. *Biochim Biophys Acta Bioenerg* 1604:135–150.

[43] Brandt U (2006) Energy converting NADH: quinone oxidoreductase (complex I). *Annu Rev Biochem* 75:69–92.

[44] Sazanov LA (2015) A giant molecular proton pump: structure and mechanism of respiratory complex I. *Nat Rev Mol Cell Biol* 16:375–388.

[45] Moparthi VK, Hägerhäll C (2011) The evolution of respiratory chain complex I from a smaller last common ancestor consisting of 11 protein subunits. *J Mol Evol* 72:484–497.

[46] Friedrich T, et al. (2004) Attempts to define distinct parts of NADH:ubiquinone oxidoreductase (complex I). *J Bioenerg Biomembr* 25:331–337.

[47] Friedrich T, Weiss H (1997) Modular evolution of the respiratory NADH:ubiquinone oxidoreductase and the origin of its modules. *J Theor Biol* 187:529–540.

[48] Sazanov LA, Hinchliffe P (2006) Structure of the hydrophilic domain of respiratory complex I from *Thermus thermophilus*. *Science* 311:1430–1436.

[49] Efremov RG, Sazanov LA (2011) Structure of the membrane domain of respiratory complex I. *Nature* 476:414–420.

[50] Hinchliffe P, Sazanov LA (2005) Organization of iron-sulfur clusters in respiratory complex I. *Science* 309:771–774.

[51] Friedrich T, Böttcher B (2003) The gross structure of the respiratory complex I: a Lego system. *Biochim BiophysActa Bioenerg* 1608:1–9.

[52] Hunte C, Zickermann V, Brandt U (2010) Functional modules and structural basis of conformational coupling in mitochondrial complex I. *Science* 329:448–452.

[53] Zickermann V, et al. (2015) Mechanistic insight from the crystal structure of mitochondrial complex I. *Science* 347:44–49.

[54] Zhu J, Vinothkumar KR, Hirst J (2016) Structure of mammalian respiratory complex I. *Nature* 536:354–358.

[55] Fiedorczuk K, et al. (2016) Atomic structure of the entire mammalian mitochondrial complex I. *Nature* 538:406–410.

[56] Blaza JN, Vinothkumar KR, Hirst J (2018) Structure of the deactive state of mammalian respiratory complex I. *Structure* 26:312–319.

[57] Agip ANA, et al. (2018) Mitochondria in two biochemically defined states. *Nat Struct Mol Biol* 25:548–556.

[58] Parey K, et al. (2018) Cryo-EM structure of respiratory complex I at work. *eLife* 7:e39213.

[59] Agip ANA, Blaza JN, Fedor JG, Hirst J (2019) Mammalian respiratory complex I through the lens of Cryo-EM. *Annu Rev Biophys* 48:165–184.

[60] Zickermann V, et al. (2009) Architecture of complex I and its implications for electron transfer and proton pumping. *Biochim Biophys Acta Bioenerg* 1787:574–583.

[61] De Vries S, Dörner K, Strampraad MJ, Friedrich T (2015) Electron tunneling rates in respiratory complex I are tuned for efficient energy conversion. *Angew Chem Int Ed Engl* 54(9):2844–2848.

[62] Hayashi T, Stuchebrukhov AA (2010) Electron tunnelling in respiratory complex I. *Proc Natl Acad Sci USA* 107:19157–19162.

[63] Verkhovskaya ML, Belevich N, Euro L, Wikström M, Verkhovsky MI (2008) Real-time electron transfer in respiratory complex I. *Proc Natl Acad Sci USA* 105:3763–3767.

[64] Sazanov LA (2007) Respiratory complex I: mechanistic and structural insights provided by the crystal structure of the hydrophilic domain. *Biochemistry* 46:2275–2288.

[65] Birrell JA, Yakovlev G, Hirst J (2009) Reactions of the flavin mononucleotide in complex I: a combined mechanism describes NADH oxidation coupled to the reduction of $APAD^+$, ferricyanide, or molecular oxygen. *Biochemistry* 48:12005–12013.

[66] Gnandt E, Schimpf J, Harter C, Hoeser J, Friedrich T (2017) Reduction of the off-pathway iron-sulphur cluster N1a of *Escherichia coli* respiratory complex I restrains $NAD^+$ dissociation. *Sci Rep* 7:8574.

[67] Saura P, Kaila VRI (2019) Energetics and dynamics of proton coupled-electron transfer reactions in the NADH/FMN site of respiratory complex I. *JACS* 141:5710–5719.

[68] Zu Y, Di Bernardo S, Yagi T, Hirst J (2002) Redox properties of the [2Fe-2S] center in the 24 kDa (NQO2) subunit of NADH:Ubiquinone oxidoreductase (complex I). *Biochemistry* 41:10056–10069.

[69] Moser CC, Keske JM, Warncke K, Farid RS, Dutton PL (1992) Nature of biological electron transfer. *Nature* 355:796–802.

[70] Pohl T, et al. (2007) Iron-sulfur cluster N7 of the NADH:ubiquinone oxidoreductase (complex I) is essential for stability but not involved in electron transfer. *Biochemistry* 46:6588–6596.

[71] Yano T, Dunham WR, Ohnishi T (2005) 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* 44:1744–1754.

[72] Verkhovskaya M, Wikström M (2014) Oxidoreduction properties of bound ubiquinone in complex I from *Escherichia coli*. *Biochim Biophys Acta Bioenerg* 1837:246–250.

[73] Kaila VRI (2018) Long-range proton-coupled electron transfer in biological energy conversion: towards mechanistic understanding of respiratory complex I. *J Royal Soc Interface* 15:20170916.

[74] Gamiz-Hernandez AP, Jussupow A, Johansson MP, Kaila VRI (2017) Terminal electron-proton transfer dynamics in the quinone reduction of respiratory complex I. *JACS* 139:16282–16288.

[75] Gutman M, Singer TP, Beinert H (1972) Relation of the respiratory chain-linked reduced nicotinamide-adenine dinucleotide dehydrogenase to energy-coupling site 1. *Biochemistry* 11:556–562.

[76] Ingledew WJ, Ohnishi T (1980) An analysis of some thermodynamics properties of iron-sulphur centres in site I of mitochondria. *Biochem J* 186:111–117.

[77] Breton NL, et al. (2017) Using hyperfine electron paramagnetic resonance spectroscopy to define the proton-coupled electron transfer reaction at Fe−S cluster N2 in respiratory complex I. *JACS* 139:16319–16326.

[78] Hirst J, Roessler MM (2016) Energy conversion, redox catalysis and generation of reactive oxygen species by respiratory complex I. *Biochim Biophys Acta Bioenerg* 1857:872–883.

[79] Lambert AJ, Brand MD (2004) Superoxide production by NADH:ubiquinone oxidoreductase (complex I) depends on the pH gradient across the mitochondrial inner membrane. *Biochem J* 382:511–517.

[80] Pryde KR, Hirst J (2011) Superoxide is produced by the reduced flavin in mitochondrial complex I: a single, unified mechanism that applies during both forward and reverse electron transfer. *J Biol Chem* 286:18056–18065.

[81] Robb EL, et al. (2018) Control of mitochondrial superoxide production by reverse electron transport at complex I. *J Biol Chem* 293:9869–9879.

[82] Murphy M (2009) How mitochondria produce reactive oxygen species. *Biochem J* 417:1–13.

[83] Esterházy D, King MS, Yakovlev G, Hirst J (2008) Production of reactive oxygen species by complex I (NADH:ubiquinone oxidoreductase) from *Escherichia coli* and comparison to the enzyme from mitochondria. *Biochemistry* 47:3964–3971.

[84] Finkel T (2011) Signal transduction by reactive oxygen species. *J Cell Biol* 194:7–15.

[85] Li C, Jackson RM (2002) Reactive species mechanisms of cellular hypoxia-reoxygenation injury. *Am J Physiol Cell Physiol* 282:C227–C241.

[86] Stowe DF, Camara AKS (2009) Mitochondrial reactive oxygen species production in excitable cells: modulators of mitochondrial and cell function. *Antioxid Redox Signal* 11:1373–1414.

[87] Vinogradov AD, Grivennikova VG (2016) Oxidation of NADH and ROS production by respiratory complex I. *Biochim Biophys Acta Bioenerg* 1857:863–871.

[88] Tocilescu MA, Zickermann V, Zwicker K, Brandt U (2010) Quinone binding and reduction by respiratory complex I. *Biochim Biophys Acta Bioenerg* 1797:1883–1890.

[89] Zickermann V, et al. (2003) Functional implications from an unexpected position of the 49-kDa subunit of NADH:ubiquinone oxidoreductase. *J Biol Chem* 278:29072–29078.

[90] Masuya T, Murai M, Ifuku K, Morisaka H, Miyoshi H (2014) Site-specific chemical labeling of mitochondrial respiratory complex I through ligand-directed tosylate chemistry. *Biochemistry* 53:2307–2317.

[91] Sinha PK, et al. (2015) Conserved amino acid residues of the NuoD segment important for structure and function of *Escherichia coli NDH-1* (complex I). *Biochemistry* 54:753–764.

[92] Sharma V, et al. (2015) Redox-induced activation of the proton pump in the respiratory complex I. *Proc Natl Acad Sci USA* 112:11571–11576.

[93] Verkhovsky M, Bloch DA, Verkhovskaya M (2012) Tightly-bound ubiquinone in the *Escherichia coli* respiratory Complex I. *Biochim Biophys Acta Bioenerg* 1817:1550–1556.

[94] Ohnishi T, Ohnishi ST, Salerno JC (2018) Five decades of research on mitochondrial NADH-quinone oxidoreductase (complex I). *Biol Chem* 399:1249–1264.

[95] Euro L, Belevich G, Verkhovsky MI, Wikström M, Verkhovskaya M (2008) Conserved lysine residues of the membrane subunit NuoM are involved in energy conversion by the proton-pumping NADH:ubiquinone oxidoreductase (complex I). *Biochim Biophys Acta Bioenerg* 1777:1166–1172.

[96] Ripple MO, Kim N, Springett R (2013) Mammalian complex I pumps 4 protons per 2 electrons at high and physiological proton motive force in living cells. *J Biol Chem* 288:5374–5380.

[97] Jones AJY, Blaza JN, Varghese F, Hirst J (2017) Respiratory complex I in *Bos taurus* and *Paracoccus denitrificans* pumps four protons across the membrane for every NADH oxidized. *J Biol Chem* 292:4987–4995.

[98] Wikström M, Hummer G (2012) Stoichiometry of proton translocation by respiratory complex I and its mechanistic implications. *Proc Natl Acad Sci USA* 109:4431–4436.

[99] Fearnley IM, Walker JE (1992) Conservation of sequences of subunits of mitochondrial complex I and their relationships with other proteins. *Biochim Biophys Acta Bioenerg* 1140:105–134.

[100] Mathiesen C, Hägerhäll C (2002) Transmembrane topology of the NuoL, M and N subunits of NADH:quinone oxidoreductase and their homologues among membrane-bound hydrogenases and bona fide antiporters. *Biochim Biophys Acta Bioenerg* 1556:121–132.

[101] Sato M, Torres-Bacete J, Sinha PK, Matsuno-Yagi A, Yagi T (2014) Essential regions in the membrane domain of bacterial complex I (NDH-1): the machinery for proton translocation. *J Bioenerg Biomembr* 46:279–287.

[102] Kaila VRI, Wikström M, Hummer G (2014) Electrostatics, hydration, and proton transfer dynamics in the membrane domain of respiratory complex I. *Proc Natl Acad Sci USA* 111:6988–6993.

[103] Nakamaru-Ogiso E, et al. (2004) Functional roles of four conserved charged residues in the membrane domain subunit NuoA of the proton-translocating NADH-quinone Ooxidoreductase from *Escherichia coli*. *J Biol Chem* 279:32360–32366.

[104] Kao MC, Nakamaru-Ogiso E, Matsuno-Yagi A, Yagi T (2005) Characterization of the membrane domain subunit NuoK (ND4L) of the NADH-quinone oxidoreductase from *Escherichia coli*. *Biochemistry* 44:9545–9554.

[105] Torres-Bacete J, Sinha PK, Castro-Guerrero N, Matsuno-Yagi A, Yagi T (2009) Features of subunit NuoM (ND4) in *Escherichia coli* NDH-1. Topology and implication of conserved Glu144 for coupling site 1. *J Biol Chem* 284:33062–33069.

[106] Torres-Bacete J, Nakamaru-Ogiso E, Matsuno-Yagi A, Yagi T (2007) Characterization of the NuoM (ND4) subunit in *Escherichia coli* NDH-1 . *J Biol Chem* 282:36914–36922.

[107] Nakamaru-Ogiso E, et al. (2010) The membrane subunit NuoL (ND5) is involved in the indirect proton pumping mechanism of *Escherichia coli* complex I. *J Biol Chem* 285:39070–39078.

[108] Sato M, Sinha PK, Torres-Bacete J, Matsuno-Yagi A, Yagi T (2013) Energy transducing roles of antiporter-like subunits in *Escherichia coli* NDH-1 with main focus on subunit NuoN (ND2). *J Biol Chem* 288:24705–24716.

[109] Michel J, DeLeon-Rangel J, Zhu S, Van Ree K, Vik SB (2011) Mutagenesis of the L, M, and N Subunits of complex I from *Escherichia coli* Indicates a common role in function. *PLoS ONE* 6:e17420.

[110] Steimle S, et al. (2012) Asp563 of the horizontal helix of subunit NuoL is involved in proton translocation by the respiratory complex I. *FEBS Lett* 586:699–704.

[111] Screpanti E, Hunte C (2007) Discontinuous membrane helices in transport proteins and their correlation with function. *J Struct Biol* 159:261–267.

[112] Forrest LR (2015) Structural symmetry in membrane proteins. *Annu Rev Biophys* 44:311–337.

[113] Belevich G, Knuuti J, Verkhovsky MI, Wikström M, Verkhovskaya M (2011) Probing the mechanistic role of the long $\alpha$-helix in subunit L of respiratory complex I from *Escherichia coli* by site-directed mutagenesis. *Mol Microbiol* 82:1086–1095.

[114] Zhu S, Vik SB (2015) Constraining the lateral helix of respiratory complex I by cross-linking does not impair enzyme activity or proton translocation. *J Biol Chem* 290:20761–20773.

[115] Cabrera-Orefice A, et al. (2018) Locking loop movement in the ubiquinone pocket of complex I disengages the proton pumps. *Nat Commun* 9:4500.

[116] Brandt U (2011) A two-state stabilization-change mechanism for proton-pumping complex I. *Biochim Biophys Acta Bioenerg* 1807:1364–1369.

[117] Ohnishi ST, Salerno JC, Ohnishi T (2010) Possible roles of two quinone molecules in direct and indirect proton pumps of bovine heart NADH-quinone oxidoreductase (complex I). *Biochim Biophys Acta Bioenerg* 1797:1891–1893.

[118] Verkhovskaya M, Bloch DA (2013) Energy-converting respiratory complex I: on the way to the molecular mechanism of the proton pump. *Int J Biochem Cell Biol* 45:491–511.

[119] Efremov RG, Sazanov LA (2012) The coupling mechanism of respiratory complex I - A structural and evolutionary perspective. *Biochim Biophys Acta Bioenerg* 1817:1785–1795.

[120] Hirst J (2009) Towards the molecular mechanism of respiratory complex I. *Biochem J* 425:327–339.

[121] Berrisford JM, Sazanov LA (2009) Structural basis for the mechanism of respiratory complex I. *J Biol Chem* 284:29773–29783.

[122] Hummer G, Wikström M (2016) Molecular simulation and modeling of complex I. *Biochim Biophys Acta Bioenerg* 1857:915–921.

[123] Haapanen O, Sharma V (2017) Role of water and protein dynamics in proton pumping by respiratory complex I. *Sci Rep* 7:7747.

[124] Di Luca A, Gamiz-Hernandez AP, Kaila VRI (2017) Symmetry-related proton transfer pathways in respiratory complex I. *Proc Natl Acad Sci USA* 114:E6314–E6321.

[125] Phillips R, Ursell T, Wiggins P, Sens P (2009) Emerging roles for lipids in shaping membrane-protein function. *Nature* 459:379–385.

[126] Paradies G, Paradies V, De Benedictis V, Ruggiero FM, Petrosillo G (2014) Functional role of cardiolipin in mitochondrial bioenergetics. *Biochim Biophys Acta Bioenerg* 1837:408–417.

[127] Fry M, Green DE (1981) Cardiolipin requirement for electron transfer in complex I and III of the mitochondrial respiratory chain. *J Biol Chem* 256:1874–1880.

[128] Dröse S, Zwicker K, Brandt U (2002) Full recovery of the NADH:ubiquinone activity of complex I (NADH:ubiquinone oxidoreductase) from *Yarrowia lipolytica* by the addition of phospholipids. *Biochim Biophys Acta Bioenerg* 1556:65–72.

[129] Paradies G, Petrosillo G, Pistolese M, Ruggiero FM (2002) Reactive oxygen species affect mitochondrial electron transport complex I activity through oxidative cardiolipin damage. *Gene* 286:135–141.

[130] Gomez B, Robinson NC (1999) Phospholipase digestion of bound cardiolipin reversibly inactivates bovine cytochrome $bc_1$. *Biochemistry* 38:9031–9038.

[131] Lange C, Nett JH, Trumpower BL, Hunte C (2001) Specific roles of protein-phospholipid interactions in the yeast cytochrome $bc_1$ complex structure. *EMBO J* 20:6591–6600.

[132] Robinson NC (1993) Functional binding of cardiolipin to cytochrome *c* oxidase. *J Bioenerg Biomembr* 25:153–163.

[133] Eble K, Coleman B, Hantgan RR, Cunningham CC (1990) Tightly associated cardiolipin in the bovine heart mitochondrial synthase as analyzed by $^{31}$P nuclear magnetic resonance spectroscopy. *J Biol Chem* 265:19434–19440.

[134] Arnarez C, Mazat JP, Elezgaray J, Marrink SJ, Periole X (2013) Evidence for cardiolipin binding sites on the membrane-exposed surface of the cytochrome $bc_1$. *JACS* 135:3112–3120.

[135] Arnarez C, Marrink SJ, Periole X (2013) Identification of cardiolipin binding sites on cytochrome *c* oxidase at the entrance of proton channels. *Sci Rep* 3:1263.

[136] Duncan AL, Robinson AJ, Walker JE (2016) Cardiolipin binds selectively but transiently to conserved lysine residues in the rotor of metazoan ATP synthases. *Proc Natl Acad Sci USA* 113:8687–8692.

[137] Mehdipour AR, Hummer G (2016) Cardiolipin puts the seal on ATP synthase. *Proc Natl Acad Sci USA* 113:8568–8570.

[138] Sharpley MS, Shannon RJ, Draghi F, Hirst J (2006) Interactions between phospholipids and NADH:ubiquinone oxidoreductase (complex I) from bovine mitochondria. *Biochemistry* 45:241–248.

[139] Wirth C, Brandt U, Hunte C, Zickermann V (2016) Structure and function of mitochondrial complex I. *Biochim Biophys Acta Bioenerg* 1857:902–914.

[140] Angerer H, et al. (2011) A scaffold of accessory subunits links the peripheral arm and the distal proton pumping module of mitochondrial complex I. *Biochem J* 437:279–288.

[141] Finel M, Skehel JM, Albracht SPJ, Fearnley IM, Walker JE (1992) Resolution of NADH:ubiquinone oxidoreductase from bovine heart mitochondria into two subcomplexes, one of which contains the redox centers of the enzyme. *Biochemistry* 31:11425–11434.

[142] Carroll J, Fearnley IM, Shannon RJ, Hirst J, Walker JE (2003) Analysis of the subunit composition of complex I from bovine heart mitochondria. *Mol Cell Proteomics* 2:117–126.

[143] Elurbe DM, Huynen MA (2016) The origin of the supernumerary subunits and assembly factors of complex I: a treasure trove of pathway evolution. *Biochim Biophys Acta Bioenerg* 1857:971–979.

[144] Vinothkumar KR, Zhu J, Hirst J (2014) Architecture of mammalian respiratory complex I. *Nature* 515:80–84.

[145] Guénebaut V, Schlitt A, Weiss H, Leonard K, Friedrich T (1998) Consistent structure between bacterial and mitochondrial NADH:ubiquinone oxidoreductase (complex I). *J Mol Biol* 276:105–112.

[146] Stroud DA, et al. (2016) Accessory subunits are integral for assembly and function of human mitochondrial complex I. *Nature* 538:123–126.

[147] Kmita K, et al. (2015) Accessory NUMM (NDUFS6) subunit harbors a Zn-binding site and is essential for biogenesis of mitochondrial complex I. *Proc Natl Acad Sci USA* 112:5685–5690.

[148] Ushakova AV, Duarte M, Vinogradov AD, Videira A (2005) The 29.9 kDa subunit of mitochondrial complex I is involved in the enzyme active/de-active transitions. *J Mol Biol* 351:327–333.

[149] Di Luca A, Kaila VRI (2018) Global collective motions in the mammalian and bacterial respiratory complex I. *Biochim Biophys Acta Bioenerg* 1859:326–332.

[150] Schilling B, et al. (2005) Mass spectrometric identification of a novel phosphorylation site in subunit NDUFA10 of bovine mitochondrial complex I. *FEBS Lett* 579:2485–2490.

[151] Babot M, Birch A, Labarbuta P, Galkin A (2014) Characterisation of the active/de-active transition of mitochondrial complex I. *Biochim Biophys Acta Bioenerg* 1837:1083–1092.

[152] Maklashina E, Kotlyar AB, Cecchini G (2003) Active/de-active transition of respiratory complex I in bacteria, fungi, and animals. *Biochim Biophys Acta Bioenerg* 1606:95–103.

[153] Minakami S, Schindler FJ, Estabrook RW (1964) Hydrogen transfer between reduced diphosphopyridine nucleotide dehydrogenase and the respiratory chain. *J Biol Chem* 239:2049–2054.

[154] Slater CE (1950) The dihydrocozymase-cytochrome *c* reductase activity of heart-muscle preparation. *Biochem J* 46:499–503.

[155] Vinogradov AD (1998) Catalytic properties of the mitochondrial NADH–ubiquinone oxidoreductase (complex I) and the pseudo-reversible active/inactive enzyme transition. *Biochim Biophys Acta Bioenerg* 1364:169–185.

[156] Vinogradov AD, Grivennikova VG (2001) The mitochondrial complex I: progress in understanding of catalytic properties. *IUBMB Life* 52:129–134.

[157] Roberts PG, Hirst J (2012) The deactive form of respiratory complex I from mammalian mitochondria is a $Na^+/H^+$ antiporter. *J Biol Chem* 287:34743–34751.

[158] Kotlyar AB, Vinogradov AD (1990) Slow active/inactive transition of the mitochondrial NADH-ubiquinone reductase. *Biochim Biophys Acta Bioenerg* 1019:151–158.

[159] Kotlyar AB, Sled VD, Burbaev DS, Moroz IA, Vinogradov AD (1990) Coupling site I and the rotenone-sensitive ubisemiquinone in tightly coupled submitochondrial particles. *FEBS Lett* 264:17–20.

[160] Grivennikova VG, Kapustin AN, Vinogradov AD (2001) Catalytic activity of NADH-ubiquinone oxidoreductase (complex I) in intact mitochondria: evidence for the slow active/inactive transition. *J Biol Chem* 276:9038–9044.

[161] Kotlyar AB, Sled VD, Vinogradov AD (1992) Effect of $Ca^{2+}$ ions on the slow active/inactive transition of the mitochondrial NADH-ubiquinone reductase. *Biochim Biophys Acta Bioenerg* 1098:144–150.

[162] Gostimskaya IS, Cecchini G, Vinogradov AD (2006) Topography and chemical reactivity of the active-inactive transition-sensitive SH-group in the mitochondrial NADH:ubiquinone oxidoreductase (complex I). *Biochim Biophys Acta Bioenerg* 1757:1155–1161.

[163] Galkin A, et al. (2008) Identification of the mitochondrial ND3 subunit as a structural component involved in the active/deactive enzyme transition of respiratory complex I. *J Biol Chem* 283:20907–20913.

[164] Ciano M, Fuszard M, Heide H, Botting CH, Galkin A (2013) Conformation-specific crosslinking of mitochondrial complex I. *FEBS Lett* 587:867–872.

[165] Babot M, et al. (2014) ND3, ND1 and 39 kDa subunits are more exposed in the de-active form of bovine mitochondrial complex I. *Biochim Biophys Acta Bioenerg* 1837:929–939.

[166] Ponder JW, Case DA (2003) Force Fields for Protein Simulations. *Adv Protein Chem* 66:27–85.

[167] Slater JC, Kirkwood JG (1931) The Van der Waals forces in gases. *Phys Rev* 37:682–697.

[168] Jones JE (1924) On the determination of molecular fields. I. From the variation of the viscosity of a gas with temperature. *Proc Royal Soc A* 106:441–462.

[169] Baker CM (2015) Polarizable force fields for molecular dynamics simulations of biomolecules. *Wiley Interdiscip Rev Comput Mol Sci* 5:241–254.

[170] Rick SW, Stuart SJ, Berne BJ (1994) Dynamical fluctuating charge force fields: application to liquid water. *J Chem Phys* 101:6141–6156.

[171] Lemkul JA, Huang J, Roux B, Mackerell AD (2016) An empirical polarizable force field based on the classical drude oscillator model: development history and recent applications. *Chem Rev* 116:4983–5013.

[172] Ponder JW, et al. (2010) Current status of the AMOEBA polarizable force field. *J Phys Chem B* 114:2549–2564.

[173] Mao Y, Demerdash O, Head-Gordon M, Head-Gordon T (2016) Assessing ion-water interactions in the AMOEBA force field using energy decomposition analysis of electronic structure calculations. *J Chem Theory Comput* 12:5422–5437.

[174] Darden T, York D, Pedersen L (1993) Particle mesh Ewald: an Nlog(N) method for Ewald sums in large systems. *J Chem Phys* 98:10089–10092.

[175] MacKerell AD, et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102:3586–3616.

[176] Best RB, et al. (2012) Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone $\phi$, $\psi$ and side chain $\chi_1$ and $\chi_2$ dihedral angles. *J Chem Theory Comput* 8:3257–3273.

[177] Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general Amber force field. *J Comput Chem* 25:1157–1174.

[178] Scott WR, et al. (1999) The GROMOS biomolecular simulation program package. *J Phys Chem A* 103:3596–3607.

[179] Lane TJ, Shukla D, Beauchamp KA, Pande VS (2013) To milliseconds and beyond: challenges in the simulation of protein folding. *Curr Opin Struct Biol* 23:58–65.

[180] Marrink SJ, Risselada HJ, Yefimov S, Tieleman DP, de Vries AH (2007) The MARTINI force field: coarse grained model for biomolecular simulations. *J Phys Chem B* 111:7812–7824.

[181] de Jong DH, et al. (2015) Atomistic and coarse grain topologies for the cofactors associated with the photosystem II core complex. *J Phys Chem B* 119:7791–7803.

[182] Periole X, Cavalli M, Marrink SJ, Ceruso MA (2009) Combining an elastic network with a coarse-grained molecular force field: structure, dynamics, and intermolecular recognition. *J Chem Theory Comput* 5:2531–2543.

[183] Whisstock JC, Lesk AM (2003) Prediction of protein function from protein sequence and structure. *Q Rev Biophys* 36:307–340.

[184] Meireles L, Gur M, Bakan A, Bahar I (2011) Pre-existing soft modes of motion uniquely defined by native contact topology facilitate ligand binding to proteins. *Protein Sci* 20:1645–1658.

[185] Echave J (2012) Why are the low-energy protein normal modes evolutionarily conserved? *Pure Appl Chem* 84:1931–1937.

[186] Tirion MM (1996) Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys Rev Letters* 77:1905–1908.

[187] López-Blanco JR, Chacón P (2016) New generation of elastic network models. *Curr Opin Struct Biol* 37:46–53.

[188] Atilgan AR, et al. (2001) Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys J* 80:505–515.

[189] Bahar I, Atilgan AR, Erman B (1997) Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold Des* 2:173–181.

[190] Haliloglu T, Bahar I, Erman B (1997) Gaussian dynamics of folded proteins. *Phys Rev Letters* 79:3090–3093.

[191] Tama F, Sanejouand YH (2001) Conformational change of proteins arising from normal mode calculations. *Protein Eng* 14:1–6.

[192] Zheng W, Brooks BR, Hummer G (2007) Protein conformational transitions explored by mixed elastic network models. *Proteins* 69:43–57.

[193] Maragakis P, Karplus M (2005) Large amplitude conformational change in proteins explored with a plastic network model: adenylate kinase. *J Mol Biol* 352:807–822.

[194] Pu J, Karplus M (2008) How subunit coupling produces the $\gamma$-subunit rotary motion in $F_1$-ATPase. *Proc Natl Acad Sci USA* 105:1192–1197.

[195] Hohenberg P, Kohn W (1964) Inhomogeneous electron gas. *Phys Rev* 136:864–871.

[196] Kohn W, Sham LJ (1965) Self-consistent equations including exchange and correlation effects. *Phys Rev* 140:A1133–A1138.

[197] Lee C, Yang W, Parr RG (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys Rev B* 37:785–789.

[198] Becke AD (1993) Density-functional thermochemistry. III. The role of exact exchange. *J Chem Phys* 98:5648–5652.

[199] Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ (1994) Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J Phys Chem* 98:11623–11627.

[200] Becke AD (1988) Density functional exchange-energy approximation with correct asymptotic behavior. *Phys Rev A* 36:3098–3100.

[201] Perdew JP (1986) Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys Rev B* 33:8822–8824.

[202] Grimme S, Antony J, Ehrlich S, Krieg H (2010) A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J Chem Phys* 132:154104.

[203] Warshel A, Levitt M (1976) Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J Mol Biol* 103:227–249.

[204] Senn HM, Thiel W (2009) QM/MM methods for biomolecular systems. *Angew Chem Int Ed* 48:1198–1229.

[205] Amara P, Field MJ (2003) Evaluation of an ab initio quantum mechanical/molecular mechanical hybrid-potential link-atom method. *Theor Chem Acc* 109:43–52.

[206] Gao J, Amara P, Alhambra C, Field MJ (1998) A generalized hybrid orbital (GHO) method for the treatment of boundary atoms in combined QM/MM calculations. *J Phys Chem A* 102:4714–4721.

[207] Lin H, Truhlar DG (2007) QM/MM: What have we learned, where are we, and where do we go from here? *Theor Chem Acc* 117:185–199.

[208] Senn HM, Thiel W (2007) QM/MM studies of enzymes. *Curr Opin Chem Biol* 11:182–187.

[209] Verlet L (1967) Computer "Experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys Rev* 159:98–103.

[210] Swope WC, Andersen HC, Berens PH, Wilson KR (1982) A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: application to small water clusters. *J Chem Phys* 76:637–649.

[211] van Gunsteren WF, Berendsen HJC (1988) A leap-frog algorithm for stochastic dynamics. *Mol Simulat* 1:173–185.

[212] Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* 23:327–341.

[213] Hess B, Bekker H, Berendsen HJ, Fraaije JG (1997) LINCS: a linear constraint solver for molecular simulations. *J Comput Chem* 18:1463–1472.

[214] Oostenbrink C, Villa A, Mark AE, nan Gunsteren WF (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem* 25:1656–1676.

[215] Woodcock LV (1971) Isothermal molecular dynamics calculations for liquid salts. *Chem Phys Lett* 10:257–261.

[216] Evans DJ, Holian BL (1985) The Nose-Hoover thermostat. *J Chem Phys* 83:4069–4074.

[217] Berendsen HJC, Postma JP, van Gunsteren WF, Di Nola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. *J Chem Phys* 81:3684–3690.

[218] Feller SE, Zhang Y, Pastor RW, Brooks BR (1995) Constant pressure molecular dynamics simulation: the Langevin piston method. *J Chem Phys* 103:4613–4621.

[219] Kirkwood JG (1935) Statistical mechanics of fluid mixtures. *J Chem Phys* 3:300–313.

[220] Zwanzig RW (1954) High-temperature equation of state by a perturbation method. I. Nonpolar gases. *J Chem Phys* 22:1420–1426.

[221] Kollman P (1993) Free energy calculations: applications to chemical and biochemical phenomena. *Chem Rev* 93:2395–2417.

[222] Sugita Y, Okamoto Y (1999) Replica exchange molecular dynamics method for protein folding. *Chem Phys Lett* 314:141–151.

[223] Sugita Y, Kitao A, Okamoto Y (2000) Multidimensional replica-exchange method for free-energy calculations. *J Chem Phys* 113:6042–6051.

[224] Murata K, Sugita Y, Okamoto Y (2004) Free energy calculations for DNA base stacking by replica-exchange umbrella sampling. *Chem Phys Lett* 385:1–7.

[225] Kästner J (2011) Umbrella sampling. *Wiley Interdiscip Rev Comput Mol Sci* 1:932–942.

[226] Curuksu J, Zacharias M (2009) Enhanced conformational sampling of nucleic acids by a new hamiltonian replica exchange molecular dynamics approach. *J Chem Phys* 130:104110.

[227] Ferrenberg AM, Swendsen RH (1989) Optimized Monte Carlo data analysis. *Phys Rev Letters* 63:1195–1198.

[228] Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J Comput Chem* 13:1011–1021.

[229] Souaille M, Roux B (2001) Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Comput Phys Commun* 135:40–57.

[230] Shirts MR, Chodera JD (2008) Statistically optimal analysis of samples from multiple equilibrium states. *J Chem Phys* 129:1–10.

[231] Genheden S, Ryde U (2015) The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin Drug Discov* 10:449–461.

[232] Aldeghi M, Bodkin MJ, Knapp S, Biggin PC (2017) Statistical analysis on the performance of molecular mechanics Poisson-Boltzmann surface area versus absolute binding free energy calculations: bromodomains as a case study. *J Chem Inf Model* 57:2203–2221.

[233] Kollman PA, et al. (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc Chem Res* 33:889–897.

[234] Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 98:10037–10041.

[235] Bashford D, Case DA (2000) Generalized born models of macromolecular solvetion effects. *Annu Rev Phys Chem* 51:129–152.

[236] David CC, Jacobs DJ (2014) Principal component analysis: a method for determining the essential dynamics of proteins. *Methods Mol Biol* 1084:193–226.

[237] Amadei A, Linssen ABM, Berendsen HJC (1993) Essential dynamics of proteins. *Proteins* 17:412–425.

[238] Sittel F, Jain A, Stock G (2014) Principal component analysis of molecular dynamics: on the use of cartesian vs. internal coordinates. *J Chem Phys* 141:014111.

[239] Antoniou D, Schwartz SD (2011) Toward identification of the reaction coordinate directly from the transition state ensemble using the kernel PCA method. *J Phys Chem B* 115:2465–2469.

[240] Naritomi Y, Fuchigami S (2011) Slow dynamics in protein fluctuations revealed by time-structure based independent component analysis: the case of domain motions. *J Chem Phys* 134.

[241] Kaila VRI, Sharma V, Wikström M (2011) The identity of the transient proton loading site of the proton-pumping mechanism of cytochrome *c* oxidase. *Biochim Biophys Acta Bioenerg* 1807:80–84.

[242] Li H, Robertson AD, Jensen JH (2005) Very fast empirical prediction and rationalization of protein p$K_a$values. *Proteins* 61:704–721.

[243] Seybold PG, Shields GC (2015) Computational estimation of p$K_a$ values . *Wiley Interdiscip Rev Comput Mol Sci* 5:290–297.

[244] Rabenstein B, Knapp EW (2001) Calculated pH-dependent population and protonation of carbon-monoxy-myoglobin conformers. *Biophys J* 80:1141–1150.

[245] Kieseritzky G, Knapp EW (2008) Optimizing p$K_a$computation in proteins with pHadapted conformations. *Proteins* 71:1335–1348.

[246] Ullmann GM, Knapp EW (1999) Electrostatic models for computing protonation and redox equilibria in proteins. *Eur Biophys J* 28:533–551.

[247] Smart OS, Goodfellow JM, Wallace BA (1993) The pore dimensions of gramicidin A. *Biophys J* 65:2455–2460.

[248] Mackerell AD, Feig M, Brooks CL (2004) Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulation. *J Comput Chem* 25:1400–1415.

[249] Huang J, MacKerell AD (2013) CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J Comput Chem* 34:2135–2145.

[250] Weigend F, Ahlrichs R (2005) Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: design and assessment of accuracy. *Phys Chem Chem Phys* 7:3297–3305.

[251] Ahlrichs R, Bär M, Häser M, Horn H, Kölmel C (1989) Electronic structure calculations on workstation computers: the program system Turbomole. *Chem Phys Lett* 162:165–169.

[252] de Jong DH, et al. (2013) Improved parameters for the Martini coarse-grained protein force field. *J Chem Theory Comput* 9:687–697.

[253] Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935.

[254] Jo S, Kim T, Iyer VG, Im W (2008) CHARMM-GUI: A web-based graphical user interface for CHARMM. *J Comput Chem* 29:1859–1865.

[255] Brooks BR, et al. (2009) CHARMM: the biomolecular simulation program. *J Comput Chem* 30:1174–1178.

[256] Phillips JC, et al. (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26:1781–1802.

[257] Berendsen HJC, van der Spoel D, van Drunen R (1995) GROMACS: a message-passing parallel molecular dynamics implementation. *Comput Phys Commun* 91:43–56.

[258] Abraham MJ, et al. (2015) Gromacs: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 1-2:19–25.

[259] Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. *J Mol Graph* 14:33–38.

[260] Schrödinger, LLC (2015) The PyMOL molecular graphics system, version 1.8.

[261] Chovancova E, et al. (2012) CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. *PLoS Comput Biol* 8:23–30.

[262] Bakan A, Meireles LM, Bahar I (2011) ProDy: protein dynamics inferred from theory and experiments. *Bioinformatics* 27:1575–1577.

[263] Scherer MK, et al. (2015) PyEMMA 2: a software package for estimation, validation, and analysis of Markov models. *J Chem Theory Comput* 11:5525–5542.

[264] Sierka M, Hogekamp A, Ahlrichs R (2003) Fast evaluation of the Coulomb potential for electron densities using multipole accelerated resolution of identity approximation. *J Chem Phys* 118:9136–9148.

[265] Riahi S, Rowley CN (2014) The CHARMM-TURBOMOLE interface for efficient and accurate QM/MM molecular dynamics, free energies, and excited state properties. *J Comput Chem* 35:2076–2086.

[266] Grossfield A (Version 2.0.9) WHAM: the weighted histogram analysis method. Version 2.0.9, http://membrane.urmc.rochester.edu/wordpress/?page_id=126.

[267] Pomès R, Roux B (1996) Structure and dynamics of a proton wire: a theoretical study of $H^+$ translocation along the single-file water chain in the gramicidin A channel. *Biophys J* 71:19–39.

[268] Hummer G, C. RJ, P. NJ (2001) Water conduction through the hydrophobic channel of a carbon nanotube. *Nature* 414:188–190.

[269] Kaila VRI, Hummer G (2011) Energetics and dynamics of proton transfer reactions along short water wires. *Phys Chem Chem Phys* 13:13207–13215.

[270] Saura P, Frey DM, Gamiz-Hernandez AP, Kaila VRI (2019) Electric field modulated redox-driven protonation and hydration energetics in energy converting enzymes. *Chem Commun* 55:6078–6081.

[271] Agmon N (1995) The Grotthuss mechanism. *Chem Phys Lett* 244:456 – 462.

[272] Hummer G (2005) Position-dependent diffusion coefficients and free energies from Bayesian analysis of equilibrium and replica molecular dynamics simulations. *New J Phys* 7:34.

[273] Ohnishi T, Ohnishi ST, Shinzawa-Itoh K, Yoshikawa S, Weber RT (2012) EPR detection of two protein-associated ubiquinone components ($SQ_{Nf}$ and $SQ_{Ns}$) in the membrane in situ and in proteoliposomes of isolated bovine heart complex I. *Biochim Biophys Acta Bioenerg* 1817:1803–1809.

# Article I

# Symmetry-related proton transfer pathways in respiratory complex I

Andrea Di Luca[a], Ana P. Gamiz-Hernandez[a], and Ville R. I. Kaila[a,1]

[a]Department Chemie, Technische Universität München, D-85747 Garching, Germany

Edited by Peter Brzezinski, Stockholm University, Stockholm, Sweden, and accepted by Editorial Board Member Harry B. Gray June 21, 2017 (received for review April 21, 2017)

**Complex I functions as the initial electron acceptor in aerobic respiratory chains of most organisms. This gigantic redox-driven enzyme employs the energy from quinone reduction to pump protons across its complete approximately 200-Å membrane domain, thermodynamically driving synthesis of ATP. Despite recently resolved structures from several species, the molecular mechanism by which complex I catalyzes this long-range proton-coupled electron transfer process, however, still remains unclear. We perform here large-scale classical and quantum molecular simulations to study the function of the proton pump in complex I from *Thermus thermophilus*. The simulations suggest that proton channels are established at symmetry-related locations in four subunits of the membrane domain. The channels open up by formation of quasi one-dimensional water chains that are sensitive to the protonation states of buried residues at structurally conserved broken helix elements. Our combined data provide mechanistic insight into long-range coupling effects and predictions for site-directed mutagenesis experiments.**

NADH:ubiquinone oxidoreductase | proton pumping | Grotthuss mechanism | multiscale simulation | bioenergetics

Complex I (NADH:ubiquinone reductase) is the largest enzyme of the respiratory chain, generating a proton motive force (*pmf*) that is used for synthesis of adenosine triphosphate (ATP) and active transport (1, 2). Complex I catalyzes electron transfer (eT) between nicotine adenine dinucleotide (NADH) and quinone (Q), and couples the energy released to pumping of four protons across the membrane (3–9). The distance between the electron and proton transferring modules extends up to approximately 200 Å. It currently remains unclear, however, how complex I catalyzes this remarkable long-range proton-coupled electron transfer (PCET) process. In addition to its central role in biological energy conversion, elucidating the molecular mechanism of complex I is of great biomedical relevance due to dysfunction of this enzyme in about half of all mitochondrial disorders (4, 8, 10).

In recent years, several structures of complex I have been resolved from both bacteria (11–13) and eukaryotes (14–19). During the completion of this study, two new structures of the mammalian enzyme were released (18, 19), resolving at near atomic resolution the structure of the supernumerary subunits in active and deactive states, complementing information from previously resolved structures (17). Similar to other respiratory enzymes (20, 21), the inner part of complex I, comprising 14 central core subunits, is well conserved, suggesting that the overall proton-pumping process might be similar in the bacterial and eukaryotic enzymes (but also see refs. 17, 22–24).

Complex I consists of a hydrophilic domain, responsible for the eT activity, and a membrane domain, which catalyzes proton transfer (pT) across the membrane (Fig. 1*A*). In the hydrophilic domain, a chain of 8–9 iron–sulfur (FeS) centers connects the NADH/flavin mononucleotide (FMN) site with the Q site (11, 25, 26), which locates approximately 30 Å above the membrane surface (14). The FeS centers form an electron conduction wire, in which the eT takes place on a microsecond–millisecond timescale between the two ends of the hydrophilic domain (27, 28). The Q head group is located approximately 10 Å from the terminal N2 FeS center and interacts with the conserved His-38 and Tyr-87 residues of subunit Nqo4 (*T. thermophilus* nomenclature) (13, 17), as originally identified in biochemical studies (29). Recent computational work (30) suggests that reduction of Q is coupled to deprotonation of the His-38 and Tyr-87, leading to formation of $QH_2$ that further triggers coupled protonation and conformational changes in the Nqo8 subunit, supporting previous suggestions from biochemical and structural data (6, 31, 32).

The membrane domain is composed of seven subunits, Nqo7–8 and Nqo10–14 (NuoA/H/J/K/L/M/N in *Escherichia coli*), arranged in an approximately 180-Å array linked to the hydrophilic domain by subunit Nqo8 (Fig. 1*B*). This domain comprises three antiporter-like subunits, Nqo12 (NuoN), Nqo13 (NuoM), and Nqo14 (NuoL) that are homologous to each other and also to multiresistance and pH adaptation (Mrp) $Na^+/H^+$ antiporters (12). These subunits are most likely responsible for pumping three protons, whereas the fourth proton channel might reside in the Nqo8 subunit (13, 17).

Each of the antiporter-like subunits have a unique internal pseudosymmetry (12) involving a twofold screw axis and comprising five transmembrane (TM) helical bundle segments, TM4–8 and TM9–13 (Fig. 1*A*, *Inset*). TM7a/b and TM12a/b are broken by short loops (Fig. 1*A*, *Inset*), a motif that is also used in ion translocation in carrier-type transporters (33). Nqo8 contains the same five-helical TM segment, but in contrast to the Nqo12/13/14 subunits, it is not symmetrical with respect to the other helices of the subunit. Structural studies (12, 13, 17) also identified a chain of highly conserved buried charged/polar residues, spanning the entire membrane domain, and creating

## Significance

Complex I is a redox-driven proton pump, central to aerobic energy conversion in most living organisms. To elucidate the mechanism of its pumping machinery, we need a detailed molecular picture of how access across the membrane is established and regulated. In this work we find that proton pathways in complex I form at symmetry-related locations near broken transmembrane helices. The channel opening allows influx of water molecules, catalyzing rapid Grotthuss-type proton transfer reactions. The hydration of these channels is sensitive to the protonation state of conserved buried lysine residues, which are in turn coupled to conformational changes in conserved ion pairs within each subunit. Our results provide mechanistic insight into the function of the long-range proton-pumping machinery in complex I.

Author contributions: A.D.L., A.P.G.-H., and V.R.I.K. designed research; A.D.L., A.P.G.-H., and V.R.I.K. performed research; A.D.L. and A.P.G.-H. contributed new analytic tools; A.D.L., A.P.G.-H., and V.R.I.K. analyzed data; and A.D.L. and V.R.I.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. P.B. is a guest editor invited by the Editorial Board.

[1]To whom correspondence should be addressed. Email: ville.kaila@ch.tum.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1706278114/-/DCSupplemental.

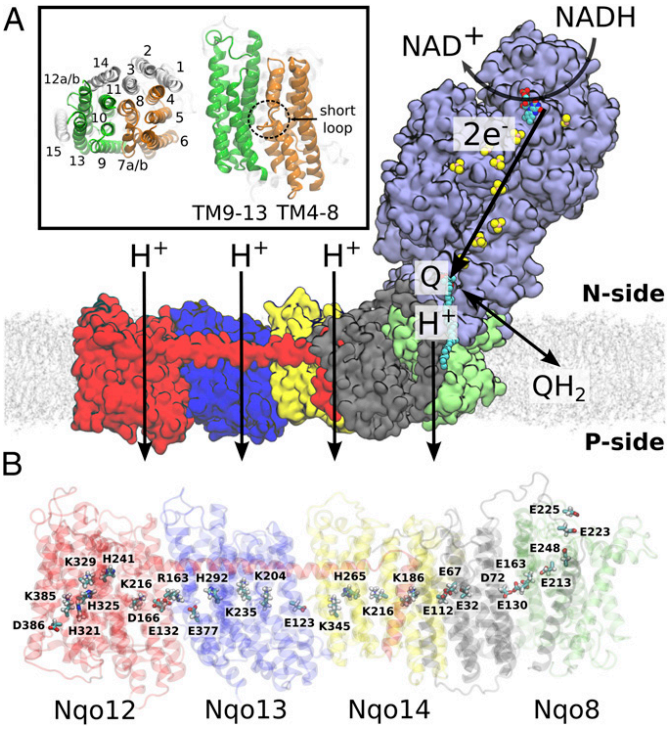


**Fig. 1.** Overall structure and function of bacterial complex I from *T. thermophilus* (PDB ID: 4HEA). (*A*) Electron transfer from NADH to Q in the hydrophilic domain (in purple) couples to proton transfer across the membrane domain, shown in red (Nqo12), blue (Nqo13), yellow (Nqo14), gray (Nqo11, Nqo10, and Nqo7), and green (Nqo8). (*Inset*) Transmembrane (TM) helix numbering of the antiporter-like subunits. Symmetry-related transmembrane helices TM4–8 (in orange) and TM9–13 (in green) in Nqo13, and TM15 (in transparent white) from Nqo12. (*B*) Array of polar and charged residues on the central axis in the membrane domain.

a connecting element between the Q site and Nqo12. Interestingly, this chain contains motifs that are common in each of the Nqo12–14 subunits: a lysine/glutamate pair, a central lysine, one or two histidine residues, and a terminal lysine (Nqo12 and Nqo14) or glutamate (Nqo13) (Fig. 1*B*). These residues may provide important functional elements in the proton-pumping machinery, because their mutation reduces the pumping activity (12, 13). Molecular simulations performed on the membrane domain of *E. coli* complex I (34) suggested that water molecules are involved in the formation of the proton pathways along these residues, but complete connections within all subunits were not resolved. The membrane domain of complex I also contains a long, approximately 110-Å, transverse horizontal helix (HL) from Nqo12 that spans along the domain and interacts with all three antiporter-like subunits. The helix is likely to have a clamping function and forms an element important for structure stability (35–37).

Here we explore the proton-pumping pathways in complex I from *T. thermophilus* (13) by large-scale classical molecular dynamics (MD) simulations on a microsecond timescale. To probe the proton transfer dynamics in the pumping process, we use a combination of MD simulations, hybrid quantum mechanics/classical mechanics (QM/MM) MD simulations, and electrostatic Poisson–Boltzmann/Monte Carlo calculations for sampling changes in protonation states. Our combined approach provides important suggestions on the pumping elements and long-range coupling effects in complex I that can further stimulate new experiments.

## Results

**General Channel Topology.** To probe the dynamics of complex I, we performed seven independent classical atomistic MD simulations, initiated from the X-ray structure of the enzyme from *T. thermophilus* (13). The protein structure was embedded in 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) membrane–water–ion surroundings and comprised in total approximately 1 million atoms. The buried conserved residues were modeled in different protonation states (*SI Appendix*, Tables S1 and S2), and the simulations extended in total 3.8 μs. In the MD simulations, we observe spontaneous hydration of the Nqo8 and Nqo12–14 subunits on a 0.2–0.4 μs timescale from the negatively charged (N) side of the membrane (Fig. 2 *A* and *B* and Movie S1). In Nqo12 and Nqo13, the water chains connect buried charged residues with the positively charged (P) side of the membrane, whereas for subunit Nqo14, we observe also a partial hydration from the P side of the membrane. The simulations suggest that the continuous water connectivity across the membrane is broken near highly conserved histidine residues, His-$325_{12}$/$292_{13}$/$265_{14}$ within the horizontal array of charged/polar residues at the center of the membrane domain (see below and Fig. 2).

The water chains form quasi one-dimensional hydrogen-bonded arrays, typical for many proton channels (38–40), and the hydration process increases the channel radius (Fig. 2*B*). We observe hydrated contacts between the bulk solvent and buried conserved residues Lys-$329_{12}$/Lys-$235_{13}$/Tyr-$288_{14}$ and Lys-$385_{12}$/Glu-$377_{13}$/Lys-$345_{14}$,

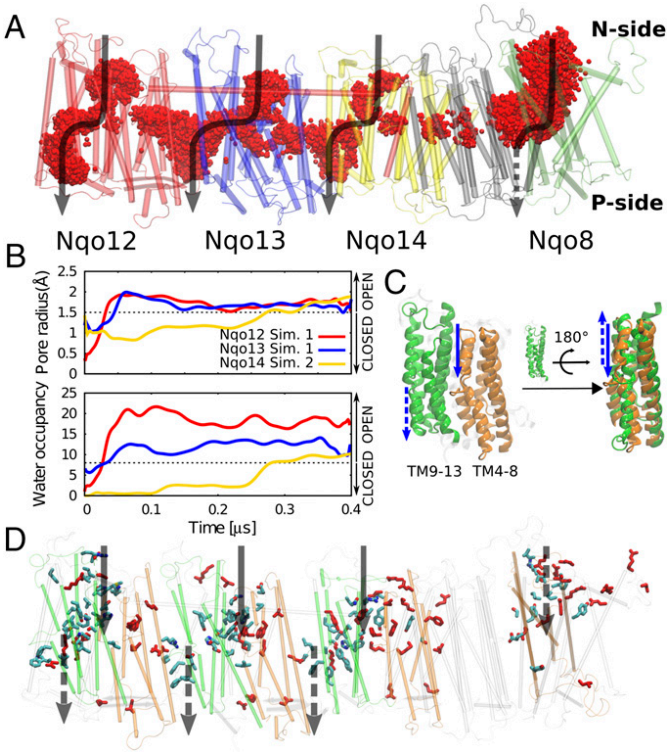


**Fig. 2.** Proton channel topology in complex I. (*A*) The membrane domain of complex I, showing overlapping water molecules (in red) obtained from seven independent simulation setups (3 μs in total). The water channels for each antiporter-like subunit and Nqo8 are shown by black arrows. (*B*) Water occupancy and channel radius of proton channels from the N side of the membrane to the center buried residues for three antiporter-like subunits calculated from two independent MD simulations (simulations 1 and 2). (*C*) The internal symmetry of the water input pathway from the N side and output to the P side shown as solid and dotted arrows, respectively, mirrors the internal symmetry of the antiporter-like subunits. Symmetry-related transmembrane helices TM4–8 (in orange) and TM9–13 (in green). (*D*) Positions of water pathways in the membrane domain, overlapped with conserved residues along the channels. Experimentally characterized residues that affect the proton-pumping activity are shown in red. The input channels from the N side in Nqo8 and output to the P side in Nqo12/13/14 are on the backside of this view.

along the N and P pathways, respectively. These pathways reflect the pseudo twofold internal symmetry of the subunits (Fig. 2*C*), with water arrays formed at similar locations on each side of the broken helices TM7b/12b (*SI Appendix*, Fig. S1). We observe that the broken helices undergo conformational changes in the opening/closure of the channels that may be important in controlling the hydration process (*SI Appendix*, Fig. S2). Due to their dynamic flexibility, broken helix elements also play an important function in carrier-type transporters (41, 42) where they are involved in establishing an alternate access between the two sides of the membrane. In carrier-type transporters, however, the channels often form along linear pathways at a dimeric protein interface, whereas in complex I the channels seem to form at the edge of every half subunit along the domain (Fig. 2*D*), consistent with their symmetry.

We find that the increase in channel radius in all antiporter-like subunits correlates well with the water occupancy in the channel (*SI Appendix*, Figs. S3 and S4). The water wires form spontaneously in many independent simulations (simulations 1–3 and 5–7) when the residues located at the end of the channels are in their charged protonation states, suggesting that the overall behavior is statistically significant. Interestingly, upon deprotonation of the central lysine residues Lys-$329_{12}$/Lys-$235_{13}$/Lys-$216_{14}$ located at the end of the N-side pathway (simulation 4), we observe a rapid dehydration of the N-side channels in <100 ns (Fig. 3). This observation suggests that the protonation state of the central buried residues might control the water connectivity across the membrane domain.

Based on structural data (12, 13) and site-directed mutagenesis experiments (43–48), the N-side proton channels in Nqo12–14 have previously been suggested to reside either near the Glu/Lys pair within each antiporter-like subunit or at the middle Lys around TM8 (9). The putative channel topology observed in our MD simulations somewhat differs from these sites and resembles more the paths suggested for *Yarrowia lipolytica* (17), with an entry site near TM7b in the middle of the two bundle segments TM4–8 and TM9–13 and output site near TM12b. Moreover, the Glu/Lys pairs, which were previously suggested to take part in the pumping process (12), are not located along the N-side input channels, and therefore the direct involvement of these residues in the proton uptake seems unlikely, based on the current simulation data. The dissociation of these residue pairs might, nevertheless, be important for propagating electrostatic forces along the central axis of conserved residues (see below).

In Nqo8, we observe hydration along the broken helix TM5 within the TM2–6 bundle, although this subunit has a somewhat different structure and sequence compared with the other antiporter-like subunits Nqo12/13/14 (13). These findings support that a fourth proton pathway may indeed reside within Nqo8, as previously suggested by Baradaran et al. (13).

Our putative channel locations thus coincide with the position of several conserved and functionally important residues, identified in previous biochemical studies (Fig. 2*D* and *SI Appendix*, Table S3). Importantly, our putative channels also comprise several residues that have not yet been experimentally probed. To this end, we have summarized central residues along the N- and



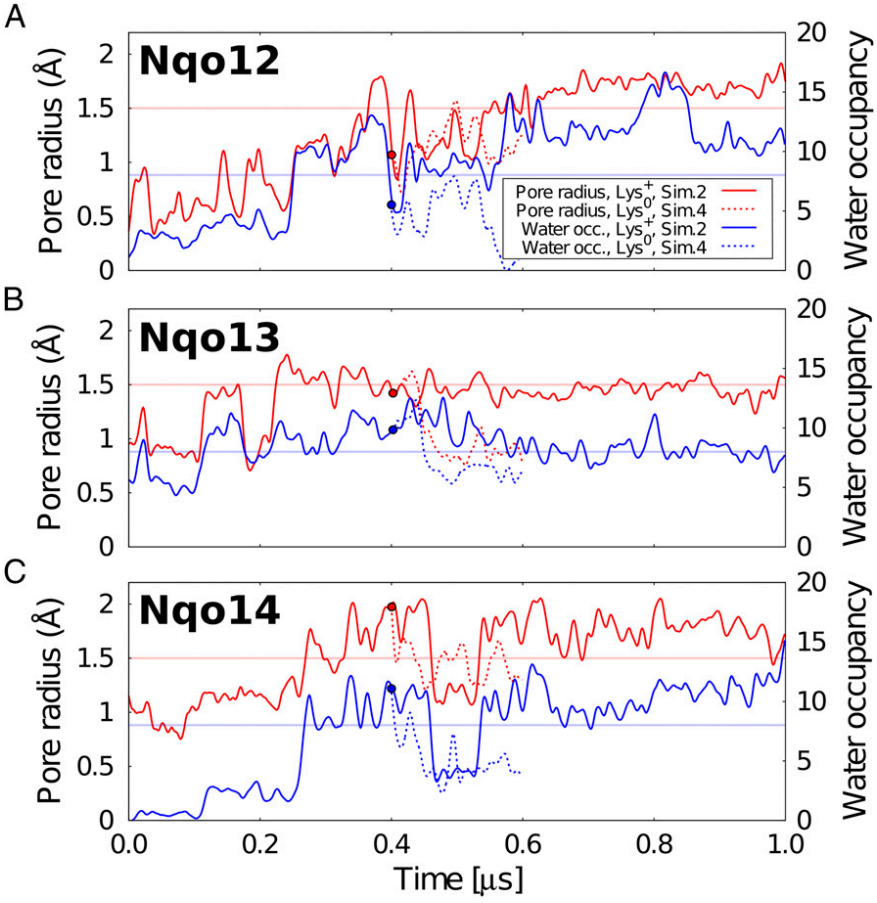


**Fig. 3.** Pore radius and water occupancy of N-side proton channels in Nqo12–14. The pore radius (in red) and water occupancy (in blue) are shown for (*A*) Nqo12, (*B*) Nqo13, and (*C*) Nqo14 in two different protonation states of the middle lysine (simulations 2 and 4, *SI Appendix*, Table S2). Circle at 0.4 μs shows simulation 4, started after 0.4 μs of simulation 2 by deprotonating the middle Lys-$329_{12}$/Lys-$235_{13}$/Lys-$216_{14}$. The channel opening threshold was defined based on the pore radius of 1.5 Å, which allows influx of water molecules. Details for simulations 1–7 are shown in *SI Appendix*.

P-side pathways in *SI Appendix*, Table S3. In the following sections, we present structural details about the putative proton pathway in each subunit.

**Proton Pathways in Nqo12 (NuoL).** Water molecules enter the Nqo12 subunit (simulations 1, 2, and 5–7) by a crevice between the helices TM7b, TM8, and TM10. This conduit provides a hydrogen-bonded connectivity between the bulk and the conserved buried residue Lys-329, which further connects via His-325 and His-321 to Lys-385/Asp-386 (Fig. 4, Nqo12/N). The channel from the P side of the membrane becomes hydrated in most simulations (*SI Appendix*, Fig. S4), and forms around the tip of the broken helix TM12b, near Lys-385 and Asp-386 (Fig. 4, Nqo12/P). This pathway lines up with TM15, close to the edge of the membrane subunit.

Although the two water pathways from the N and P sides come in close contact near Lys-329, Phe-328 and Leu-373 tightly seal the channel, and do not allow water exchange between the two sides of the membrane (*SI Appendix*, Fig. S5). A dry area is also observed in the proximity of Asp-386, suggesting that this residue might control the access to the P side of the membrane. The simulations show that the conserved Gln-302 of TM10 dynamically switches its hydrogen-bonding contacts between Lys-329 and His-321 (*SI Appendix*, Fig. S6). Although the glutamine cannot undergo protonation changes itself, its side chain can act as a hydrogen bond donor/acceptor and thus stabilize protonated intermediates in the transfer process.

**Proton Pathways in Nqo13 (NuoM).** The MD simulations suggest that movement of the broken helix TM7b on an approximately 0.1-μs timescale (simulations 1, 2, 6, and 7) opens a pathway for water molecules to enter Nqo13 from the N side of the membrane (Fig. 4, Nqo13/N). This channel leads to the buried conserved Lys-235 (TM8), as also previously suggested as a potential input channel in the *E. coli* structure (34). Connection to the P side forms at the interface between Nqo12 (TM5) and Nqo13 (TM12b) (simulation 6) from Glu-377 (Fig. 4, Nqo13/P). Formation of this hydrogen-bonded network to the P side is triggered by a conformational change in Phe-374, a residue that is conserved in all antiporter-like subunits of complex I.

During the opening of the N-side channel, we observe a dissociation of a backbone hydrogen bond between His-211 and Leu-214 (*SI Appendix*, Fig. S7). To further probe the involvement of these residues, we performed steered molecular dynamics (SMD) simulations in which this hydrogen bond was broken. These simulations suggest that the channel opening/closure can indeed be induced by perturbation of this reaction coordinate (*SI Appendix*, Fig. S7). Our unbiased MD simulations also indicate that when the N-side channel is closed (simulation 5), water molecules establish a bridging pathway between Lys-235 and Glu-377 via His-292, providing a conduction pathway toward the P side of the membrane (see below).

The channel from the N side passes close to Asp-557 (Nqo12), which is a conserved residue on the transverse HL helix. Although site-directed mutagenesis data (35) suggest that the HL helix might provide a clamping function in complex I rather than the earlier suggested piston function (12), the mutation of this aspartate to a glutamate strongly decreases the proton-pumping activity by approximately 50% (49).

**Proton Pathways in Nqo14 (NuoN).** The Nqo14 subunit undergoes the slowest hydration process in our MD simulations on an approximately 0.3-μs timescale (Fig. 2*B* and *SI Appendix*, Fig. S3). We find that water molecules enter this subunit through a crevice between TM7b, TM8, and TM10 (Fig. 4, Nqo14/N) leading to the conserved Tyr-288 and Lys-216 residues. This pathway thus forms at locations that closely resemble the channels observed in Nqo12. Similar to the P-side channel formed in Nqo13, we also observe a partial hydration at the interface between Nqo13 and

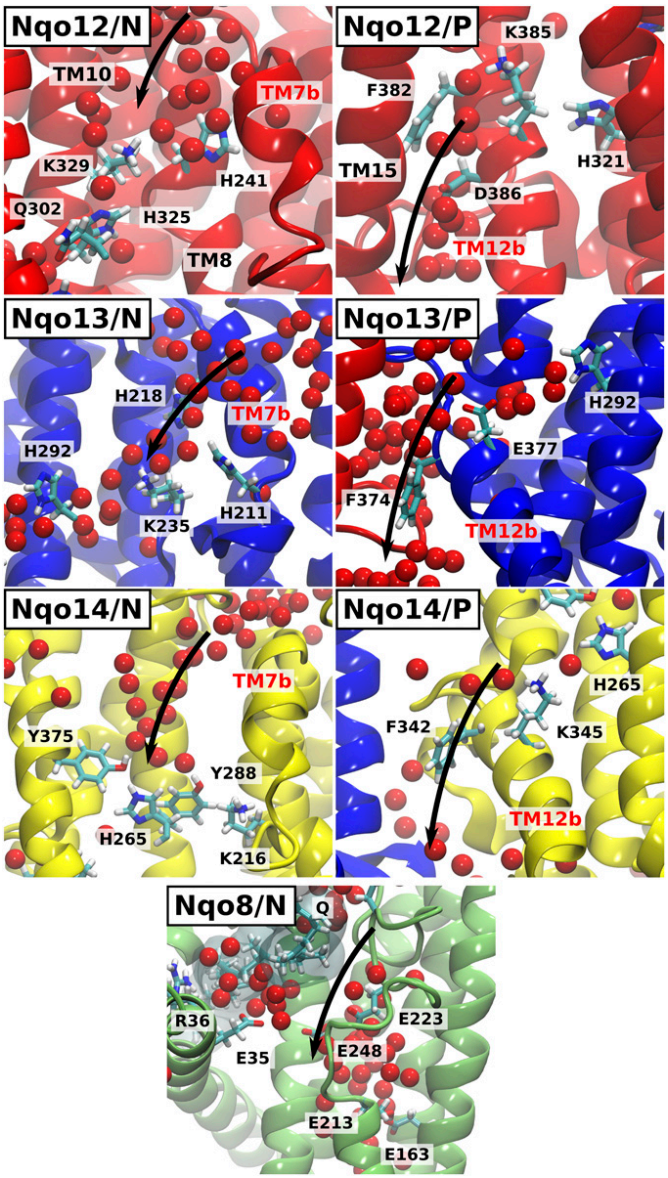


**Fig. 4.** Proton pathways from the N and P sides of the membrane in complex I. Water molecules enter between TM helices 7b, 8, and 10 (Nqo12–14) and TM5 and TM6 (Nqo8) from the N-side bulk and establish a conduction pathway to the buried residues Lys-$329_{12}$/Lys-$235_{13}$/Tyr-$288_{14}$/Glu-$163_{8}$. Nqo12/P, Nqo13/P, Nqo14/P: Water pathways next to TM12b lead from Lys-$385_{12}$/Glu-$377_{13}$/Lys-$345_{14}$ to the P side of the membrane. Some helices are not shown for visualization purpose.

Nqo14 (Fig. 4, Nqo14/P), but Phe-342 seems to partially block the pathway between the Lys-345 and the P side. The channel from the N side in Nqo14 is closed in the crystal structure by the bulky side chain of Leu-258, but its side-chain movement correlates with formation/disruption of the water wire (*SI Appendix*, Fig. S8), suggesting that this residue might have a potential gating function. In previous simulations of the *E. coli* structure (34), a channel leading to the middle Tyr-288/Lys-216 pair was also observed.

We do not observe substantial hydration in Nqo14 at the symmetry-related position where the channel forms in Nqo13. In contrast, this analogous site in Nqo14 is covered by the C-terminal loop of Nqo11 (*SI Appendix*, Fig. S9) that was recently suggested to form an element important for structural stability based on mutagenesis and cross-linking studies in *E. coli* (50).

Moreover, in contrast to the water channel in Nqo13, where the HL helix inserts Asp-557$_{12}$ near His-211, His-193 in Nqo14 lacks such an interaction partner with the transverse HL helix.

**Proton Pathways in Nqo8 (NuoH).** We observe in all independent MD simulations that the Nqo8 subunit undergoes large hydration changes on approximately 0.1-μs timescales (*SI Appendix*, Fig. S3). The Nqo8 channel contains many charged and polar residues, and it is also wider compared with the hydration sites in the other subunits. The hydration extends from the N side to Glu-213/163 within the so-called E channel (13), located approximately 15 Å below the membrane surface. This region was previously suggested to undergo proton uptake induced by Q reduction (30). Our simulations suggest this hydrated channel coincides with the lower part of the quinone cavity (Fig. 4, Nqo8/N), supporting that proton uptake/release might couple to movement of within its tunnel. Despite a significant N-side entry channel, we observe no clear exit pathways to the P side of the membrane within Nqo8.

**Proton Transfer Dynamics in Nqo13.** To probe the mechanism of proton transfer along the formed water channels, we performed QM/MM MD simulations of two putative steps of the proton-pumping cycle in the Nqo13 subunit. To this end, we studied the proton uptake from the bulk to the buried Lys-235 residue, and then further along the lateral water pathway from Lys-235 to Glu-377 (Fig. 5 *A* and *B* and *SI Appendix*, Table S4). To study the proton transfer from the solvent to Lys-235, we used a starting geometry obtained after a 0.1-μs classical MD simulation in which the channel from the N side was well hydrated (simulation 6) and Lys-235 was neutralized. We observe a rapid transfer of the excess proton placed at the entrance of the channel (N side) to the deprotonated Lys-235 (Fig. 5*A*). The transfer process takes place by a Grotthuss-type proton exchange along Zundel and hydronium species, typical for proton transfer in quasi one-dimensional water wires (40, 51) and subsequent rearrangement of the hydrogen bonded network (52). His-211 does not seem to directly take part in the transfer process, but stabilizes the water network and the proton in the cavity (Movie S2). To further probe the dependence of this proton transfer process on the channel hydration state, we performed several independent QM/MM simulations (320 ps in total), starting from different starting structures along the hydration trajectory. We find that, whereas the proton transfer takes place on the picosecond timescale in the simulations that are initiated from well-hydrated structures, the simulations initiated from dry or broken hydrogen-bonded networks have a lower probability for productive proton transfer reactions to the lysine on the sampled simulation timescale (*SI Appendix*, Fig. S10 and Table S5). This behavior suggests that, whereas there is a large thermodynamic driving force for the proton to transfer from the hydronium to the deprotonated lysine, the slow (microsecond) channel hydration might be rate limiting in the proton uptake process. Interestingly, for some partially hydrated structures, we also find that the hydronium species does itself participate in establishing a water wire toward the central lysine. Similar proton transfer mechanisms have been previously observed in carbon nanotubes (51).

In another independent QM/MM MD simulation setup, we studied the proton transfer along the Lys-235/His-292/Glu-377 triad to further elucidate the role of Lys-235 in the proton release. Starting from a structure with the N- and P-side channels closed, the proton is transferred rapidly in three independent simulations from Lys-235 via His-292 and water molecules to Glu-377 (Fig. 5*B* and Movie S3). This result supports the idea

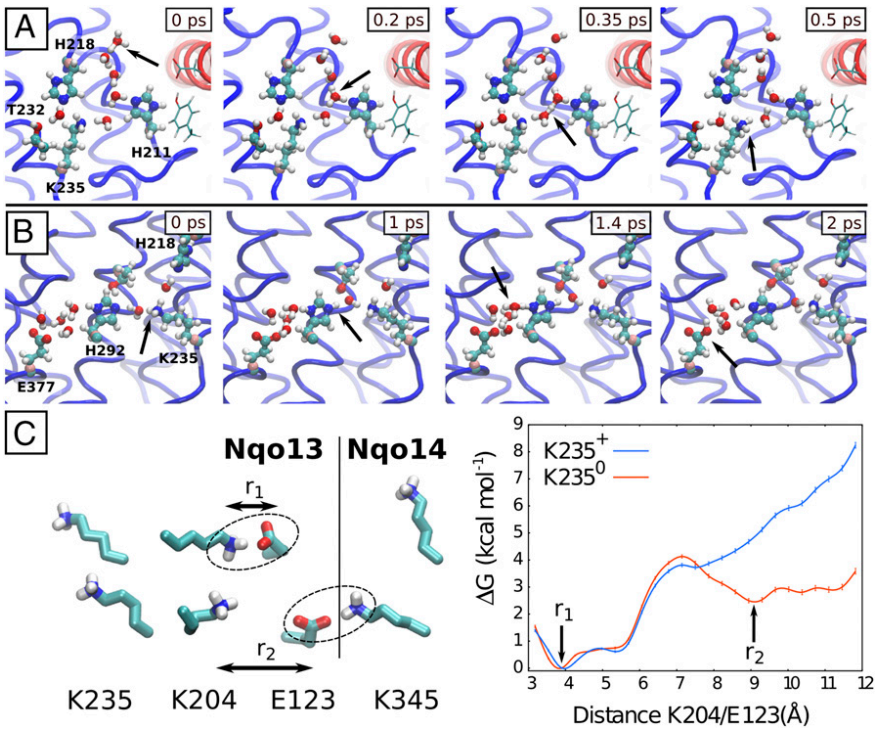


**Fig. 5.** Proton transfer dynamics and conformational switching in Nqo13. (*A*) A water chain from the N side, formed in the MD simulations (simulation 6), catalyzes Grotthuss-type proton transfer to Lys-235 in Nqo13 in QM/MM MD simulations. (*B*) Proton transfer from Lys-235 to Glu-377. The simulation was initiated from a state in which the channel from the N side is closed and the residues Lys-235 and Glu-377 are protonated and deprotonated, respectively (simulation 5). His-292 bridges the proton transfer between the K235 and E377. (*C*) Classical free-energy profiles of Lys-204/Glu-123 (Nqo13) dissociation when Lys-235 is protonated (in blue) and deprotonated (in red). Glu-123 interacts both with Lys-204 and Lys-345 (Nqo14). The second minimum is absent when Lys-235 is protonated, possibly due to electrostatic repulsion with Lys-204.

that the proton transfer along the central polar axis is indeed possible. Our continuum electrostatic calculations on the crystal structure also suggest that the p$K_a$ of Glu-377 is higher than the p$K_a$ of Lys-235 (*SI Appendix*, Table S6), further supporting this hypothesis. The QM/MM MD simulations thus support the idea that the water chains formed in the classical MD simulations provide efficient conduits for proton transfer in the membrane domain of complex I.

**The Lysine/Glutamate Pair as an Element for Modulating the Proton Pump.** The MD simulations suggest that the protonation state of the buried residues could regulate the channel hydration (Fig. 3 and *SI Appendix*, Figs. S3 and S4). More specifically, the N-side access channel in Nqo13 is open when Lys-235 is deprotonated and His-211 is modeled in its protonated state (simulation 1) or when Lys-235 is protonated and His-211 is neutral (δ-tautormer, simulations 6 and 7). In contrast, the channel is significantly less hydrated in simulations where both Lys-235 and His-211 are in their neutral states (simulation 4), suggesting that the protonation states of these residues play a role in the channel opening.

To study how the dissociation of the Lys/Glu pair couples to the protonation state of the middle Lys-235, we performed free-energy calculations using the replica exchange umbrella sampling (REUS) method (53), where we biased the dissociation of Lys-204/Glu-123 in Nqo13 and modeled Lys-235 in both its protonated and neutral forms. The free-energy simulations show a clear preference of the Lys/Glu pair to be associated when Lys-235 is protonated (Fig. 5*C*). Interestingly, we observe a second minimum in the free-energy profile when Lys-235 is deprotonated (Fig. 5*C*), in which Glu-123 forms an ion pair with Lys-345 of Nqo14. For the protonated Lys-235, this minimum is absent, possibly due to the repulsion between the positively charged residues. Based on the principle of microscopic reversibility, this difference indicates that dissociation of the Lys-204/Glu-123 ion pair could increase the deprotonation probability of the central Lys-235, and possibly triggering lateral proton transfer toward the P side.

To further explore this coupling effect in all three antiporter-like subunits, we performed SMD simulations where the dissociation of the $Lys^+/Glu^-$ ($Asp^-$ in Nqo12) pair was induced (*Methods*) in the three antiporter-like subunits, followed by p$K_a$ calculations. These calculations suggest that the p$K_a$ of the middle lysines ($Lys\text{-}329_{12}$, $Lys\text{-}235_{13}$, and $Lys\text{-}265_{14}$) are affected by the dissociation of the ion pairs (*SI Appendix*, Fig. S11), further supporting the findings from the REUS simulations. Benchmarking studies suggest that PB electrostatic calculations with MC sampling have a rmsd of approximately 1 pK unit from experimental values (54). These p$K_a$ benchmarking studies, however, do not include large membrane proteins such as complex I, which might be subjected to larger errors.

## Discussion

Our combined results from classical MD simulations, QM/MM calculations, and PB electrostatic models suggest that proton pumping in complex I takes place by hydration of proton channels that form along conserved broken helices, TM7b and TM12b in Nqo12-14, and TM5 in Nqo8, at symmetry-related locations. The hydration is sensitive to the protonation state of buried residues, with charged and neutral states inducing channel opening and closure, respectively. Our simulations suggest that the electrostatic field induced by these buried residues pull in water molecules from the solvent, and the quasi one-dimensionally oriented chains might further strengthen the electrostatic couplings between the residues. Our free-energy and steered molecular simulations further suggest that the dynamics of the conserved Glu/Lys ion pair within TM5/7a modulate the p$K_a$ of the buried central residues. These results indicate that complex I may operate within each antiporter-like subunit by using a unit dedicated for signal transmission (TM4–8), and another unit responsible for the proton-pumping activity (TM9–13). The lateral proton transfer within the TM9–13, as explored in our QM/MM simulations, leads to protonation of the terminal protonatable residue ($Lys_{12/14}/Glu_{13}$), and could in turn modulate the dynamics of the signal transduction unit in the neighboring subunit, as suggested by the MD data where intersubunit contacts are formed (*SI Appendix*, Fig. S12).

Although the X-ray structure of complex I from *T. thermophilus* represents a single state of limited resolution (3.3 Å), it nevertheless provides a good starting point to probe the dynamics of transient states along the pumping cycle. We observe similar hydration dynamics in several independent simulations, supporting the idea that the overall findings are robust and statistically significant. Moreover, although the timescales of the individual trajectories are shorter than the overall millisecond turnover of complex I, computations performed on transient states provide valuable information on the structure and dynamics of the pumping cycle and information that is complementary to that obtained from many experiments. Further work is, nevertheless, needed to elucidate the exact role of the interplay between the protonation states of the central Lys and/or His residues as well as the involvement of the Lys/Glu ion pairs in modulating the hydration and proton transfer energetics.

Our findings provide mechanistic ideas into how the force might propagate from the Q site across the antiporter-like subunits, schematically outlined in Fig. 6. In this putative model, the pumping is initiated by reduction of quinone that triggers protonation changes in Nqo8 (30) along water molecules near TM5 (Fig. 4, Nqo8/N). The protonation change further triggers dissociation of the Lys/Glu pair in Nqo14, which in turn could initiate proton transfer and release to the P side along the pathway near TM12b. Deprotonation of the middle lysine in Nqo14 is expected to close the connectivity to the N side (Fig. 3), and could help to prevent the pumped proton from leaking to the N side. Protonation changes in the Nqo14 would then induce dissociation of the Lys/Glu ion pair in Nqo13, which in turn would propagate by similar steps to Nqo12. Relaxation of the Lys/Glu ion pairs reestablishes the original protonation states by protonation from the N side along the TM7b pathways, whereas quinone/quinol exchange and rereduction by NADH initiates a new pumping cycle.

This schematic model favors a sequential propagation of the electrostatic force from the Q site along Nqo8 to Nqo14, and further from Nqo14 to Nqo13 and Nqo12. Local electrostatic coupling between neighboring charged residues is strong, which could also allow pumping without large-scale conformational changes within the antiporter-like subunits. Other mechanistic models, such as the wave-spring model (7, 45), have also been presented, in which Nqo8/Nqo13 and Nqo12/Nqo14 pump and release protons during synchronized steps. It is currently unclear, however, whether complex I operates by direct (electrostatic) or indirect (conformational) coupling (6, 55, 56) using one- (22, 27, 57) or two-stroke (31) mechanisms. Our simulations nevertheless support the idea that both electrostatic coupling between titratable residues as well as conformational changes in broken helices might be important for the pumping function in complex I.

## Conclusions

We have performed here multiscale classical and quantum molecular simulations based on the experimentally resolved structure of complex I from *T. thermophilus* to probe how protonation and conformational states of the protein are linked to the proton-pumping process. We observed that hydrated channels form at symmetry-related locations, near TM7b and TM12b in Nqo12–14 and TM5b in subunit Nqo8, mirroring the internal pseudo twofold screw axis of the antiporter-like subunits in complex I. The channels to the N and P sides initiate from buried charged residues, and the protonation states of these residues

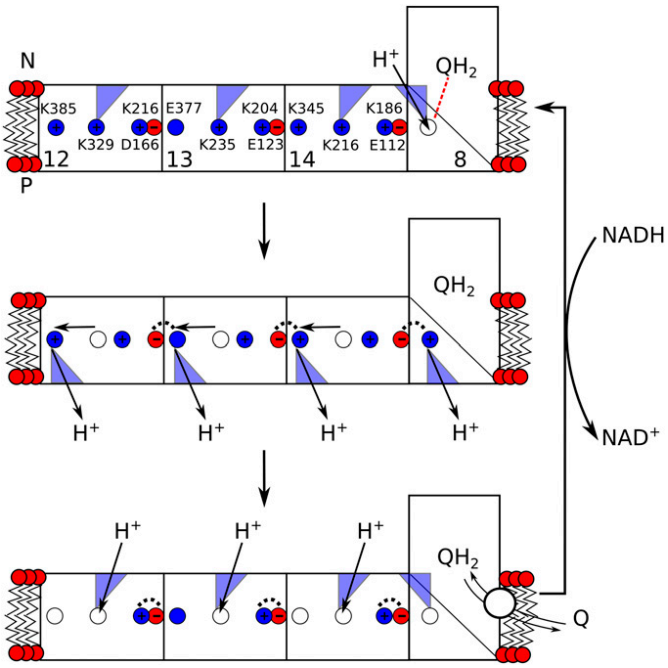


**Fig. 6.** Schematic outline of a putative pumping cycle in complex I. (*Top*) Reduction of Q induces proton uptake in Nqo8 that leads to sequential opening of the $K^+/E^-$ ion pairs in each subunit. (*Middle*) The $K^+/E^-$ opening triggers proton transfer along the later hydrophilic axis and release of the proton to the P side. The open $K^+/E^-$ ion pairs are stabilized by interactions with neighboring subunits (dashed lines). Deprotonation of the middle buried residues prevents the pumped proton from leaking back to the N side by blocking water connectivity. (*Bottom*) $QH_2$/Q exchange and relaxation of the $K^+/E^-$ ion pairs opens up water access from the N side, reforming the initial protonation states. Reduction of Q by NADH reinitiates a new pumping cycle.

themselves were found to correlate well with the hydration states of the channels as well as with conformational changes in TM7b/12b. Our simulations suggested that the electrostatic coupling effects are important for the proton-pumping process. We also found that the conformation of the conserved Glu/Lys ion pairs might modulate the p$K_a$ of the central buried residues, with important mechanistic implications. Our QM/MM MD simulations further suggested that the proton transfer is rapid, taking place on picosecond timescales, after the water wires have formed on microsecond timescales. The simulations also support a previously postulated lateral proton transfer along the polar buried axis in states where the channel to the N side is closed. These combined results may stimulate new site-directed mutagenesis experiments (*SI Appendix*, Table S3) and provide insights into the remarkable long-range proton-coupled electron transfer machinery in complex I.

## Methods

**Classical MD Simulations.** The X-ray structure from *T. thermophilus* (PDB ID: 4HEA) (13) was embedded in a POPC membrane and solvated with TIP3P water. Sodium and chloride ions were added to give an ionic strength of approximately 100 mM. The Q site was modeled by inserting ubiquinone ($Q_{10}$) in the cavity, identified using the software HOLE (58), and by placing the Q head group next to His-38 and Tyr-87. The total system size comprised approximately 830,000 atoms. Seven independent simulation systems, each studied for 0.1–1 μs, were created to probe the effect of different protonation states (*SI Appendix*, Table S2). All classical MD simulations were performed using NAMD2 (59) with the CHARMM27 (60, 61) force field, and density functional theory (DFT)-derived parameters for all cofactors (FeS centers, Q, FMN) (30). The simulations were performed in an *NPT* ensemble with $T = 310$ K, $P = 1$ atm, using a time step of 2 fs, and treating long-range electrostatics by the Particle Mesh Ewald (PME) method. Water occupations were calculated from data averaged over the MD trajectories by selecting water molecules from residues shown in *SI Appendix*, Table S4. Pore radii were calculated using the software HOLE (58). SMD simulations were performed based on structures obtained after dynamics (*SI Appendix*, Table S1).

A 20-ns SMD simulation was performed by perturbing the backbone His-211(O)/Leu-214(N) atom distances between 3 and 6 Å using a force constant of 50 kcal $mol^{-1}$ $Å^{-2}$ to probe channel hydration in Nqo13 (*SI Appendix*, Fig. S7). The 3 × 16 ns simulations were performed by perturbing the head groups distances of Lys-216 (Nζ)/Asp-166 (Cγ) in Nqo12, Lys-204 (Nζ)/Glu-123 (Cδ) in Nqo13, and Lys-186 (Nζ)/Glu-112 (Cδ) in Nqo14 between 3 and 12 Å, using force constants of 50 kcal $mol^{-1}$ $Å^{-2}$. REUS simulations (53) (simulations 12 and 13, *SI Appendix*, Table S1) were performed using 25 replicas started from the SMD structures (simulation 10, *SI Appendix*, Table S1), modeling Lys-235 both in its neutral and charged states. A harmonic biasing potential with a force constant of 10 kcal $mol^{-1}$ $Å^{-2}$ was used to restrain the distance of Lys-204 (Nζ) and Glu-123 (Cδ) head groups between 3 and 12 Å. Exchange moves were attempted every 2 ps, and distances were collected every 20 ps. The free-energy profiles converged after 14 × 25 ns REUS simulations and were reconstructed using the weighted-histogram analysis method (WHAM) (62, 63) and estimating errors by bootstrapping analysis. Visual Molecular Dynamics (64) was used for visualization and analysis of all MD simulations. The total MD simulation time was approximately 3.8 μs.

**Poisson–Boltzmann Continuum Electrostatics with Monte Carlo Sampling.** p$K_a$ values were calculated using PB continuum electrostatic calculations with MC sampling of $2^N$ protonation states of a nine subunit model of complex I, comprising the membrane domain and subunits Nqo4/6. The PB calculations were performed using the Adaptive Poisson–Boltzmann Solver (APBS) (65), and the MC sampling was performed using Karlsberg+ (54, 66). The protein was described using explicit atoms with atomic partial charges, embedded in an inhomogeneous dielectric continuum with a dielectric constant of 4. The bulk water was described by a homogeneous dielectric continuum with a dielectric constant of 80. The boundary interface between the protein and solvent was calculated by the molecular surface routine implemented in APBS, using a solvent probe radius of 1.4 Å, and modeling an implicit ionic strength with 100 mM potassium chloride. Protonation probabilities were probed along all classical simulations every 0.1 ns. The Lys/Glu (Asp) pairs of each antiporter-like subunit were fixed in their standard protonation states in the p$K_a$ calculations of the SMD trajectories.

**QM/MM MD Simulations.**

***Proton transfer from the N side to Lys-235$_{13}$.*** Setup 1 was as follows: QM/MM MD simulations were performed based on eight independent structures obtained from simulation 6. A water molecule at the channel entrance was replaced with a hydronium ion and classically relaxed for 10 ps for each of the eight starting structures. A total of 56 independent QM/MM MD simulations were initiated from these relaxed structures with different levels of water chain connectives between the N side and Lys-235. Lys-325, His-211, His-218, and Thr-232 together with the water molecules in the cavity were treated at the DFT level using the BP86 functional (67, 68), the multipole accelerated resolution of identity (MARIJ) approximation (69), and def2-SVP basis sets (70). The QM region comprised approximately 75 atoms and was coupled to the MM region by link atoms, introduced between the Cα and Cβ atoms. Additional eight QM/MM systems comprising 125 atoms were used to model the systems with broken water connectives from the N side (*SI Appendix*, Table S4). The MM system comprised the Nuo12/13/14 subunits and surrounding lipids and water molecules, with a total system size of approximately 75,000 atoms. The QM/MM MD simulations were simulated for 5 ps each, using a 1-fs integration time step, and a temperature of $T = 310$ K. The simulations were performed using the CHARMM/TURBOMOLE interface (71–73).

***Proton transfer from Lys-235$_{13}$ to Glu-377$_{13}$.*** Setup 2 was as follows: Structures used for the QM/MM setup 2 were taken from simulation 5, after 0.6 μs of dynamics. Residues His-218, Lys-235, His-292, Thr-322, and Glu-377 together with six bridging water molecules were included in the QM region, comprising 78 atoms, and treated as described above. Three independent QM/MM MD simulations were simulated for 5 ps each, using a 1-fs integration time step at a temperature of $T = 310$ K.

**ACKNOWLEDGMENTS.** We thank Prof. Mårten Wikström and Dr. Gerhard Hummer for insightful discussions. This work was supported by the German Research Foundation. The computing time was provided by the SuperMuc at the Leibniz Rechenzentrum (Computing Grant pr48de).

1. Mitchell P (1961) Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* 191:144–148.
2. Yoshida M, Muneyuki E, Hisabori T (2001) ATP synthase: A marvellous rotary engine of the cell. *Nat Rev Mol Cell Biol* 2:669–677.
3. Wikström M (1984) Two protons are pumped from the mitochondrial matrix per electron transferred between NADH and ubiquinone. *FEBS Lett* 169:300–304.
4. Brandt U (2006) Energy converting NADH:quinone oxidoreductase (complex I). *Annu Rev Biochem* 75:69–92.
5. Dröse S, et al. (2011) Functional dissection of the proton pumping modules of mitochondrial complex I. *PLoS Biol* 9:e1001128.
6. Wikström M, Hummer G (2012) Stoichiometry of proton translocation by respiratory complex I and its mechanistic implications. *Proc Natl Acad Sci USA* 109:4431–4436.
7. Verkhovskaya M, Bloch DA (2013) Energy-converting respiratory Complex I: On the way to the molecular mechanism of the proton pump. *Int J Biochem Cell Biol* 45:491–511.
8. Hirst J (2013) Mitochondrial complex I. *Annu Rev Biochem* 82:551–575.
9. Sazanov LA (2015) A giant molecular proton pump: Structure and mechanism of respiratory complex I. *Nat Rev Mol Cell Biol* 16:375–388.
10. Mimaki M, Wang X, McKenzie M, Thorburn DR, Ryan MT (2012) Understanding mitochondrial complex I assembly in health and disease. *Biochim Biophys Acta* 1817:851–862.
11. Sazanov LA, Hinchliffe P (2006) Structure of the hydrophilic domain of respiratory complex I from *Thermus thermophilus*. *Science* 311:1430–1436.
12. Efremov RG, Sazanov LA (2011) Structure of the membrane domain of respiratory complex I. *Nature* 476:414–420.
13. Baradaran R, Berrisford JM, Minhas GS, Sazanov LA (2013) Crystal structure of the entire respiratory complex I. *Nature* 494:443–448.
14. Hunte C, Zickermann V, Brandt U (2010) Functional modules and structural basis of conformational coupling in mitochondrial complex I. *Science* 329:448–451.
15. Vinothkumar KR, Zhu J, Hirst J (2014) Architecture of mammalian respiratory complex I. *Nature* 515:80–84.
16. Wirth C, Brandt U, Hunte C, Zickermann V (2016) Structure and function of mitochondrial complex I. *Biochim Biophys Acta* 1857:902–914.
17. Zickermann V, et al. (2015) Structural biology. Mechanistic insight from the crystal structure of mitochondrial complex I. *Science* 347:44–49.
18. Fiedorczuk K, et al. (2016) Atomic structure of the entire mammalian mitochondrial complex I. *Nature* 538:406–410.
19. Zhu J, Vinothkumar KR, Hirst J (2016) Structure of mammalian respiratory complex I. *Nature* 536:354–358.
20. Kaila VRI, Verkhovsky MI, Wikström M (2010) Proton-coupled electron transfer in cytochrome oxidase. *Chem Rev* 110:7062–7081.
21. Wikström M, Sharma V, Kaila VRI, Hosler JP, Hummer G (2015) New perspectives on proton pumping in cellular respiration. *Chem Rev* 115:2196–2221.
22. Efremov RG, Sazanov LA (2012) The coupling mechanism of respiratory complex I: A structural and evolutionary perspective. *Biochim Biophys Acta* 1817:1785–1795.
23. Kotlyar AB, Vinogradov AD (1990) Slow active/inactive transition of the mitochondrial NADH-ubiquinone reductase. *Biochim Biophys Acta* 1019:151–158.
24. Maklashina E, Kotlyar AB, Cecchini G (2003) Active/de-active transition of respiratory complex I in bacteria, fungi, and animals. *Biochim Biophys Acta* 1606:95–103.
25. Ohnishi T (1998) Iron-sulfur clusters/semiquinones in complex I. *Biochim Biophys Acta* 1364:186–206.
26. Hinchliffe P, Sazanov LA (2005) Organization of iron-sulfur clusters in respiratory complex I. *Science* 309:771–774.
27. Verkhovskaya ML, Belevich N, Euro L, Wikström M, Verkhovsky MI (2008) Real-time electron transfer in respiratory complex I. *Proc Natl Acad Sci USA* 105:3763–3767.
28. de Vries S, Dörner K, Strampraad MJ, Friedrich T (2015) Electron tunneling rates in respiratory complex I are tuned for efficient energy conversion. *Angew Chem Int Ed Engl* 54:2844–2848.
29. Tocilescu MA, et al. (2010) The role of a conserved tyrosine in the 49-kDa subunit of complex I for ubiquinone binding and reduction. *Biochim Biophys Acta* 1797:625–632.
30. Sharma V, et al. (2015) Redox-induced activation of the proton pump in the respiratory complex I. *Proc Natl Acad Sci USA* 112:11571–11576.
31. Brandt U (2011) A two-state stabilization-change mechanism for proton-pumping complex I. *Biochim Biophys Acta* 1807:1364–1369.
32. Berrisford JM, Sazanov LA (2009) Structural basis for the mechanism of respiratory complex I. *J Biol Chem* 284:29773–29783.
33. Screpanti E, Hunte C (2007) Discontinuous membrane helices in transport proteins and their correlation with function. *J Struct Biol* 159:261–267.
34. Kaila VRI, Wikström M, Hummer G (2014) Electrostatics, hydration, and proton transfer dynamics in the membrane domain of respiratory complex I. *Proc Natl Acad Sci USA* 111:6988–6993.
35. Belevich G, Knuuti J, Verkhovsky MI, Wikström M, Verkhovskaya M (2011) Probing the mechanistic role of the long α-helix in subunit L of respiratory Complex I from Escherichia coli by site-directed mutagenesis. *Mol Microbiol* 82:1086–1095.
36. Torres-Bacete J, Sinha PK, Matsuno-Yagi A, Yagi T (2011) Structural contribution of C-terminal segments of NuoL (ND5) and NuoM (ND4) subunits of complex I from Escherichia coli. *J Biol Chem* 286:34007–34014.
37. Zhu S, Vik SB (2015) Constraining the lateral helix or respiratory complex I by cross-linking does not impair enzyme activity or proton translocation. *J Biol Chem* 290: 20761–20773.
38. Pomès R, Roux B (1996) Structure and dynamics of a proton wire: a theoretical study of $H^+$ translocation along the single-file water chain in the gramicidin A channel. *Biophys J* 71:19–39.
39. Hummer G, Rasaiah JC, Noworyta JP (2001) Water conduction through the hydrophobic channel of a carbon nanotube. *Nature* 414:188–190.
40. Kaila VRI, Hummer G (2011) Energetics and dynamics of proton transfer reactions along short water wires. *Phys Chem Chem Phys* 13:13207–13215.
41. Jardetzky O (1966) Simple allosteric model for membrane pumps. *Nature* 211:969–970.
42. Forrest LR, et al. (2008) Mechanism for alternating access in neurotransmitter transporters. *Proc Natl Acad Sci USA* 105:10338–10343.
43. Nakamaru-Ogiso E, et al. (2010) The membrane subunit NuoL(ND5) is involved in the indirect proton pumping mechanism of Escherichia coli complex I. *J Biol Chem* 285: 39070–39078.
44. Amarneh B, Vik SB (2003) Mutagenesis of subunit N of the Escherichia coli complex I. Identification of the initiation codon and the sensitivity of mutants to decylubiquinone. *Biochemistry* 42:4800–4808.
45. Euro L, Belevich G, Verkhovsky MI, Wikström M, Verkhovskaya M (2008) Conserved lysine residues of the membrane subunit NuoM are involved in energy conversion by the proton-pumping NADH:ubiquinone oxidoreductase (Complex I). *Biochim Biophys Acta* 1777:1166–1172.
46. Michel J, DeLeon-Rangel J, Zhu S, Van Ree K, Vik SB (2011) Mutagenesis of the L, M, and N subunits of Complex I from Escherichia coli indicates a common role in function. *PLoS One* 6:e17420.
47. Torres-Bacete J, Sinha PK, Castro-Guerrero N, Matsuno-Yagi A, Yagi T (2009) Features of subunit NuoM (ND4) in Escherichia coli NDH-1: TOPOLOGY AND IMPLICATION OF CONSERVED GLU144 FOR COUPLING SITE 1. *J Biol Chem* 284:33062–33069.
48. Torres-Bacete J, Nakamaru-Ogiso E, Matsuno-Yagi A, Yagi T (2007) Characterization of the NuoM (ND4) subunit in Escherichia coli NDH-1: Conserved charged residues essential for energy-coupled activities. *J Biol Chem* 282:36914–36922.
49. Steimle S, et al. (2012) Asp563 of the horizontal helix of subunit NuoL is involved in proton translocation by the respiratory complex I. *FEBS Lett* 586:699–704.
50. Zhu S, Canales A, Bedair M, Vik SB (2016) Loss of Complex I activity in the Escherichia coli enzyme results from truncating the C-terminus of subunit K, but not from cross-linking it to subunits N or L. *J Bioenerg Biomembr* 48:325–333.
51. Cao Z, et al. (2010) Mechanism of fast proton transport along one-dimensional water chains confined in carbon nanotubes. *J Am Chem Soc* 132:11395–11397.
52. Agmon N (1995) The Grotthuss mechanism. *Chem Phys Lett* 244:456–462.
53. Sugita Y, Kitao A, Okamoto Y (2000) Multidimensional replica-exchange method for free energy calculations. *J Chem Phys* 113:6042–6051.
54. Kieseritzky G, Knapp EW (2008) Optimizing pKa computation in proteins with pH adapted conformations. *Proteins* 71:1335–1348.
55. Ohnishi T, Nakamaru-Ogiso E, Ohnishi ST (2010) A new hypothesis on the simultaneous direct and indirect proton pump mechanisms in NADH-quinone oxidoreductase (complex I). *FEBS Lett* 584:4131–4137.
56. Treberg JR, Brand MD (2011) A model of the proton translocation mechanism of complex I. *J Biol Chem* 286:17579–17584.
57. Verkhovsky M, Bloch DA, Verkhovskaya M (2012) Tightly-bound ubiquinone in the *Escherichia coli* respiratory complex I. *Biochim Biophys Acta* 1817:1550–1556.
58. Smart OS, Neduvelil JG, Wang X, Wallace BA, Sansom MSP (1996) HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *J Mol Graph* 14:354–360, 376.
59. Phillips JC, et al. (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26: 1781–1802.
60. MacKerell AD, et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102:3586–3616.
61. Klauda JB, et al. (2010) Update of the CHARMM all-atom additive force field for lipids: Validation on six lipid types. *J Phys Chem B* 114:7830–7843.
62. Grossfield A (2013) WHAM: the weighted histogram analysis method, version 2.0.9. Available at membrane.urmc.rochester.edu/content/wham. Accessed November 15, 2013.
63. Kumar A, Bouzida D, Swendsen RH, Kollman PA, Rosenberg JM (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J Comput Chem* 13:1011–1021.
64. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. *J Mol Graph* 14:33–8, 27–28.
65. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 98:10037–10041.
66. Rabenstein B, Knapp EW (2001) Calculated pH-dependent population and protonation of carbon-monoxy-myoglobin conformers. *Biophys J* 80:1141–1150.
67. Becke AD (1988) Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys Rev A Gen Phys* 38:3098–3100.
68. Perdew JP (1986) Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys Rev B Condens Matter* 33:8822–8824.
69. Sierka M, Hogekamp A, Ahlrichs R (2003) Fast evaluation of the Coulomb potential for electron densities using multipole accelerated resolution of identity approximation. *J Chem Phys* 118:9136–9148.
70. Weigend F, Ahlrichs R (2005) Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys Chem Chem Phys* 7:3297–3305.
71. Ahlrichs R, Bär M, Häser M, Hom H, Kölmel C (1989) Electronic structure calculations on workstation computers. *Chem Phys Lett* 162:165–169.
72. Riahi S, Rowley CN (2014) The CHARMM-TURBOMOLE interface for efficient and accurate QM/MM molecular dynamics, free energies, and excited state properties. *J Comput Chem* 35:2076–2086.
73. Brooks BR, et al. (2009) CHARMM: The biomolecular simulation program. *J Comput Chem* 30:1545–1614.

# Article II

# Correlating kinetic and structural data on ubiquinone binding and reduction by respiratory complex I

Justin G. Fedor[a], Andrew J. Y. Jones[a], Andrea Di Luca[b], Ville R. I. Kaila[b], and Judy Hirst[a,1]

[a]Medical Research Council Mitochondrial Biology Unit, University of Cambridge, Cambridge, CB2 0XY, United Kingdom; and [b]Department of Chemistry, Technical University of Munich, D-85747 Garching, Germany

Edited by Douglas C. Rees, Howard Hughes Medical Institute, California Institute of Technology, Pasadena, CA, and approved October 16, 2017 (received for review August 9, 2017)

**Respiratory complex I (NADH:ubiquinone oxidoreductase), one of the largest membrane-bound enzymes in mammalian cells, powers ATP synthesis by using the energy from electron transfer from NADH to ubiquinone-10 to drive protons across the energy-transducing mitochondrial inner membrane. Ubiquinone-10 is extremely hydrophobic, but in complex I the binding site for its redox-active quinone headgroup is ∼20 Å above the membrane surface. Structural data suggest it accesses the site by a narrow channel, long enough to accommodate almost all of its ∼50-Å isoprenoid chain. However, how ubiquinone/ubiquinol exchange occurs on catalytically relevant timescales, and whether binding/dissociation events are involved in coupling electron transfer to proton translocation, are unknown. Here, we use proteoliposomes containing complex I, together with a quinol oxidase, to determine the kinetics of complex I catalysis with ubiquinones of varying isoprenoid chain length, from 1 to 10 units. We interpret our results using structural data, which show the hydrophobic channel is interrupted by a highly charged region at isoprenoids 4–7. We demonstrate that ubiquinol-10 dissociation is not rate determining and deduce that ubiquinone-10 has both the highest binding affinity and the fastest binding rate. We propose that the charged region and chain directionality assist product dissociation, and that isoprenoid stepping ensures short transit times. These properties of the channel do not benefit the exhange of short-chain quinones, for which product dissociation may become rate limiting. Thus, we discuss how the long channel does not hinder catalysis under physiological conditions and the possible roles of ubiquinone/ubiquinol binding/dissociation in energy conversion.**

bioenergetics | coenzyme Q10 | electron transport chain | mitochondria | NADH:ubiquinone oxidoreductase

Respiratory complex I (NADH:ubiquinone oxidoreductase) (1) is a major entry point to the electron transport chain of oxidative phosphorylation in mammalian mitochondria. It catalyzes NADH oxidation coupled to ubiquinone reduction and captures the free energy produced to transport protons (2, 3) across the mitochondrial inner membrane, supporting ATP synthesis and transport processes. Because complex I is essential for regenerating $NAD^+$ to sustain the tricarboxylic acid cycle and fatty acid oxidation, and an important contributor to cellular reactive oxygen species production (4), mutations in its subunits and assembly factors cause a wide range of inherited neuromuscular and metabolic diseases (5).

Mammalian complex I is a large (1 MDa) membrane-bound enzyme of 45 subunits. Due to advances in single-particle electron cryomicroscopy (cryoEM), knowledge of its structure (6–8) has surged forward recently, and descriptions of the 14 core subunits, in mammalian (6–8), fungal (9) and bacterial (10) enzymes, have laid a new foundation for mechanistic studies. NADH is oxidized by a flavin near the top of the hydrophilic domain of the L-shaped complex. Then, electrons are transferred along a chain of seven iron–sulfur (FeS) clusters to ubiquinone, bound at the interface of the hydrophilic and membrane domains. The membrane domain contains four antiporter-like units, considered to each transport one proton per cycle. They are connected by elements indicative of ion-transport activities, including π-bulges, loops in transmembrane helices (TMHs), and a series of buried charged residues, which also connects them to the quinone-binding region. Although molecular simulations have suggested how conformational, protonation, and hydration changes could propagate through the membrane domain to drive proton transfer events (11), how the energy from quinone reduction is captured and transferred to proton translocation is currently unknown. Reactions that may initiate the proton-transfer cascade include movement of a conserved aspartate upon quinone reduction (12), double reduction of Q to $Q^{2-}$ or its subsequent protonation to $QH_2$ (13), and/or quinone binding/quinol dissociation. It has also been suggested that a permanently bound quinone shuttles between two positions in the channel, requiring an additional site for the exchangeable substrate to bind (14).

Here, we focus on ubiquinone binding and reduction by mammalian complex I. Strikingly, the binding site for the redox-active ubiquinone headgroup is ∼20 Å above the membrane interface and thought to be accessed by a long, narrow channel (Fig. 1) that has been identified in all of the structures described so far, but not yet confirmed experimentally. The headgroup binds in a cleft between the 49 kDa and PSST subunits (we use the nomenclature for the bovine enzyme throughout) and both mutational studies in *Yarrowia lipolytica* (15) and *Escherichia coli* (16), and densities observed in structural data from *Thermus thermophilus* complex I (10), indicate that it forms hydrogen bonds with $H59^{49\ kDa}$ and $Y108^{49\ kDa}$, placing it within 12 Å of the terminal FeS cluster, N2.

## Significance

Respiratory complex I, a redox-coupled proton pumping enzyme, is central to aerobic metabolism in mammalian mitochondria and implicated in many neuromuscular disorders. One of its substrates, ubiquinone-10, binds in an unusually long and narrow channel, which is at the intersection of the enzyme's electron and proton transfer modules and a hotspot for disease-causing mutations. Here, we use a minimal, self-assembled respiratory chain to study complex I catalyzing with ubiquinones of different isoprenoid chain lengths. We show that the channel enhances the affinity of long-chain quinones, assists in their transfer along the channel, and organizes them for product release. Finally, we discuss how efficient binding and dissociation processes may help to link redox catalysis to proton pumping for energy conversion.

Author contributions: J.G.F. and J.H. designed research; J.G.F., A.J.Y.J., A.D.L., and V.R.I.K. performed research; J.G.F., A.D.L., V.R.I.K., and J.H. analyzed data; and J.G.F. and J.H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.



Data deposition: The modeled structures with Q1–Q10 bound and with the empty site are available from the University of Cambridge data repository (https://doi.org/10.17863/CAM.13786).

[1]To whom correspondence should be addressed. Email: jh@mrc-mbu.cam.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1714074114/-/DCSupplemental.

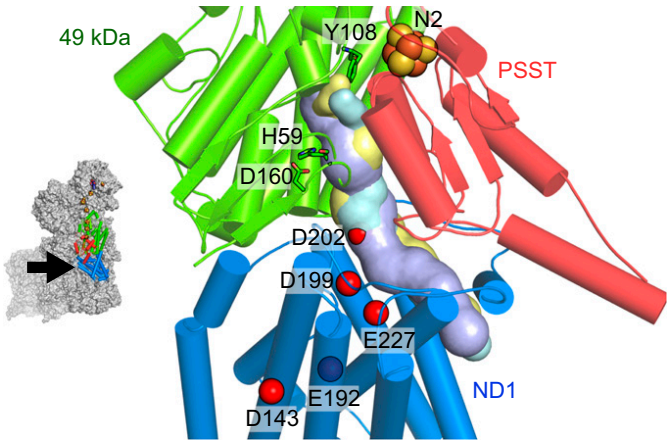


**Fig. 1.** The proposed ubiquinone-binding channel in mammalian complex I. The 49-kDa, PSST, and ND1 subunits from bovine complex I (5LC5.pdb) (6) are in cartoon, with the surfaces of predicted quinone-binding cavities in the aligned structures of the bovine (6) (cyan), porcine (7) (yellow), and *T. thermophilus* (10) (purple) enzymes. The quinone headgroup is considered to hydrogen bond to Y108 and H59, and H59 to D160. The carboxylate groups of acidic residues connecting the quinone-binding region to the proton-pumping subunits are shown by red spheres. Cavities were generated using the Caver 3.0 PyMOL plugin (33) with a 1.4-Å probe.

At the base of the cleft, the predicted channel meets subunit ND1, and runs along its interface with the 49-kDa and PSST subunits before exiting into the membrane; in total, it is long enough to accommodate most of the ∼50-Å-long isoprenoid tail of ubiquinone-10. Many structurally diverse inhibitors are thought to bind in the channel (17) and it is a hotspot for both pathophysiological mutations (ref. 5 and https://www.mitomap.org) and site-directed variants that affect catalysis (18, 19).

Here, we have used proteoliposomes (PLs) to determine the kinetics of complex I catalysis with a series of ubiquinone substrates of varying isoprenoid tail length, from ubiquinone-1 (Q1) to ubiquinone-10 (Q10). Previous attempts to investigate the effects of quinone tail length used native membranes supplemented with exogenous quinones, following removal of the endogenous Q10 by lyophilization and pentane extraction (20–22). However, these studies were compromised because (*i*) the quinone exchange procedures were detrimental to catalysis (the specific activities were 10-fold lower than observed here); and (*ii*) effects on complex I activity were obscured by the catalysis of other enzymes: Lenaz et al. (22) reconstituted Q1 to Q10 into lyophilized pentane-extracted mitochondria and measured $O_2$ consumption, which requires complexes III and IV, whereas Estornell et al. (20) similarly reconstituted Q10 and measured cytochrome *c* reduction, which requires complex III. More recently, Fato et al. (21) assayed lyophilized pentane-extracted bovine mitochondria reconstituted with Q3, Q5, and Q10 with an NADH-Q1 reductase assay. In contrast, the PLs used here (23) contain highly active bovine complex I (CI, to oxidize NADH and reduce ubiquinone to ubiquinol), the cyanide-insensitive nonprotonmotive alternative oxidase from *Trypanosoma brucei brucei* (AOX, to reoxidize ubiquinol to ubiquinone and reduce $O_2$) and varying concentrations of different ubiquinones. The system and its kinetic properties can be defined precisely through measurements of the protein, phospholipid, and ubiquinone contents and have been developed to ensure complex I is rate determining. We compare the kinetic parameters of different ubiquinones with the structural and physicochemical properties of the quinone-binding channel to reveal determinants of catalysis.

## Results

**Ubiquinone Reduction by Complex I Is Rate Limiting for Catalysis.** The CI-AOX PLs described here catalyze NADH:$O_2$ oxidoreduction (the NADH:$O_2$ reaction) by redox cycling ubiquinone/ubiquinol: complex I reduces ubiquinone to ubiquinol and AOX reoxidizes it. To use the NADH:$O_2$ reaction to investigate ubiquinone reduction by complex I it must be rate limiting. Thus, the NADH (200 μM) and $O_2$ (200–250 μM) concentrations used are substantially higher than the $K_M$ values of 79 ± 8 μM for complex I (measured with the NADH:$O_2$ reaction) and 10–20 μM for AOX (24). Furthermore, the $k_{cat}$ value for NADH oxidation by the flavin in bovine complex I is >5,000 $s^{-1}$ (25), more than 10 times faster than the maximum rate of NADH:ubiquinone oxidoreduction, so NADH oxidation does not limit catalysis.

The complex I and AOX used here display similar turnover rates in solution assays with ubiquinone-1, typically 250–300 $s^{-1}$ and 200–250 $s^{-1}$, respectively. Thus, to increase the rate of ubiquinol oxidation, PLs were created starting from a molar ratio of 1:25 CI:AOX. The ratio then effectively doubles to 1:46 ± 1.1 (mean value ± SD for all samples) because NADH can only access complex I oriented with its active site outwards (77 ± 10% of the total), whereas AOX substrates can access it in both orientations, and because AOX incorporates into PLs more efficiently than complex I (80 ± 15% and 59 ± 11%, respectively). This high ratio provides a strong expectation for complex I being rate limiting, and thus for the ubiquinone/ubiquinol pool being predominantly oxidized during catalysis (23). To test this expectation, titrations on the NADH:$O_2$ reaction by Q10-containing PLs with a high (1:51) or low (1:1.5) CI:AOX ratio (defined for outward facing complex I) were performed with complex I (piericidin A)- and AOX (colletochlorin B) (26)-specific inhibitors. For both ratios, piericidin A inhibition builds rapidly at low concentrations (Fig. 2*A*) and the $IC_{50}$ values are similar (2.0 ± 1.1 and 3.4 ± 1.1 nM, or 8.7 and 8.0 piericidin A per oriented complex I, respectively). In contrast, the high ratio PLs were unresponsive to low colletochlorin B concentrations (Fig. 2*B*) and displayed an $IC_{50}$ value 30 times greater (15.9 ± 0.03 and 0.52 ± 0.03 nM, respectively, or 0.7 and 1.4 colletochlorin B per AOX). Thus, in the high ratio (typical) preparation, most of the AOX can be inhibited with little effect on catalysis, confirming that AOX is not rate limiting. High ratio PLs containing Q8, Q6, and Q4, or tested with Q2, all gave colletochlorin B $IC_{50}$ values of 17.6 ± 1.3 nM, matching that for Q10. Only for Q1 was the $IC_{50}$ value for colletochlorin B different; its much higher value of ∼5 μM suggests colletochlorin B is a poor inhibitor of ubiquinol-1 oxidation by AOX, or that Q1 is a poor complex I substrate.

**Measurement of $K_M$ and $k_{cat}$ Values.** Fig. 3 shows Michaelis–Menten curves for complex I catalyzing with Q10, Q8, Q6, Q4, Q2, and Q1. The highly hydrophobic Q10, Q8, Q6, and Q4 molecules [calculated logP values (cyclohexane/water) 19.4, 15.7, 12, and 8.3, respectively (https://www.ncbi.nlm.nih.gov/pccompound)] were incorporated into the PL membranes during reconstitution so each point is from a different preparation. The samples for each curve were prepared together in batches, then their quinone, phospholipid, complex I,

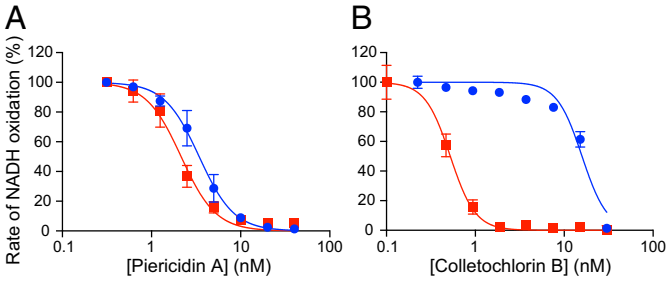


**Fig. 2.** Complex I is rate limiting for catalysis in CI-AOX PLs. Normalized rates of the NADH:$O_2$ reaction by PLs containing high (1:51, blue) and low (1:1.5, red) ratios of oriented-CI:AOX are shown. Both preparations contained ∼10 mM Q10. (*A*) Inhibition by piericidin A. (*B*) Inhibition by colletochlorin B.

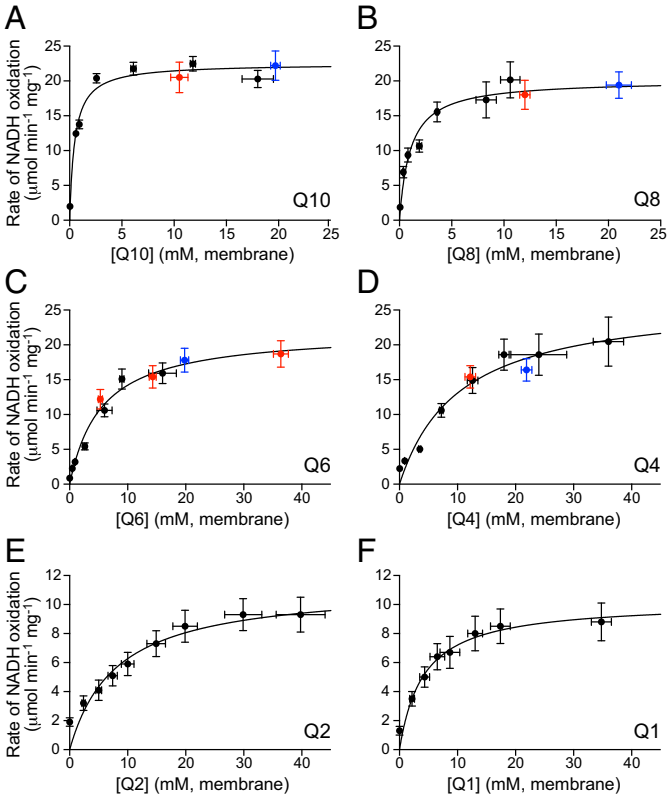


**Fig. 3.** Michaelis–Menten curves for reduction of Q1–Q10 by complex I. (*A*) Q10, (*B*) Q8, (*C*) Q6, (*D*) Q4, (*E*) Q2, and (*F*) Q1. The black points are from sets of PLs with different quinone concentrations; the datasets for Q1–Q8 were adjusted to the Q10 dataset using the red and blue datasets and scaling factors of 1.44 (Q8), 1.35 (Q6), 1.86 (Q4), 1.57 (red points), and 1.60 (blue points). Each value is the mean of at least three replicates ± the SD propagated from each underlying measurement.

and AOX concentrations were determined. Quinone concentrations in the membrane were calculated by assuming that 1 mg of phospholipid occupies ∼1 μL (27), so 1 nmol of quinone per milligram phospholipid is equivalent to 1 mM. The more hydrophilic Q2 and Q1 molecules were added from ethanolic stock solutions to quinone-free PLs and considered to partition between the aqueous and membrane phases. Their membrane concentrations were calculated from the phase volumes and membrane/water logP values, 4.0 for Q2 and 2.9 for Q1 (28). Micelle formation was not considered because the critical micelle concentration of Q2 of 14.0 μM (29) is above the maximum 10-μM concentration added, and it was confirmed that all of the Q2 added could be reduced by complex I. Addition of protonophore uncouplers, such as gramicidin, did not increase the rate so they were not included. Finally, to account for variations in enzyme activity between the quinone-specific batch preparations, two further batches containing samples from each different quinone were prepared (see red and blue points in Fig. 3); the Q1 and Q2 datasets are from quinone-free PLs included in these batches. The Q10 $K_M$ dataset was used as the reference and scaling parameters for the other datasets derived by simultaneous nonlinear least squares fitting; the scaling parameters were applied uniformly and so affect only $V_{max}$, not $K_M$.

**Dependence of $K_M$ and $V_{max}$ on Isoprenoid Chain Length.** The $K_M$ and $V_{max}$ values for each ubiquinone (from Fig. 3) are summarized in Fig. 4. Fig. 4*A* shows that $V_{max}$ (or $k_{cat}$) is biphasic and averages to 23 ± 2 μmol·min$^{-1}$·mg$^{-1}$ (380 s$^{-1}$) for Q10, Q8, Q6, and Q4, and 9.3 ± 0.9 μmol·min$^{-1}$·mg$^{-1}$ (150 s$^{-1}$) for Q2 and Q1. Previous studies, using pentane-extracted mitochondria, also observed the highest NADH:$O_2$ activities from Q7 to Q10, with Q1 to Q4 supporting only ∼30% of the Q10 value (21, 22). Fig. 4*B* shows that $K_M$ displays a bell-shaped curve with the highest value at Q4. We note that the Q10 $K_M$ value reported here of 0.48 mM is lower than that of 3.9 mM we measured previously (23) due to (*i*) improvements to the AOX preparation that have increased its specific activity two- to threefold, (*ii*) improvements to our quinone quantification protocol, and (*iii*) the omission of alamethicin. Alamethicin was used previously (23) to open pores in the membrane to allow NADH to access all of the complex I, but we have now found that 11.25 μg·mL$^{-1}$ increases the apparent Q10 $K_M$ value from 0.5 ± 0.1–1.3 ± 0.4 mM. Alamethicin has been reported to sequester membrane-bound fatty acids (30) so it may sequester quinones also, and to induce structural changes in the membrane (31). Finally, Fig. 4*C* shows that the catalytic efficiency or pseudo second order rate constant, $k_{cat}/K_M$, increases markedly for Q8 and Q10. Previously, $k_{cat}/K_M$ was reported to exhibit a bell-shaped dependence on substrate hydrophobicity (32) but the earlier study used only quinones with alkyl chains up to 11 carbons long, whereas the same trend is not replicated here using polyisoprenoid ubiquinones with partition coefficients that vary much more.

**Analysis of Structural Data and Docking of Q10.** To investigate how our data correlate with structural features in complex I, we used the Caver software (33) to detect and compare the proposed quinone-binding sites in available structures. Fig. S1 shows that the bovine (6), porcine (7), *T. thermophilus* (10), and *Y. lipolytica* (9) structures describe a common channel. The same channel is truncated in the ovine structure (8) by the 49-kDa subunit β1–β2 loop, and similarly constricted in the *Y. lipolytica* structure (9). Variations in the channels may arise from their medium resolutions and/or different enzyme states; both the ovine and *Y. lipolytica* enzymes were proposed to be in the "deactive" state (8, 9) that is unable to reduce quinone, whereas the cryoEM dataset for the bovine structure (6) was classified into three states and only the structure allocated to the "active" state was used here. The bovine, porcine, and *T. thermophilus* channels are overlaid in Fig. 1, highlighting their similarity.

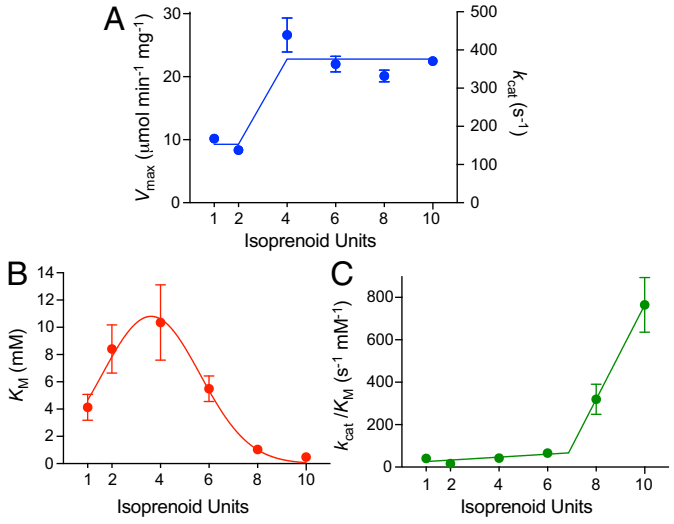


**Fig. 4.** Dependence of the Michaelis–Menten parameters on isoprenoid chain length. (*A*) $V_{max}$ and $k_{cat}$ values shown with average values of 9.3 ± 0.9 μmol·min$^{-1}$·mg$^{-1}$ (150 ± 15 s$^{-1}$) for Q1–Q2 and 23 ± 2 μmol·min$^{-1}$·mg$^{-1}$ (380 ± 39 s$^{-1}$) for Q4–Q10. (*B*) $K_M$ values; the curve is only to guide the eye. (*C*) $k_{cat}/K_M$ values with two linear fits (gradients 7.0 and 220 mM$^{-1}$·s$^{-1}$·isoprenoid unit$^{-1}$). Values are from the data in Fig. 3 and also given in Table S1.

To better define how quinones bind in the channel, we modeled Q10 into the channel identified in the bovine complex [Protein Data Bank (PDB) ID: 5LC5] (6) and relaxed the system by atomistic molecular dynamics (MD) simulations (Fig. 5*A*). Subsequently, by truncating the relaxed Q10 model, we performed further MD simulations for Q8, Q6, Q4, Q2, and Q1 (Fig. 5*B*). Analyses of the Q10-bound structure showed that its channel closely matches the original channel (Fig. S1) and, in all cases, the modeled protein structures around the bound quinones remain in similar conformations except that, for Q1–Q6 and the empty site, the sidechain of $F224^{ND1}$ moves into the position of isoprenoid-7 (Fig. S2*A*). Notably, the root-mean-square fluctuations (RMSFs) of the bound quinone variants suggest that the long isoprenoid chains of Q8 and Q6, which overlay closely on that of Q10 (Fig. 5*B*), are spatially tightly constrained by the channel, whereas the shorter chains of Q4, Q2, and Q1 have much greater conformational freedom (Fig. 5*C*).

Strikingly, Fig. 6 shows how the properties of residues close to the modeled Q10 vary along it (Fig. S3 presents equivalent data for Q8–Q1). The environment of the headgroup plus isoprenoids 1–3, in the cleft between the PSST and 49-kDa subunits, is primarily hydrophobic and uncharged. Conversely, that of isoprenoids 4–7 contains many charged residues. In particular, a group of highly conserved arginines (Fig. 5*A*) form a channel elbow that produces a ~100° kink in the modeled Q10 between isoprenoids 4 and 5. Notably, the hydrophobic faces of the Arg guanidiniums form π-stacking interactions with the isoprenoids, while their edges form polar and electrostatic interactions with a set of conserved Glu/Asp residues (Fig. 5*A*), consistent with the known behavior of guanidium groups (34). In the modeled structures, interactions between the charged residues separate them into two groups (Fig. S2*B*), with the group involving residues on ND1 TMH1 showing less positional variation than that involving primarily residues on the ND1 TMH5–6 loop (Fig. S2*C*),

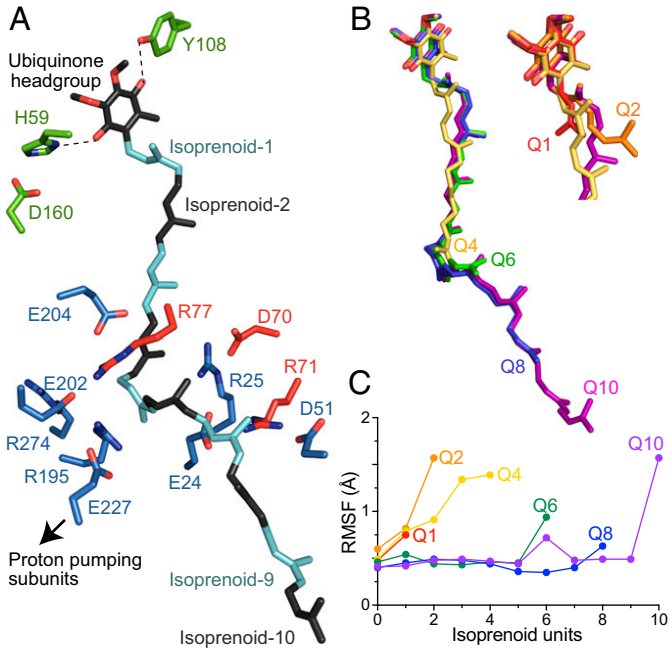


**Fig. 5.** Modeled structures of bovine complex I with Q1–Q10 bound. (*A*) The modeled Q10 molecule with its isoprenoids in black and cyan, alternately, alongside the sidechains of charged residues within 5 Å (red, PSST; blue, ND1). Green: 49-kDa subunit residues hydrogen bonded to the ubiquinone headgroup. (*B*) The overlaid modeled structures for Q10, Q8, Q6, and Q4, and *Inset* for Q10, Q4, Q2, and Q1. (*C*) RMSFs for each quinone species from MD simulations, colored as in *B*.

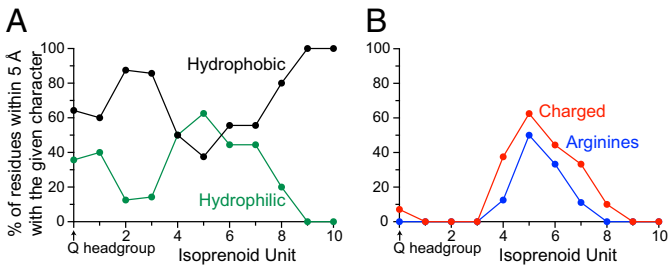


**Fig. 6.** The protein environment of the modeled bound Q10 molecule in the structure of bovine CI. (*A*) Percentage of residues within 5 Å of each isoprenoid that are hydrophobic (A, F, I, L, M, P, V, W, and Y) or hydrophilic (C, D, E, G, H, K, N, Q, R, S, and T). (*B*) Percentage of residues within 5 Å of each isoprenoid that are canonically charged (D, E, H, K, and R) and that are arginines.

noted previously for its flexibility (6). Finally, for isoprenoids 8 and 9 the expected hydrophobic nature of the ubiquinone-binding channel is reestablished.

## Discussion

**Entry and Exit of Quinone/Quinol to and from the Channel.** In the classical Michaelis–Menten model, $k_{cat}$ encapsulates all of the steps that follow formation of the enzyme–substrate complex, so the fact that $k_{cat}$ is constant for Q4–Q10 (Fig. 4*A*) suggests that product dissociation, which is expected to be chain-length dependent, is not rate limiting. Alternative rate-limiting steps include the electron–proton transfers for quinone reduction, the coupled processes leading to proton translocation, and all of the reactions required to complete the catalytic cycle.

For Q2 and Q1, $k_{cat}$ (150 $s^{-1}$) is substantially lower than for Q4–Q10 (380 $s^{-1}$) (Fig. 4*A*) so the rate-limiting step has either slowed or changed identity. Both possibilities indicate that a chain-length dependent step, such as product dissociation, has slowed down. Furthermore, the pseudo second-order rate constant $k_{cat}/K_M$ [where $K_M = (k_{off} + k_{cat})/k_{on}$] provides a guide to the rate constant for substrate binding ($k_{on}$) since it approximates to it when $k_{cat} > k_{off}$ and reflects how fast the rate increases with increasing concentration, in substrate-limiting conditions. $k_{cat}/K_M$ increases substantially with isoprenoid tail length (Fig. 4*C*). Both observations are consistent with rates of binding and dissociation (i.e., channel transit rates) increasing with isoprenoid chain length, and being limiting only for Q2 and Q1. We propose three explanations. First, the short isoprenoid tails of Q1, Q2, and Q4 do not overlay their respective Q10 isoprenoids (Fig. 5*B*) so they are conformationally mobile within the site (Fig. 5*C*): their dissociation may be hindered by lack of directionality, due to lack of a guiding anchor extending down the channel. The tail of Q10 extends into the membrane and may guide the entire dissociation process. Second, for the shorter quinones, an additional molecule(s) may enter the channel behind the substrate, impeding its dissociation. Third, the channel entrance is at a similar depth below the membrane interface to the favored position for the ubiquinone-10 headgroup (35), whereas shorter tail quinones may distribute differently in the membrane, affecting their binding rates.

Finally, while the simple Michaelis–Menten mechanism provides a convenient framework for interpreting our data, it does not account for the unusual nature of the ubiquinone-binding channel in complex I. Transfer along the whole channel is unlikely to occur over a single transition state but over a complex energy surface that may form local minima and transient binding sites. Longer chain species may move faster along the channel because their stepwise movements incur minimal changes. For example, moving a Q10 molecule by one step requires only the residues (and water molecules) around the headgroup and the

leading isoprenoid to reorganize (at all other positions one isoprenoid simply replaces another), whereas moving a short chain quinone requires the region behind it to reorganize also. The energy surface for Q10 transit along the channel may therefore be flatter. In this way, the unusually long substrate-binding channel in complex I need not exert a rate-limiting effect on catalysis.

**Effect of Isoprenoid Chain Length on Binding Affinity.** In classical Michaelis–Menten kinetics, $K_M$ is expressed as $(k_{off} + k_{cat})/k_{on}$ so, for Q4–Q10, where $k_{cat}$ is constant (Fig. 4*A*) but $K_M$ decreases with increasing chain length (Fig. 4*B*), $k_{off}$ decreases and/or $k_{on}$ increases. Both possibilities suggest that $K_d$ (= $k_{off}/k_{on}$) decreases, such that Q4 binds most weakly and Q6, Q8, and Q10 progressively more tightly. The comparison does not extend to Q2 and Q1 because $k_{cat}$ is different, but as their $k_{cat}$ values are similar, the same considerations suggest Q1 binds more tightly than Q2. Thus, although a robust comparison between Q4 and Q2 is currently not possible, we may infer that the curve describing $K_d$ qualitatively resembles that of $K_M$.

The relative free-energy changes that result from moving each quinone species from a low dielectric (membrane) environment into its complex I binding site were evaluated using continuum electrostatics calculations (36), revealing a linear dependence on isoprenoid number (Fig. S4). The trend is consistent with the decreasing $K_d$ proposed for Q4–Q10, but not with the decrease proposed for Q4–Q1. Intriguingly, the discrepancy can be explained by considering that $k_{off}$ for Q4, Q2, and Q1 may be decreased, as discussed above, by their lacking a guiding chain for dissociation and by additional substrate molecules blocking their exit. By decreasing $k_{off}$, these factors "cage" short chain species in the binding site, independently of their intrinsic affinity for it, and decrease $K_d$. Notably, many tight-binding hydrophobic complex I inhibitors match the dimensions of Q1–Q4 and are considered to occupy the same binding site; density attributed to piericidin A, which resembles Q3, was observed to overlay density attributed to decylubiquinone in crystallographic maps from *T. thermophilus* complex I (10). Thus, the potent inhibition of these molecules may also partly result from caging effects.

The concept of a highly charged channel for the hydrophobic isoprenoid chain, as observed for isoprenoids 4–7 (Figs. 5 and 6), is intrinsically challenging. Charged cavities in proteins tend to fill with water molecules (37), and indeed, in MD simulations on the structure of *T. thermophilus* complex I, waters accumulate in the channel, especially in the charged region (11). If waters are present it may be that (*i*) the channel has become hydrated artificially during the extended handling required for structural work, or (*ii*) during every catalytic cycle the quinone displaces waters, making an entropic contribution to offset the enthalpy loss of breaking charged and polar interactions. In fact, an unfavorable binding enthalpy from the charged region may mitigate the increasing affinity due to binding an increasing number of hydrophobic units within a hydrophobic channel and be important for efficient product release. Higher resolution structural data are required for further understanding of this intriguing structural feature.

Finally, increased contributions to the binding affinity from the second, nonpolar region at isoprenoids 8 and 9 cause a marked increase in $k_{cat}/K_M$ (Fig. 4*C*). Because $k_{cat}$ is constant from Q4 to Q10, $k_{cat}/K_M$ is dominated by the reciprocal of $K_M$ and the apparent discontinuity in Fig. 4*C* results simply from a point of inflection in the $K_M$ curve.

**How Is Proton Translocation Coupled to Quinone Reduction?** The identity of the coupling point, at which the redox reaction initiates proton translocation, is currently the most important unknown feature of the complex I mechanism. Molecular dynamics simulations of the movement of $D160^{49\ kDa}$ away from $H59^{49\ kDa}$, as a result of proton transfer from H59 to the nascent quinol, have been used to illustrate one possibility (12). Alternatively, could conformational changes triggered by quinone/quinol moving along the channel, particularly through the charged section, trigger proton translocation? If so, short chain quinones must activate the mechanism as effectively as long chain quinones, since they elicit the same proton-pumping stoichiometry (2, 3). As Fig. 5*A* shows, the charged residues surrounding the channel are, like D160, connected to the center of the membrane domain by a chain of acidic residues (6–10, 12), suggesting how quinone/quinol binding/dissociation may be communicated to the proton translocation machinery. To elucidate, challenge, and add to proposals for the mechanistic coupling point in complex I require higher resolution structures set in different states, alongside mutational, functional, and computational studies to provide complementary strategies and perspectives on tackling this difficult problem.

## Experimental Methods

**Preparation of AOX.** The construct described previously for overexpression of AOX from *T. brucei brucei* (38) was modified by removing the N-terminal mitochondrial targeting peptide (residues 1–24) and replacing the 6xHis tag with a Twin-Strep tag (IBA GmbH) (39). The modified AOX was overexpressed in *E. coli* strain FN102 (38) in a 60-L fermenter as described previously (23, 38), except by using 100 μg $mL^{-1}$ ampicillin instead of carbenicillin. The culture was incubated at 30 °C, 60% $O_2$ saturation until $OD_{600}$ ∼0.6, then expression induced with 25 μM isopropyl β-D-1-thiogalactopyranoside for 12 h. Cells (200–300 g) were collected by centrifugation, then membranes were prepared immediately (23) and resuspended to ∼30 mg protein $mL^{-1}$ in 50 mM Tris·HCl (pH 8.0 at 4 °C) for storage at −80 °C.

All of the following steps were at 4 °C: 20 mL of membranes were solubilized at 6 mg protein $mL^{-1}$ in 25 mM Tris·HCl (pH 8.0), 200 mM $MgSO_4$, 1.4% (wt/vol) *n*-octyl-glucopyranoside (Anagrade, Anatrace) and 20% (vol/vol) glycerol for 1 h then centrifuged (165,000 × *g*, 30 min). The supernatant was loaded onto a ∼8-mL column of Strep-Tactin Superflow high capacity resin (IBA GmbH) preequilibrated in strep buffer: 20 mM Tris·HCl (pH 8.0 at 4 °C), 50 mM $MgSO_4$, 160 mM NaCl and 20% (vol/vol) glycerol (39). AOX was eluted in strep buffer supplemented with 2.5 mM desthiobiotin (Sigma-Aldrich) and 0.042% *n*-dodecyl-β-D-maltopyranoside (DDM) (Anagrade, Anatrace) (39). Pooled AOX-containing fractions were concentrated 10-fold, then dialyzed for 6 h against 2 L of strep buffer plus 0.042% DDM (23). Typical preparations yielded 3–5 mg AOX and were ≥95% pure by SDS/PAGE.

**Preparation of Complex I.** Mitochondrial membranes were prepared from *Bos taurus* (bovine) heart (40) then complex I was prepared as described previously (23), with minor modifications. Membranes were solubilized at 5 mg $mL^{-1}$ protein in 1% DDM. The Q-Sepharose buffers consisted of 20 mM Tris·HCl (pH 7.55 at 4 °C), 2 mM EDTA, 10% (vol/vol) ethylene glycol, 0.2% (wt/vol) DDM, 0.02% asolectin (total soy lipid extract, Avanti Polar Lipids), and 0.02% 3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate (CHAPS). Size exclusion chromatography was conducted on a Superose 6 Increase 10/300 column (GE Healthcare) in 20 mM Tris·HCl (pH 7.55 at 4 °C), 200 mM NaCl, 0.05% DDM, and 10% (vol/vol) glycerol. Samples were flash frozen for storage at ∼20 mg·$mL^{-1}$ in 30% (vol/vol) glycerol.

**Preparation of CI-AOX PLs.** Chloroform solutions (25 mg $mL^{-1}$) of bovine heart phosphatidylcholine (PC), phosphatidylethanolamine (PE) and cardiolipin (CL) were from Avanti Polar Lipids. Chloroform stock solutions of Q10 (Sigma Aldrich), Q8 (Avanti Polar Lipids), Q6 (Avanti Polar Lipids) and Q4 (Santa Cruz Biotechnology) were created at 2–4 mM. PLs were prepared as described previously (23), starting from mixtures of 8 mg PC, 1 mg PE, 1 mg CL and quinone (as required).

**Characterization of PLs.** To quanitify the quinone present, 90 μL of ethanol (HPLC grade Chromasolv; Sigma-Aldrich) was added to 10 μL of PL solution, sonicated (1 min) and centrifuged (16,300 × *g*, 10 min). A total of 50 μL of supernatant was injected onto a Nucleosil 100–5C18 (Hichrom) column, and run at 30 °C at 800 μL·$min^{-1}$ on an Agilent 1100 series HPLC in 70% ethanol, 30% methanol, 0.07% $HClO_4$, and 50 mM $NaClO_4$ for Q10 and Q8, or in 100% methanol, 0.07% $HClO_4$ and 50 mM $NaClO_4$ for Q6 and Q4 (retention times 9, 6, 10, and 5 min, respectively). Concentrations were determined by comparison with known standards. Total phospholipid contents were determined as detailed previously (23). Total protein contents were quantified by the amido black assay, which is insensitive to detergents and high phospholipid concentrations (41). The NADH:$APAD^+$ oxidoreduction assay

BIOCHEMISTRY

was used to determine the total complex I content and orientation together with 15 $\mu g \cdot mL^{-1}$ alamethicin to allow NADH into the PL lumen (23). The AOX content was taken as the difference between the total protein and complex I contents.

**Catalytic Activity Assays.** Activity assays were at 32 °C in 10 mM Tris-$SO_4$ (pH 7.5 at 32 °C), 50 mM KCl, 200 μM NADH with ∼1.5 mg protein $mL^{-1}$ PLs (23). NADH oxidation was monitored at 340–380 nm ($\varepsilon$ = 4.81 $mM^{-1} \cdot cm^{-1}$) on a Molecular Devices Spectramax 384 Plus platereader. Q2 and Q1 (Sigma-Aldrich) were added in ethanol. Inhibitor-insensitive rates were determined in 0.5 μM piericidin A as 0.30 ± 0.03 μmol NADH $min^{-1} \cdot mg\ Cl^{-1}$ for all quinones except Q1 (which exhibited a concentration-dependent inhibitor-insensitive rate) and have been subtracted from the data reported.

**Molecular Modeling of Ubiquinone-Bound Structures.** Ubiquinone-10 was docked into a channel identified by the HOLE software (42) in the structure of bovine complex I [PDB ID: 5LC5 (6)] with missing amino acid sidechains modeled. Then, 20 ns of classical MD simulation (integration time step 2 fs, 310 K) were performed using NAMD2 (43) with the CHARMM36 force field (44), and density functional theory (DFT)-derived parameters for all cofactors (12). Backbone atoms of the 49 kDa, ND1, and PSST subunits were harmonically restrained (force constant 2 $kcal \cdot mol^{-1} \cdot Å^{-2}$), with positional constraints for all other subunits. The A57–V61 and L106–S110 loops of the 49-kDa subunit that carry H59 and Y108 were not restrained. For the Q1–Q8-bound structures, the Q10 molecule was truncated and each variant relaxed for a further 10 ns. Free energies for quinone solvation into complex I were estimated by the molecular mechanics–generalized Born surface area (MM-GBSA) method (36), using 50 snapshots extracted from 10-ns MD simulations.

**ACKNOWLEDGMENTS.** We thank C. Humphreys & Sons Abattoir (Chelmsford) for providing bovine hearts and Sotiria Tavoulari for help with the amido black assay. Computing time was provided by the SuperMuc at The Leibniz Rechenzentrum (Grant pr48de). This work was supported by the Medical Research Council (Grant U105663141 to J.H.) and by the German Research Foundation (V.R.I.K.).

1. Hirst J (2013) Mitochondrial complex I. *Annu Rev Biochem* 82:551–575.
2. Jones AJY, Blaza JN, Varghese F, Hirst J (2017) Respiratory complex I in *Bos taurus* and *Paracoccus denitrificans* pumps four protons across the membrane for every NADH oxidized. *J Biol Chem* 292:4987–4995.
3. Galkin AS, Grivennikova VG, Vinogradov AD (1999) →$H^+$/2e- stoichiometry in NADH-quinone reductase reactions catalyzed by bovine heart submitochondrial particles. *FEBS Lett* 451:157–161.
4. Murphy MP (2009) How mitochondria produce reactive oxygen species. *Biochem J* 417:1–13.
5. Fassone E, Rahman S (2012) Complex I deficiency: Clinical features, biochemistry and molecular genetics. *J Med Genet* 49:578–590.
6. Zhu J, Vinothkumar KR, Hirst J (2016) Structure of mammalian respiratory complex I. *Nature* 536:354–358.
7. Wu M, Gu J, Guo R, Huang Y, Yang M (2016) Structure of mammalian respiratory supercomplex $I_1III_2IV_1$. *Cell* 167:1598–1609.
8. Fiedorczuk K, et al. (2016) Atomic structure of the entire mammalian mitochondrial complex I. *Nature* 538:406–410.
9. Zickermann V, et al. (2015) Structural biology. Mechanistic insight from the crystal structure of mitochondrial complex I. *Science* 347:44–49.
10. Baradaran R, Berrisford JM, Minhas GS, Sazanov LA (2013) Crystal structure of the entire respiratory complex I. *Nature* 494:443–448.
11. Di Luca A, Gamiz-Hernandez AP, Kaila VRI (2017) Symmetry-related proton transfer pathways in respiratory complex I. *Proc Natl Acad Sci USA* 114:E6314–E6321.
12. Sharma V, et al. (2015) Redox-induced activation of the proton pump in the respiratory complex I. *Proc Natl Acad Sci USA* 112:11571–11576.
13. Efremov RG, Sazanov LA (2012) The coupling mechanism of respiratory complex I–A structural and evolutionary perspective. *Biochim Biophys Acta* 1817:1785–1795.
14. Wikström M, Sharma V, Kaila VRI, Hosler JP, Hummer G (2015) New perspectives on proton pumping in cellular respiration. *Chem Rev* 115:2196–2221.
15. Tocilescu MA, et al. (2010) The role of a conserved tyrosine in the 49-kDa subunit of complex I for ubiquinone binding and reduction. *Biochim Biophys Acta* 1797:625–632.
16. Sinha PK, et al. (2015) Conserved amino acid residues of the NuoD segment important for structure and function of *Escherichia coli* NDH-1 (complex I). *Biochemistry* 54: 753–764.
17. Murai M, Miyoshi H (2016) Current topics on inhibitors of respiratory complex I. *Biochim Biophys Acta* 1857:884–891.
18. Angerer H, et al. (2012) Tracing the tail of ubiquinone in mitochondrial complex I. *Biochim Biophys Acta* 1817:1776–1784.
19. Sinha PK, et al. (2009) Critical roles of subunit NuoH (ND1) in the assembly of peripheral subunits with the membrane domain of *Escherichia coli* NDH-1. *J Biol Chem* 284:9814–9823.
20. Estornell E, et al. (1992) Saturation kinetics of coenzyme Q in NADH and succinate oxidation in beef heart mitochondria. *FEBS Lett* 311:107–109.
21. Fato R, et al. (1996) Steady-state kinetics of the reduction of coenzyme Q analogs by complex I (NADH:ubiquinone oxidoreductase) in bovine heart mitochondria and submitochondrial particles. *Biochemistry* 35:2705–2716.
22. Lenaz G, Castelli A, Littarru GP, Bertoli E, Folkers K (1971) Specificity of lipids and coenzyme Q in mitochondrial NADH and succin-oxidase of beef heart and *S. cerevisiae*. *Arch Biochem Biophys* 142:407–416.
23. Jones AJY, et al. (2016) A self-assembled respiratory chain that catalyzes NADH oxidation by ubiquinone-10 cycling between complex I and the alternative oxidase. *Angew Chem Int Ed Engl* 55:728–731.
24. Young L, et al. (2014) Probing the ubiquinol-binding site of recombinant *Sauromatum guttatum* alternative oxidase expressed in *E. coli* membranes through site-directed mutagenesis. *Biochim Biophys Acta* 1837:1219–1225.
25. Birrell JA, Yakovlev G, Hirst J (2009) Reactions of the flavin mononucleotide in complex I: A combined mechanism describes NADH oxidation coupled to the reduction of $APAD^+$, ferricyanide, or molecular oxygen. *Biochemistry* 48:12005–12013.
26. Shiba T, et al. (2013) Structure of the trypanosome cyanide-insensitive alternative oxidase. *Proc Natl Acad Sci USA* 110:4580–4585.
27. Nagle JF, Tristram-Nagle S (2000) Structure of lipid bilayers. *Biochim Biophys Acta* 1469:159–195.
28. Fato R, Battino M, Degli Esposti M, Parenti Castelli G, Lenaz G (1986) Determination of partition and lateral diffusion coefficients of ubiquinones by fluorescence quenching of *n*-(9-anthroyloxy)stearic acids in phospholipid vesicles and mitochondrial membranes. *Biochemistry* 25:3378–3390.
29. Battino M, Fahmy T, Lenaz G (1986) Determination of the critical micelle concentration of short-chain ubiquinones in model systems. *Biochim Biophys Acta* 851:377–384.
30. Afanasyeva EF, Syryamina VN, Dzuba SA (2017) Alamethicin can capture lipid-like molecules in the membrane. *J Chem Phys* 146:011103.
31. Wang KF, Nagarajan R, Camesano TA (2014) Antimicrobial peptide alamethicin insertion into lipid bilayer: A QCM-D exploration. *Colloids Surf B Biointerfaces* 116: 472–481.
32. Degli Esposti M, et al. (1996) The specificity of mitochondrial complex I for ubiquinones. *Biochem J* 313:327–334.
33. Chovancova E, et al. (2012) CAVER 3.0: A tool for the analysis of transport pathways in dynamic protein structures. *PLOS Comput Biol* 8:e1002708.
34. Armstrong CT, Mason PE, Anderson JLR, Dempsey CE (2016) Arginine side chain interactions and the role of arginine as a gating charge carrier in voltage sensitive ion channels. *Sci Rep* 6:21759.
35. Galassi VV, Arantes GM (2015) Partition, orientation and mobility of ubiquinones in a lipid bilayer. *Biochim Biophys Acta* 1847:1560–1573.
36. Genheden S, Ryde U (2015) The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin Drug Discov* 10:449–461.
37. Zhang L, Hermans J (1996) Hydrophilicity of cavities in proteins. *Proteins* 24:433–438.
38. Nihei C, et al. (2003) Purification of active recombinant trypanosome alternative oxidase. *FEBS Lett* 538:35–40.
39. Schmidt TGM, et al. (2013) Development of the Twin-Strep-tag® and its application for purification of recombinant proteins from cell culture supernatants. *Protein Expr Purif* 92:54–61.
40. Blaza JN, Serreli R, Jones AJY, Mohammed K, Hirst J (2014) Kinetic evidence against partitioning of the ubiquinone pool and the catalytic relevance of respiratory-chain supercomplexes. *Proc Natl Acad Sci USA* 111:15735–15740.
41. Kaplan RS, Pedersen PL (1985) Determination of microgram quantities of protein in the presence of milligram levels of lipid with amido black 10B. *Anal Biochem* 150: 97–104.
42. Smart OS, Neduvelil JG, Wang X, Wallace BA, Sansom MSP (1996) HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *J Mol Graph* 14:354–360, 376.
43. Phillips JC, et al. (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26: 1781–1802.
44. Best RB, et al. (2012) Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain $\chi(_1)$ and $\chi(_2)$ dihedral angles. *J Chem Theory Comput* 8:3257–3273.

# Article III

BBA - Bioenergetics 1859 (2018) 734–741

Contents lists available at ScienceDirect

BBA - Bioenergetics

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

# How inter-subunit contacts in the membrane domain of complex I affect proton transfer energetics

Andrea Di Luca[1], Max E. Mühlbauer[1], Patricia Saura, Ville R.I. Kaila*

*Department Chemie, Technische Universität München, Lichtenbergstr. 4, Garching, D-85747, Germany*

ARTICLE INFO

*Keywords:*
Bioenergetics
Proton transfer
NADH:ubiquinone oxidoreductase
Enzyme dynamics

ABSTRACT

The respiratory complex I is a redox-driven proton pump that employs the free energy released from quinone reduction to pump protons across its complete ca. 200 Å wide membrane domain. Despite recently resolved structures and molecular simulations, the exact mechanism for the proton transport process remains unclear. Here we combine large-scale molecular simulations with quantum chemical density functional theory (DFT) models to study how contacts between neighboring antiporter-like subunits in the membrane domain of complex I affect the proton transfer energetics. Our combined results suggest that opening of conserved Lys/Glu ion pairs within each antiporter-like subunit modulates the barrier for the lateral proton transfer reactions. Our work provides a mechanistic suggestion for key coupling effects in the long-range force propagation process of complex I.

## 1. Introduction

The respiratory complex I (NADH:ubiquinone oxidoreductase) is a redox-driven proton pump that serves as an initial electron entry point in prokaryotic and eukaryotic respiratory chains [1–4]. By reducing quinone (Q) to quinol ($QH_2$), complex I transports four protons across a biological membrane [1,5,6] and establishes a proton motive force (*pmf*) that is employed for active transport and synthesis of adenosine triphosphate (ATP) [7,8]. Complex I is by far the largest and most intricate member of the respiratory chain, and despite recently resolved structures [9–12], the molecular mechanism by which it pumps protons still remains unclear.

Complex I is a 0.5–1 MDa L-shaped enzyme that comprises up to 45 subunits in eukaryotes [13], and is organized into a hydrophilic domain and a membrane domain. The electron transfer process takes places in the ca. 100 Å long hydrophilic domain, whereas the 200 Å long membrane domain is responsible for the proton pumping function (Fig. 1). The 14 conserved core subunits of the enzyme constitute the machinery needed to catalyze the long-range proton-coupled electron transfer (PCET) process. The remaining 31 supernumerary subunits in the eukaryotic enzyme are organized around the core subunits [14–16] and are possibly involved in the regulation of enzyme functions, e.g., by the *active*-to-*deactive* transition [11,12,17–20], which modulates complex I activity.

The electron transport process is initiated by the oxidation of nicotinamide adenine dinucleotide (NADH), which transfers its two electrons via a non-covalently bound flavin mononucleotide (FMN) cofactor to a chain of 8–9 iron sulfur centers (ISC) and further to the Q-binding site, located ca. 30 Å above the membrane surface (Fig. 1) [9,21,22]. The electron transfer (eT) between NADH and the terminal N2 iron-sulfur center takes place on ca. 90 μs timescales [23,24], which is fast relative to the millisecond turnover of complex I [2], and thus not rate-limiting for the proton pumping process [2,25].

The Q pocket is located at the interface between the three subunits Nqo4, Nqo6, and Nqo8 (*T. thermophilus* nomenclature), and extends ca. 40 Å toward the membrane domain. There are to date no experimentally resolved structures of complex I with bound Q, but computational studies [26,27] suggest that $Tyr_4$-87 and $His_4$-38 stabilize the Q headgroup and function as local proton donors in the Q-reduction process. These findings are also supported by site-directed mutagenesis studies [28,29]. The Q cavity has a non-uniform polarity, and it comprises a kink region with many polar and charged amino acids [9,30]. From this kink, a chain of conserved charged/polar residues extends in the middle of the membrane domain toward the terminal Nqo12 subunit [9,31].

Of the seven membrane domain subunits, three antiporter-like subunits, Nqo12, Nqo13, and Nqo14 are evolutionary related to each other and to multi-resistance and pH adaptation (Mrp) $Na^+/H^+$ antiporters. The antiporter-like subunits share an internal pseudo-

* Corresponding author.
*E-mail address:* ville.kaila@ch.tum.de (V.R.I. Kaila).
[1] Contributed equally to this work.

https://doi.org/10.1016/j.bbabio.2018.06.001
Received 1 February 2018; Received in revised form 8 May 2018; Accepted 2 June 2018
Available online 05 June 2018

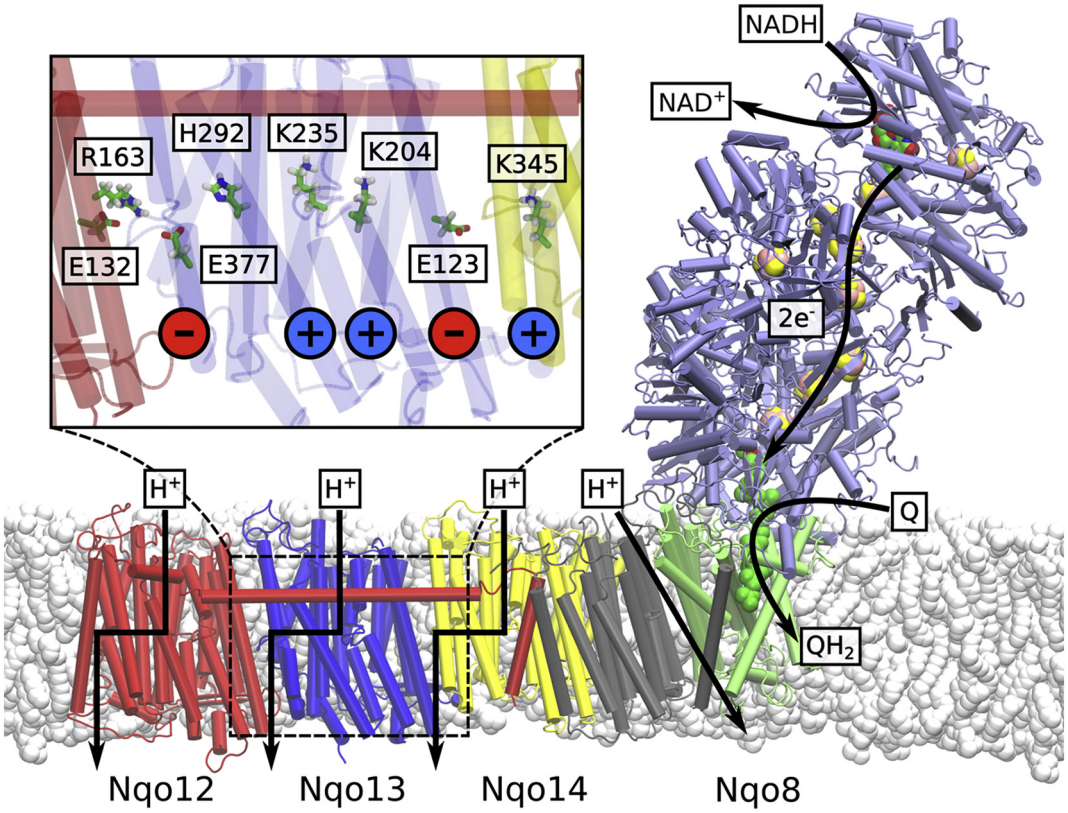


**Fig. 1.** Structure of the bacterial complex I from *T. thermophilus* (PDB ID: 4HEA). Electron transfer takes place in the hydrophilic domain (in purple) between NADH and Q. Energy released from the Q reduction process is employed to transfer four protons across the membrane. The antiporter-like subunits Nqo12 (in red), Nqo13 (in blue) and Nqo14 (in yellow) are likely to transfer one proton each, whereas the location of the fourth proton pathway is still not fully clear (but cf. [33,34]). *Inset*: Chain of conserved charged and polar residues in Nqo13. From right to left: terminal charged residue of Nqo14, the Glu/Lys ion pair, the central Lys, the bridging His, the terminal charged residue of Nqo13 (Glu), and the interface to Nqo12.

symmetry, with two trans-membrane (TM) helix bundles, TM4-8 and TM9-13, that contain a broken-helix element. A similar five-helical bundle, TM2-6, is also present in the Nqo8 subunits, and it comprises a part of the Q channel. The proton pumping is likely to occur in the Nqo12, Nqo13, and Nqo14 subunits, pumping one proton each. The location of the fourth proton pathway is still under debate, but a possible location is the region between Nqo8 and Nqo11 [9,10]. Recent simulations [32–34] show that the proton channels are likely to form at the broken-helix segment, similar as in other transporters [35,36].

It is possible to identify conserved repeated residue motifs in each antiporter-like subunit. These include a Lys/Glu ion pair (Arg/Glu in Nqo12), a central Lys residue, one or more bridging His residues, and a terminal charged Lys or Glu residue (Fig. 1, *inset*). Site-directed mutagenesis experiments [37–42] and molecular simulations suggest that these residues are crucial for the proton pumping activity. Moreover, the Q reduction activity is also affected by mutations of these residues, suggesting that the electron and proton transfer processes are tightly coupled in complex I. Importantly, to achieve such tight coupling between the Q reaction and the terminal proton transfer in Nqo12, the "Q-reduction signal" needs to propagate through the complete membrane domain.

Biochemical, structural, and computational studies have probed possible proton pumping mechanisms ([25,33,42–46], cf. [44] and refs. therein). Although the overall pumping process takes place on milliseconds timescales, individual transitions that couple to the pumping process may take place on much shorter timescales once a rate-limiting step has been overcome. Therefore, relaxation of such "non-equilibrium" state created here, e.g., by protonation changes, can be employed to obtain mechanistic information of rare events using molecular dynamics simulations that are shorter than the overall turnover timescale.

Recently, we suggested a molecular mechanism where conformational changes in the Lys/Glu ion pairs are involved in the long-range force propagation process and transmit the signal between neighboring antiporter-like subunits. This mechanism involves sequential Lys/Glu ion-pair dissociation and lateral proton transfer processes, propagating the signal from Nqo14 to Nqo12. We found that the energetics of ion-pair dissociations depends on the protonation state of the central Lys residues, making also the reverse effect possible, i.e., that the ion-pair dynamics modulate the p$K_a$ of the neighboring amino acids. The protonic connectivity to the two membrane sides (N- and P-side) was further suggested to be regulated by the hydration state of *input* and *output* channels, which in turn is controlled by the state of buried charged residues. The water channels "*open*" and "*close*" on the sub-μs timescale, which may provide a rate-limiting element in the proton pumping process. It was suggested that the proton N-side input and P-side output channels are located at symmetry-related positions. More specifically, the channels form at the interface between subunits and the 5-TM helices bundles, sharing the same symmetry and connecting the buried central Lys and terminal charged residues within each subunit to the N- and P-sides of the membrane [33]. The lateral proton transfer takes place between the central Lys residue and the charged residue facing the subsequent subunit. Each of these events triggers the following process and propagates across the complete ca. 200 Å membrane domain. We also suggested based on thermodynamic considerations [46] that the proton pumping process involves a "*backwave*" that couples to proton release across the membrane.

To study the energetics and dynamics of this coupling principle, we perform here classical molecular dynamics (MD) simulations in combination with quantum chemical density functional theory (DFT)

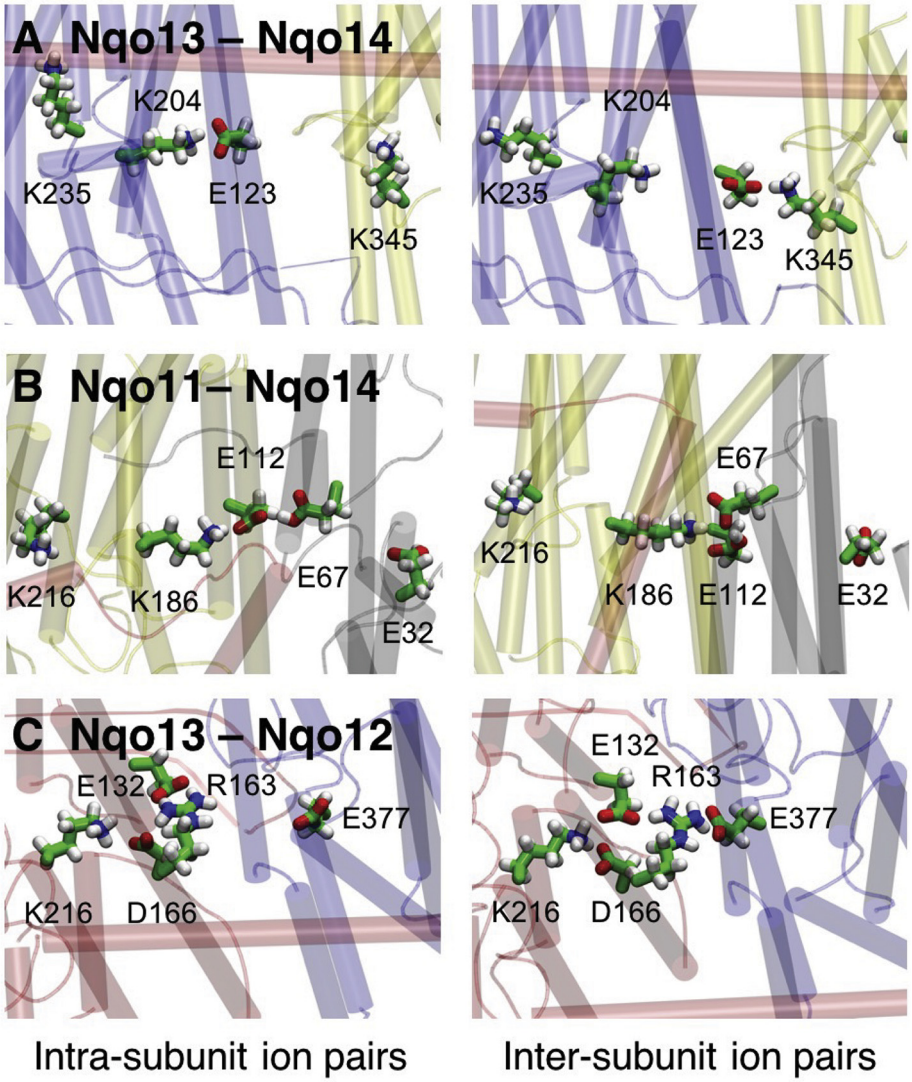


**Fig. 2.** Structure of the antiporter-like subunit interfaces in the membrane domain of complex I. The figure shows A) the Nqo13-Nqo14 interface, B) the Nqo11-Nqo14 interface, and C) the Nqo13-Nqo12 interface. Ion pairs between conserved charged residues form alternating contacts within the same and neighboring subunits, with snapshots of *intra*-subunit contacts (to the left) and *inter*-subunit contacts (to the right).

calculations on the experimentally resolved X-ray structure of complex I from *Thermus thermophilus*. Our data suggest how the *inter*-subunit contacts are established and how these interactions could modulate the proton transfer energetics.

## 2. Models and methods

### 2.1. Classical molecular dynamics

The X-ray structure of *T. thermophilus* complex I [9] was embedded in a POPC membrane and solvated with TIP3P water, and the system was neutralized with a ca. 100 mM NaCl concentration. Ubiquinone ($Q_{10}$) was modeled in the Q-cavity, which was identified using the HOLE [47] software, and the Q headgroup was placed between $His_4$-38 and $Tyr_4$-87 of the Nqo4 subunit. The system comprised ca. 830,000 atoms. A constant temperature of 310 K and pressure of 1 bar were modeled in an *NPT* ensemble, and long-range electrostatics were treated by the Particle Mesh Ewald (PME) method [48]. The simulations were performed using NAMD2 [49] and the CHARMM27 force field [50,51] using a 2 fs integration timestep. Force field parameters for the cofactors were derived from density functional theory (DFT/B3LYP/def2-TZVP) calculations. p$K_a$ values were estimated using Poisson-Boltzmann (PB) continuum electrostatic calculations and performed using the Adaptive PB solver (APBS) [52], by performing the Monte Carlo sampling of the $2^N$ possible protonation states with Karlsberg+ [53]. The system was described by explicit partial atomic charges embedded in an inhomogeneous medium with an $\varepsilon = 4$, and bulk water by a homogeneous medium with $\varepsilon = 80$. Part of the simulation data were also employed in Ref. [33] and are reported in Table S1. Principal component analysis (PCA) [54,55] of the MD data was performed using the position of the Cα atoms of subunits during 200–600 ns of dynamics (simulations 2 and 4) after 400 ns of simulation. The PCA and related analyses were performed with ProDy [56].

### 2.2. Quantum chemical density functional theory models

Quantum chemical DFT models consisting of the $Lys_{13}$-235, $His_{13}$-292 and $Glu_{13}$-377 residues of Nqo13 and $Lys_{14}$-216, $His_{14}$-265 and $Lys_{14}$-345 residues of Nqo14, with three intervening water molecules in each case, were constructed based on 100 ns relaxed MD simulations. The amino acid residues were cut at the Cβ (for His and Glu) or Cδ (for Lys) positions, which were fixed during structure optimization at the B3LYP-D3/def2-SVP level [57–60]. The protein environment was treated as a polarizable medium with $\varepsilon = 4$ using the conductor-like

screening model (COSMO) [61]. Transition states were also optimized at the same level of theory. Electronic energies were computed at B3LYP-D3/def2-TZVP/ε = 4 level with zero-point vibrational (ZPE) energy corrections obtained at B3LYP-D3/def2-SVP/ε = 4 level. To study the effect of the $Lys_{13}$-204/$Lys_{14}$-186, we added a Lys α-amino group at 10.5 Å from the $Lys_{13}$-235/$Lys_{14}$-216. All calculations were performed with TURBOMOLE v 6.6 [62].

## 3. Results and discussion

### 3.1. Inter-subunit contacts affect intra-subunit residue conformations

Starting from the crystal structure of complex I from *Thermus thermophilus*, we performed ca. 3 μs classical molecular dynamics (MD) simulations with the central polar residues modeled in both their protonated and deprotonated forms (Table S1). In the MD simulations, we find that the Lys/Arg-Glu ion pairs in each antiporter-like subunit form transient contacts with neighboring subunits that are stabilized by interactions with oppositely charged residues at their interface (Fig. 2). For each of the three interfaces Nqo11/Nqo14, Nqo13/Nqo14, and Nqo12/Nqo13, we observe a qualitatively similar behavior, with the key residues showing a two-state conformational switching behavior, which could be important for the signal propagation in the membrane domain of complex I.

We find that the Nqo13/Nqo14 interface has the clearest switching behavior (Fig. 3A). The MD simulations suggest that $Glu_{13}$-123 can form both an *intra*-subunit salt-bridge with $Lys_{13}$-204 and an *inter*-subunit contact with $Lys_{14}$-345 upon conformational switching (Fig. 2A, Fig. 3A). This switching is coupled with a decrease in the distance between $Lys_{13}$-204 and the central $Lys_{13}$-235 (Fig. 3A, *upper panel*), which could function as a primary proton donor in the pumping process [33], as also supported by site-directed mutagenesis experiments [41]. The distance distribution for the $Lys_{13}$-235/$Lys_{13}$-204 pair shows two major sidechain conformations (Fig. 2A). When either of the residues is modeled in the deprotonated state, we obtain a mean distance of ca. 10.5 Å, whereas when both residues are in their protonated (charged) states, the electrostatic repulsion increases their mean distance to ca. 14.0 Å (Fig. 3A, *lower panel*), showing that the ion-pair dynamics is tightly coupled to the protonation state of the residues (Fig. 3A). Interestingly, in the crystal structure of complex I from *Thermus thermophilus* (PDB ID: 4HEA), the $Lys_{13}$-204/$Glu_{13}$-123 ion pair has been refined in the dissociated state, with $Glu_{13}$-123 flipped toward the Nqo14 subunit.

Similar as for Nqo13/Nqo14, we also observe a conformational switching at the interface between Nqo14 and Nqo11 (Fig. 3B). However, in contrast to the Nqo13/Nqo14 interface where $Glu_{13}$-123 interacts with two oppositely charged residues, $Glu_{14}$-112 is surrounded by $Lys_{14}$-186 and two acidic residues, $Glu_{11}$-67 and $Glu_{11}$-32 (Fig. 2B). Here we observe two distinct conformations of $Glu_{11}$-67, which result in a ca. 7 Å distance to $Lys_{14}$-186 when it is modeled in a protonated state, and a ca. 3.2 Å distance when is modeled in a deprotonated state (Fig. 3B), forming a hydrogen-bonded contact. When deprotonated, $Glu_{11}$-67 faces away from $Glu_{11}$-32, possibly due to electrostatic repulsion. Similar as in Nqo13, this conformational change correlates with an increase in the $Lys_{14}$-186/$Lys_{14}$-216 distance (Fig. 3B). $Lys_{14}$-216 is likely to function as the proton donor in subunit Nqo14 [33,44].

We next analyzed the ion-pair dynamics at the Nqo12/Nqo13 interface, which is structurally different from the other antiporter-like subunits. The Nqo12/Nqo13 interface comprises two positively charged residues, $Lys_{12}$-216 and $Arg_{12}$-163, which are compensated by three negatively charged residues, $Glu_{13}$-377, $Glu_{12}$-132, and $Asp_{12}$-166 (Fig. 2C). $Arg_{12}$-163 has been suggested to replace a putative $Na^+$-binding site in their evolutionary ancestral $Na^+$-pumping Mrp transporters [4]. Our MD simulations suggests that $Arg_{12}$-163 can form a salt-bridge with the surrounding acidic residues. In addition to the two distinct conformational states observed for the other interfaces, we also observe a third intermediate state (Fig. 3C, *lower panel*), which could arise from simultaneous interaction with both its acidic neighbors at the same time. We find that the strong interaction between $Arg_{12}$-163 and $Asp_{12}$-166 is anti-correlated with the opening of the $Glu_{13}$-377/$Arg_{12}$-163 ion pair (Fig. 3C, *upper panel*).

In Nqo12, the distance between the putative proton donor, $Lys_{12}$-329, and residues $Lys_{12}$-216/$Asp_{12}$-166 at the *inter*-subunit interface is larger than 18 Å, making a direct electrostatic coupling between the two sites somewhat weaker, as compared to the interaction in the other subunits. However, the conserved $His_{12}$-241 of TM8, located in the same position as the central Lys in Nqo13 and Nqo14, could provide a link necessary to couple the ion-pair dynamics to the proton transfer process. The higher complexity in Nqo12 might be related to the fact that it is the terminal antiporter-like subunit, and the coupling might be weaker than for the other subunits [4] (however, cf. also [46]).

### 3.2. Channel hydration and coupling between subunits

We recently observed a connection between the protonation state of the central Lys residues and the opening/closing dynamics of the proton channels in the antiporter-like subunits [33]. Our MD simulations suggest that in Nqo13, the water connectivity next to the broken helix TM7a is established when $Lys_{13}$-235 is protonated, and is lost upon its deprotonation (Fig. 4A). The hydration state of the antiporter-like subunit is, interestingly, also coupled with subtle structural changes (Fig. 4A). We observe that upon deprotonation of the middle Lys, the hydration level of the channel next to the broken helix TM7a drops significantly (Fig. 4B). Only one water molecule, which interacts with the deprotonated Lys, remains close to the TM4-8 helix bundle, whereas the remaining channel water is pushed toward the N-side by $His_{13}$-211 and $Leu_{13}$-214, which move closer together to form a gating element. Both residues are located on the upper part of the broken helix (TM7) of the antiporter-like subunit. Communication between the two parts of the helix (TM7a and TM7b) could be mediated by $Trp_{13}$-213 in TM7a, as indicated by a hydrogen-bond between the tryptophan and the backbone of $Leu_{13}$-203 of the lower helix (TM7b). The tilt of TM7a relative to TM7b also changes significantly after deprotonation of $Lys_{13}$-235 (Fig. 4B), which also couples to a subtle shift in the π-kink of TM8 with $Lys_{13}$–235. These conformational changes lead to a decrease in the channel radius by ca. 2 Å measured from the distance between $Leu_{13}$-214 on TM7a and $Lys_{13}$-287 on TM8, and a decrease in mean channel hydration by ca. 50% (within 4 Å of the gate, Fig. 4B).

The MD simulations performed with different protonation states reveal possible effects driving the opening/closure of channels. To further probe global coupling effects, we performed a principal component analysis (PCA) that projects out global slow relaxing degrees of freedom. The dynamical correlation between the antiporter-like subunits, calculated based on the PCA correlation matrix on different trajectories, shows how the coupling of motions could depend on both the hydration level of the subunits and the protonation states of the buried residues. When the putative proton channels are hydrated, we observe a strong coupling between the subunits (Fig. 5A). However, upon dehydration of the water channels by deprotonation of the central Lys, the coupling between subunits weakens as indicated by a reduced correlation (Fig. 5B).

Although analysis of more intermediate states is needed to clarify details of the coupling between subunits, our data nevertheless indicate that the protonation state of conserved residues and channel hydration affect complex I dynamics. This suggests that not only the protein structure, but also the water molecules play an active role in the pumping process, by providing essential coupling elements that transmit the signal in addition to their role as “proton wires”.

### 3.3. Inter-subunit contacts modulate proton transfer energetics

To “push” the proton horizontally within the antiporter-like

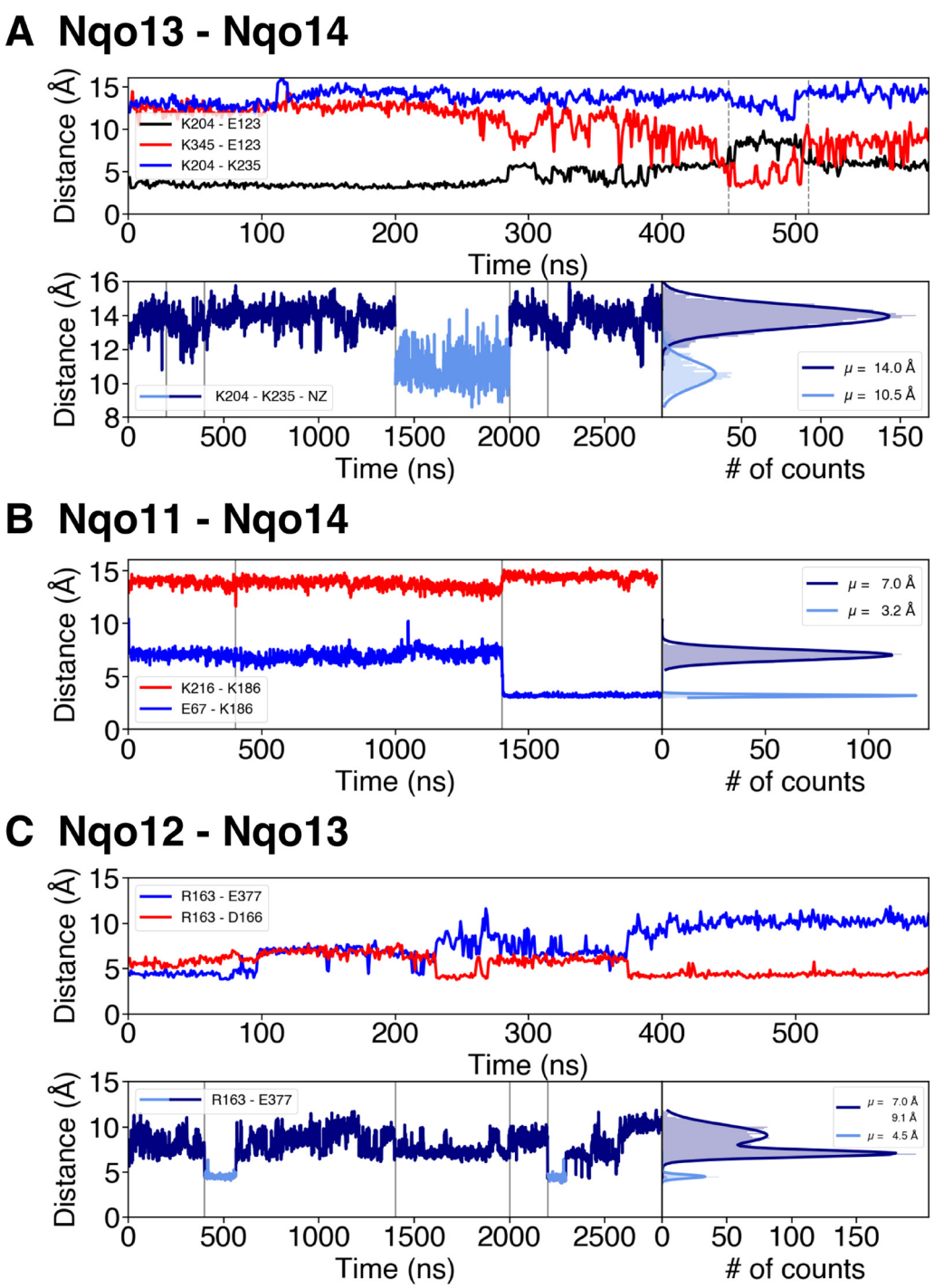


**Fig. 3.** Ion-pair dynamics at the subunit interfaces in the membrane domain of complex I. A) The Nqo13 – Nqo14 interface. *Upper panel*: Distances of the $Lys_{13}$-204/$Glu_{13}$-123 ion pair (in black), the $Lys_{14}$-345/$Glu_{13}$-123 ion pair (in red) during 0.6 μs of MD simulation (simulation 5). The flip of the ion pair correlates with a decrease in the $Lys_{13}$-204/$Lys_{13}$-235 distance (in blue). *Lower panel*: $Lys_{13}$-235/$Lys_{13}$-204 distances and their distribution during 2.8 μs of dynamics (simulations 1 to 5). The dark blue curve represents a '*repulsive*' regime where both residues are positively charged, whereas deprotonation of $Lys_{13}$-204 leads to sampling of shorter distances (light blue curve). Vertical grey lines indicate boundaries between separate simulations. B) The Nqo11 – Nqo14 interface. Distances of the $Glu_{11}$-67/$Lys_{14}$-186 (in blue) and $Lys_{14}$-186/$Lys_{14}$-216 (in red) ion pairs during 2 μs MD simulations (simulations 1, 2, 5) and the distribution of the distances between $Glu_{11}$-67/$Lys_{14}$-186 during the MD simulations. Upon deprotonation of $Glu_{11}$-67 at 1.4 μs, the inter-subunit contact forms, which correlates with an increase in the $Lys_{14}$-186/$Lys_{14}$-216 distance. Deprotonation of $Glu_{11}$-67 leads to formation of contact with $Lys_{14}$-186. C) The Nqo12 – Nqo13 interface. *Upper panel*: Distances of the $Arg_{12}$-163/$Glu_{13}$-377 (in blue) and the $Arg_{12}$-163/$Asp_{12}$-166 (in red) ion pairs during 0.6 μs of MD simulation (simulation 5). $Arg_{12}$-163 forms *inter-* and *intra*-subunit salt-bridges. *Lower panel*: Distances between $Arg_{12}$-163/$Glu_{13}$-377 during 2.8 μs of simulation (simulations 1 to 5) and their respective distribution. *Intra*-subunit and intermediate states are drawn in dark-blue and the *inter*-subunit state in light blue. Vertical grey lines indicate boundaries between separate simulations.

domains, complex I must invest energy by destabilizing the protonated middle Lys. As described above, this could be achieved by opening of the Lys/Glu ion pair, which is expected to result in a charge repulsion between $Lys_{13}$-204 and $Lys_{13}$-235 (in Nqo13), as also suggested by recent free energy simulations [33]. To qualitatively probe such coupling principles, we built quantum chemical model systems comprising the sidechains of residues Lys-His-Lys and Lys-His-Glu (proton donor – bridging residue – proton acceptor), bridging water molecules, and using a protonated Lys sidechain as "*triggering signal*" for the proton transfer process (Fig. 6). The structures were extracted from the classical MD simulations of complex I after ca. 100 ns. After DFT optimization of intermediate and transitions states, we studied the energetics of the proton transfer reaction for this model system by probing the effect of a positive charge next to the proton donor, mimicking the charge of the unpaired Lys resulting from the flip of the interface salt-bridges to the neighboring subunit.

In subunit Nqo13 (Fig. 6A), our quantum chemical models suggest that proton transfer from $Lys_{13}$-235 to $His_{13}$-292 via one bridging water molecule (A1 to A3 in Fig. 6A) has an energy barrier of ca. 4 kcal $mol^{-1}$ and takes place via a hydronium-like transition state structure. The process is exergonic by ca. 5 kcal $mol^{-1}$. The effect of the uncompensated $Lys_{13}$-204 lowers this barrier by ca. 2 kcal $mol^{-1}$ and renders the reaction more exergonic by ca. 3 kcal $mol^{-1}$. The subsequent proton transfer from $His_{13}$-292 to $Glu_{13}$-377 via two water molecules (A3 to A5 in Fig. 6A) has a barrier of ca. 6 kcal $mol^{-1}$, with a transition state resembling a Zundel ion, and a reaction energy of ca.

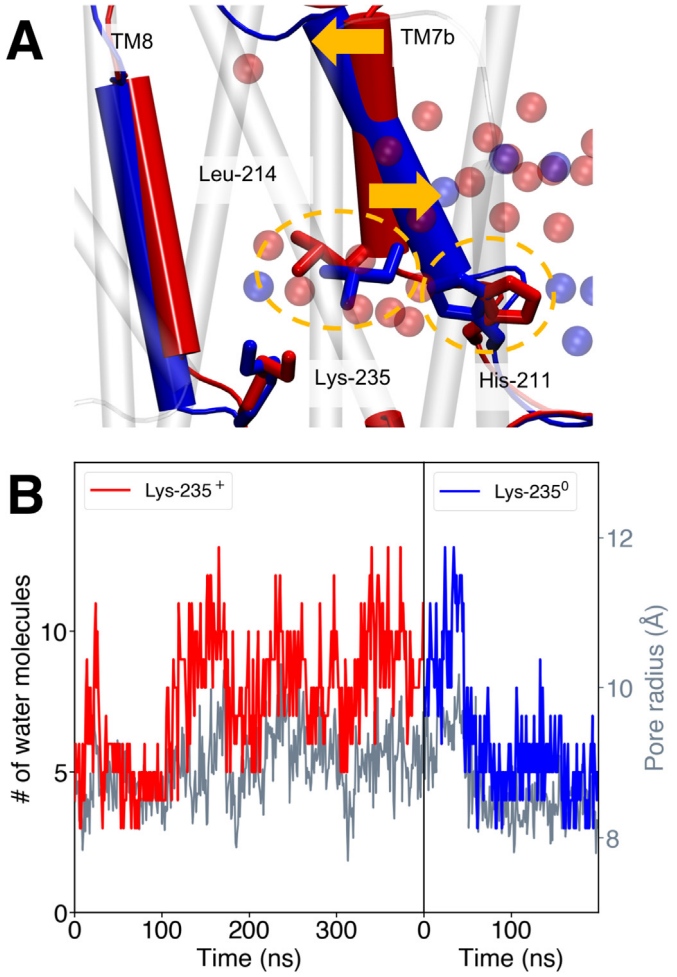


**Fig. 4.** A) Conformational changes associated with deprotonation of the central $Lys_{13}$-235 in subunit Nqo13 in the membrane domain of complex I. When $Lys_{13}$-235 is simulated in a protonated state (red, simulation 2) the gate formed by the sidechains of $Leu_{13}$-214 and $His_{13}$-211 (yellow circles) is open, whereas the gate closes when $Lys_{13}$-235 is modeled in a deprotonated state (blue, simulation 4), inducing dehydration of the channel. The opening/closing is coupled to a structural tilting of the TM7b broken helix (yellow arrows). B) The channel hydration state (in red and blue) and channel radius (in grey) for subunit Nqo13 with $Lys_{13}$-235 modeled the protonated (simulation 2, red) and deprotonated states (simulation 4, blue), respectively. The hydration state is measured as the number of water molecules within 4 Å around the $Leu_{13}$-214/$His_{13}$-211 gate. The pore radius is measured as distance between $Leu_{13}$-214 (TM7a) and $Lys_{13}$-287 (TM8). The mean hydration changes from four water molecules to eight water molecules during channel opening.

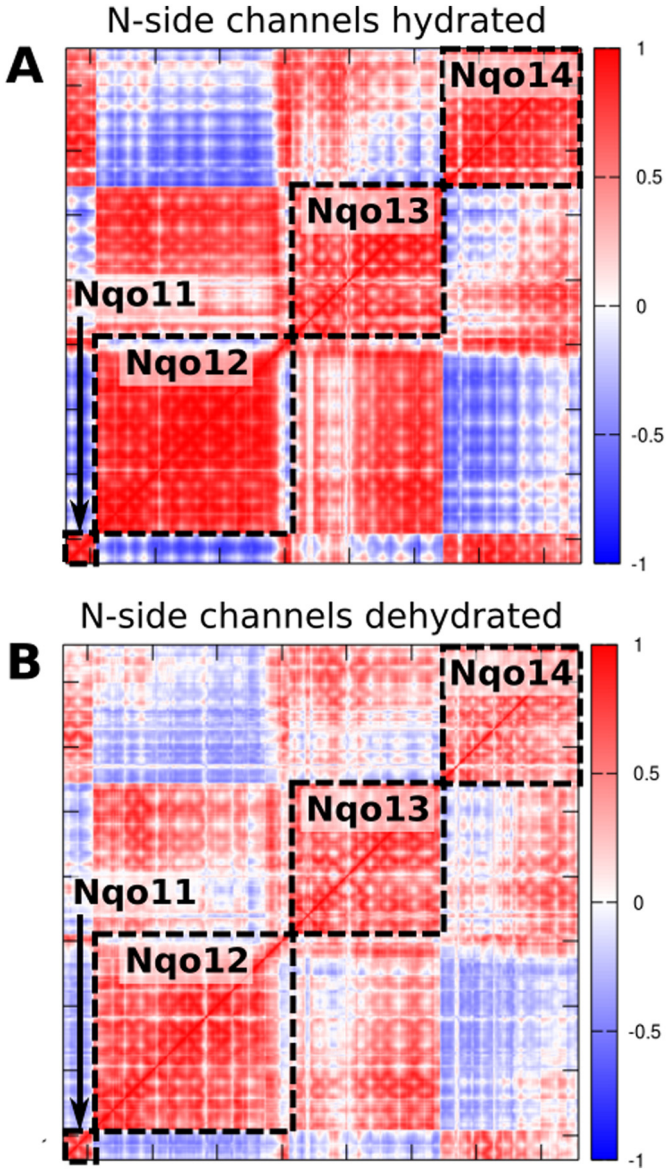


**Fig. 5.** Correlation matrices of the membrane domain of complex I obtained from principal component (PC) analysis of the MD simulation data (simulations 2 and 4, Table S1). Correlation of PCs A) with open channels between the N-side bulk water and central Lys residues, B) upon deprotonation of the central Lys residues and closure of the N-side channels. Neutralization and/or dehydration of the subunits cause a decreased coupling between subunits. The colors refer to the inter-residue correlation from 1 (directly correlated), to −1 (anti-correlated).

$-14\ kcal\ mol^{-1}$. The exergonicity of this process is likely to be overestimated due to the neutralization of an uncompensated negative Glu in a low dielectric environment. The presence of the unpaired Lys sidechain at 10.5 Å from $Lys_{13}$-235 stabilizes both the intermediate and final states (A3 to A5 in Fig. 6A) by ca. $3\ kcal\ mol^{-1}$, conserving the barrier of ca. $5\ kcal\ mol^{-1}$. In subunit Nqo14 (Fig. 6B), the energetics of the proton transfer from $Lys_{14}$-216 to $His_{14}$-292 (B1 to B3 in Fig. 6B) is similar as in the Nqo13 subunit model, with a barrier of ca. $5\ kcal\ mol^{-1}$ and an exergonicity of ca. $5\ kcal\ mol^{-1}$. We find that the charge of $Lys_{14}$-186 could stabilize this hydronium-ion transition state by ca. $2\ kcal\ mol^{-1}$, but it slightly destabilizes the protonated $His_{14}$-262 by ca. $1\ kcal\ mol^{-1}$. Proton transfer from $His_{14}$-262 to $Lys_{14}$-345 via two water molecules (B3 to B5 in Fig. 6B) has a relatively high energy barrier of ca. $15\ kcal\ mol^{-1}$ in the model system, and it is endergonic by ca. $7\ kcal\ mol^{-1}$, yielding a final state which is ca. $2\ kcal\ mol^{-1}$ higher in energy than the initial state. However, the charge of the uncompensated Lys lowers the barrier by ca. $4\ kcal\ mol^{-1}$ and makes the reaction more exergonic by ca. $4\ kcal\ mol^{-1}$, now being stabilized relative to the initial state by ca. $2\ kcal\ mol^{-1}$. The intermediates and transition states observed in our QM cluster models are consistent with the ones from QM/MM-models probed in our previous study [33].

Although the proton transfer model employed here is simple, it nevertheless captures qualitative features on how modulation of the proton transfer energetics could be achieved in complex I. The model also describes some of the essential features needed to create a tightly coupled proton pumping machinery. While here the cost to change the energetics of the system is given by the energy required to "create" a positive charge next to our proton transfer chain, the same process in the complete system would also include the free energy caused by the separation of the ion pair and creation of the "excess charge". Free energy computations [33] estimate these effects as 3–4 $kcal\ mol^{-1}$ with

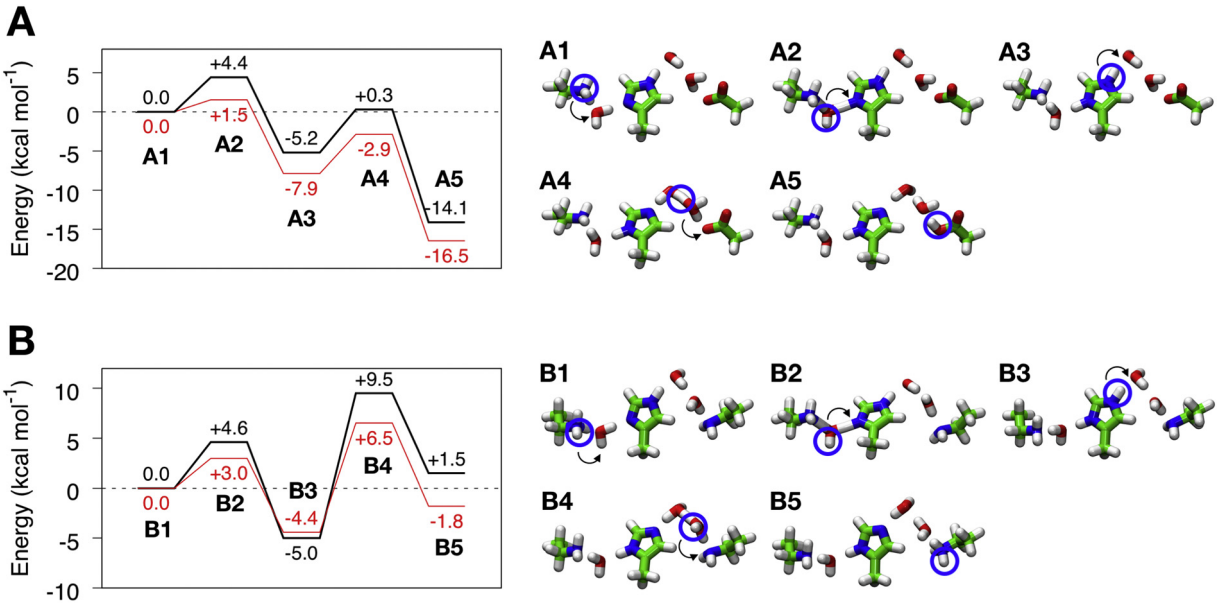


**Fig. 6.** Effect of dissociation of the "Glu-Lys" ion pair on proton transfer energetics in minimal quantum chemical models of the proton pathway in A) Nqo13 and B) Nqo14. Energy profiles (in kcal $mol^{-1}$) are shown to the left and the corresponding structures of optimized intermediates and transition states are drawn on the right. Profiles in red show the effect of adding a polarizing $Lys^{+}$, that forms upon dissociation of the Lys/Glu ion pair. The energetics were estimated at B3LYP/def2-TZVP/ $\varepsilon = 4$ level with zero-point energy (ZPE) corrections calculated at the B3LYP/def2-SVP level.

a deprotonated central Lys and > 8 kcal $mol^{-1}$ when $Lys_{13}$–235 is modeled in a protonated state. The models thus support that the inter-subunit ion pairs could influence the lateral proton transfer reaction and coupling between subunits propagated.

## 4. Conclusions

Our molecular simulations presented here provide insight on how *inter*-subunit contacts can modulate the proton transfer processes in the antiporter-like subunits of complex I. The proton pumping in the membrane domain of complex I, involves proton transfer reactions between the N-side bulk and buried titratable lysine residues, as well as horizontally across the membrane domain. Our simulations indicate that the energetics and kinetics of these proton transfer reactions are coupled to the charged state of a conserved ion pair within each antiporter-like subunit, and in turn, regulated by the charged state of the neighboring subunit. Our data indicate that $Glu_{13}$-123 in Nqo13, could act as a two-state conformational switch, by interaction with $Lys_{14}$-345 of Nqo14. The resulting uncompensated charge of $Lys_{13}$-204 could therefore lead to proton transfer to $Glu_{13}$-377. Quantum chemical model calculations of these processes further support that the conformational state of the Lys/Glu ion pair indeed strongly modulates the proton transfer energetics. Moreover, our principal component analysis of global dynamics in the membrane domain of complex I suggests that the ion-pair dynamics is also linked with channel hydration and *inter*-subunit couplings. We observe similar effects as described for Nqo13, also in the other antiporter-like subunits. Our combined results suggest that a combination of conformational and electrostatic switching provide an important functional principle to achieve an *action-at-a-distance* effect in complex I.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbabio.2018.06.001.

## Transparency document

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

## Acknowledgements

This work received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program/grant agreement 715311. The Leibniz-Rechenzentrum (LRZ), SuperMuc (projects: pr48de and pr27xu) provided computational resources.

## References

[1] M. Wikström, Two protons are pumped from the mitochondrial matrix per electron transferred between NADH and ubiquinone, FEBS Lett. 169 (2) (1984) 300–304.
[2] J. Hirst, Mitochondrial complex I, Annu. Rev. Biochem. 82 (1) (2013) 551–575.
[3] U. Brandt, Energy converting NADH: ubiquinone oxidoreductase (complex I), Annu. Rev. Biochem. 75 (1) (2006) 69–92.
[4] L.A. Sazanov, A giant molecular proton pump: structure and mechanism of respiratory complex I, Nat. Rev. Mol. Cell Biol. 16 (6) (2015) 375–388.
[5] A.J.Y. Jones, J.N. Blaza, F. Varghese, J. Hirst, Respiratory complex I in *Bos taurus* and *Paracoccus denitrificans* pumps four protons across the membrane for every NADH oxidised, J. Biol. Chem. 292 (12) (2017) 4987–4995.
[6] A.S. Galkin, V.G. Grivennikova, A.D. Vinogradov, $H^{+}/2e^{-}$ stoichiometry in NADH-quinone reductase reactions catalyzed by bovine heart submitochondrial particles, FEBS Lett. 451 (2) (1999) 157–161.
[7] P. Mitchell, Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism, Nature 191 (4784) (1961) 144–148.
[8] M. Yoshida, E. Muneyoki, T. Hisabori, ATP synthase — a marvelous rotary engine of the cell, Nat. Rev. Mol. Cell Biol. 2 (9) (2001) 669–677.
[9] R. Baradaran, J.M. Berrisford, G.S. Minhas, L.A. Sazanov, Crystal structure of the entire respiratory complex I, Nature 494 (7438) (2013) 443–448.
[10] 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 (6217) (2015) 44–49.
[11] J. Zhu, K.R. Vinothkumar, J. Hirst, Structure of mammalian respiratory complex I, Nature 536 (7616) (2016) 354–358.
[12] 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 (7625) (2016) 406–410.
[13] J. Carroll, I.M. Fearnley, J.M. Skehel, R.J. Shannon, J. Hirst, J.E. Walker, Bovine complex I is a complex of 45 different subunits, J. Biol. Chem. 281 (43) (2006) 32724–32727.
[14] K.R. Vinothkumar, J. Zhu, J. Hirst, Architecture of mammalian respiratory complex I, Nature 515 (7525) (2014) 80–84.
[15] C. Wirth, U. Brandt, C. Hunte, V. Zickermann, Structure and function of mitochondrial complex I, Biochim. Biophys. Acta 1857 (7) (2016) 902–914.
[16] K. Kmita, V. Zickermann, Accessory subunits of mitochondrial complex I, Biochem. Soc. Trans. 41 (5) (2013) 1272–1279.
[17] E. Maklashina, A.B. Kotlyar, G. Cecchini, Active/de-active transition of respiratory

complex I in bacteria, fungi and animals, Biochim. Biophys. Acta 1606 (1) (2003) 95–103.
[18] A.D. Vinogradov, Catalytic properties of the mitochondrial NADH-ubiquinone oxidoreductase (complex I) and the pseudo-reversible active/inactive enzyme transition, Biochem Biophys Acta 1364 (2) (1998) 169–185.
[19] S. Dröse, A. Stepanova, A. Galkin, Ischemic A/D transition of mitochondrial complex I and its role in ROS generation, Biochim. Biophys. Acta 1857 (7) (2016) 946–957.
[20] A. Di Luca, V.R.I. Kaila, Global collective motions in the mammalian and bacterial respiratory complex I, Biochim. Biophys. Acta 1859 (5) (2018) 326–332.
[21] C. Hunte, V. Zickermann, U. Brandt, Functional modules and structural basis of conformational coupling in mitochondrial complex I, Science 329 (5990) (2010) 448–451.
[22] L.A. Sazanov, P. Hinchliffe, Structure of the hydrophilic domain of respiratory complex I from *Thermus thermophilus*, Science 311 (5766) (2006) 1430–1536.
[23] M.L. Verkhovskaya, N. Belevich, L. Euro, M. Wikström, Real-time electron transfer in respiratory complex I, Proc. Natl. Acad. Sci. U. S. A. 105 (10) (2008) 3763–3767.
[24] S. De Vries, K. Dörner, M.J.F. Strampraad, T. Friedrich, Electron tunneling rates in respiratory complex I are tuned for efficient energy conversion, Angew. Chem. Int. Ed. 54 (2015) 2844–2848.
[25] 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 (2) (2013) 491–511.
[26] 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 (37) (2015) 11571–11576.
[27] A.P. Gamiz-Hernandez, A. Jussupow, M.P. Johansson, V.R.I. Kaila, Terminal electron-proton transfer dynamics in the quinone reduction of respiratory complex I, J. Am. Chem. Soc. 139 (45) (2017) 16282–16288.
[28] M.A. Tocilescu, U. Fendel, K. Zwicker, S. Dröse, 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 (22) (2010) 625–632.
[29] P.K. Sinha, N. Castro-Guerrero, G. Patki, M. Sato, J. Torres-Bacete, S. Sinha, H. Miyoshi, A. Mitsuno-Yagi, T. Yagi, Conserved amino acid residues of the NuoD segment important for structure and function of *Escherichia coli* NDH-1 (complex I), Biochemistry 54 (3) (2015) 753–764.
[30] J.G. Fedor, A.J.Y. Jones, A. Di 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 (48) (2017) 12737–12742.
[31] R.G. Efremov, L.A. Sazanov, Structure of the membrane domain of the respiratory complex I, Nature 465 (7297) (2011) 414–420.
[32] V.R.I. Kaila, M. Wikström, G. Hummer, Electrostatics, hydration and proton transfer dynamics in the membrane domain of respiratory complex I, Proc. Natl. Acad. Sci. U. S. A. 111 (19) (2014) 6988–6993.
[33] 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 (31) (2017) E6314–E6321.
[34] O. Haapanen, V. Sharma, Role of water and protein dynamics in proton pumping by respiratory complex I, Sci. Rep. 7 (7747) (2017) 1–12.
[35] E. Screpanti, C. Hunte, Discontinuous membrane helices in transport proteins and their correlation with function, J. Struct. Biol. 159 (2007) 261–267.
[36] L.R. Forrest, et al., Mechanism for alternating access in neurotransmitter transporters, Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 10338–10343.
[37] J. Torres-Bacete, E. Nakamaru-Ogiso, A. Matsuno-Yagi, T. Yagi, Characterization of the NuoM (ND4) subunit in *Escherichia coli* NDH-1: conserved charged residues essential for energy-coupled activities, J. Biol. Chem. 282 (51) (2007) 36914–36922.
[38] M.C. Kao, S. Di Bernardo, M. Perego, E. Nakamaru-Ogiso, A. Matsuno-Yagi, T. Yagi, Functional role of four conserved charged residues on the membrane domain subunit NuoA of the proton-translocating NADH-quinone oxidoreductase from *Escherichia coli*, J. Biol. Chem. 279 (31) (2004) 32360–32366.
[39] J. Torres-Bacete, P.K. Sinha, N. Castro-Guerrero, A. Matsuno-Yagi, T. Yagi, Features of subunit NuoM (ND4) subunit in *Escherichia coli* NDH-1: topology and implication of conserved Glu144 for coupling site 1, J. Biol. Chem. 284 (48) (2009) 33062–33069.
[40] E. Nakamaru-Ogiso, M.C. Kao, H. Chen, S.C. Sinha, T. Yagi, T. Ohnishi, The membrane subunit NuoL (ND5) is involved in the indirect proton pumping mechanism of *Escherichia coli* Complex I, J. Biol. Chem. 285 (50) (2010) 39070–39078.
[41] J. Michel, J. Deleon-Rangel, S. Zhu, K. Van Ree, S.B. Vik, Mutagenesis of the L, M, and N subunits of complex I from *Escherichia coli* indicates a common role in function, PLoS One 6 (2) (2011) e17420.
[42] L. Euro, G. Belevich, M.I. Verkhovsky, M. Wikström, M. Verkhovskaya, Conserved lysine residues of the membrane subunit NuoM are involved in energy conversion by the proton-pumping NADH:ubiquinone oxidoreductase (complex I), Biochim. Biophys. Acta 1777 (9) (2008) 1166–1172.
[43] M. Wikström, G. Hummer, Stoichiometry of the proton translocation by respiratory complex I and its mechanistic implications, Proc. Natl. Acad. Sci. U. S. A. 109 (12) (2012) 4431–4436.
[44] M. Wikström, V. Sharma, V.R.I. Kaila, J.P. Hosler, G. Hummer, New perspectives on the proton pumping in cellular respiration, Chem. Rev. 115 (5) (2015) 2196–2221.
[45] U. Brandt, A two-state stabilization-change mechanism for proton-pumping complex I, Biochim. Biophys. Acta 1807 (10) (2011) 1364–1369.
[46] V.R.I. Kaila, Long-range proton-coupled electron transfer in biological energy conversion: towards mechanistic understanding of respiratory complex I, J. Roy Soc. Interfaces 141 (15) (2018) 20170916.
[47] O.S. Smart, J.G. Neduvelil, X. Wang, B.A. Wallace, M.S.P. Sansom, HOLE: a program for the analysis of the pore dimensions of ion channel structural models, J. Mol. Graph. 14 (6) (1996) 354–360.
[48] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an *N*·log(*N*) method for Ewald sums in large systems, J. Chem. Phys. 98 (12) (1993) 10089–10092.
[49] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, K. Laxmikant, K. Schulten, Scalable molecular dynamics with NAMD, J. Comput. Chem. 26 (16) (2005) 1781–1802.
[50] A.D. MacKerell, D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, All atom empirical potential for molecular modeling and dynamics studies of proteins, J. Phys. Chem. B 102 (18) (1998) 3586–3616.
[51] J.B. Klauda, R.M. Venable, J.A. Freites, J.W. O'Connor, D.J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A.D. MacKerell Jr., R.W. Pastor, Update of the CHARMM all-atom additive force field for lipids: validation of six lipid types, J. Phys. Chem. B 114 (23) (2010) 7830–7843.
[52] N.A. Baker, D. Sept, S. Joseph, M.J. Holst, J.A. McCammon, Electrostatic of nanosystems: application to microtubules and the ribosome, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 10037–10041.
[53] G. Kieseritzky, E.W. Knapp, Optimizing pKa computation in proteins with pH adapted conformations, Proteins 71 (3) (2008) 1335–1348.
[54] I. Bahar, T.R. Lezon, A. Bakan, I.H. Shrivastava, Normal mode analysis of biomolecular structures: functional mechanisms of membrane proteins, Chem. Rev. 110 (3) (2010) 1463–1497.
[55] A.W. Van Wynsberghe, Q. Cui, Interpreting correlated motions using normal mode analysis, Structure 14 (1) (2006) 1647–1653.
[56] A. Bakan, L.M. Meireles, I. Bahar, Prody: protein dynamics inferred from theory and experiments, Bioinformatics 27 (11) (2011) 1575–1577.
[57] A.D. Becke, Density-functional thermochemistry. III. The role of exact-exchange, J. Chem. Phys. 98 (7) (1993) 5648–5652.
[58] C. Lee, W. Yang, R.G. Parr, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density, Phys. Rev. B Condens. Matter 37 (2) (1988) 785–789.
[59] F. Weigend, R. Ahlrichs, Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: design and assessment of accuracy, Phys. Chem. Chem. Phys. 7 (18) (2005) 3297–3305.
[60] S. Grimme, J. Antony, S. Ehrlich, H.J. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, J. Chem. Phys. 132 (15) (2010) 154104.
[61] A. Klamt, Schüürmann, COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient, J. Chem. Soc. Perkin Trans. 2 (1993) 799–805.
[62] R. Ahlrichs, M. Bär, M. Häser, H. Horn, C. Kölmel, Electronic structure calculations on workstation computers: the program system Turbomole, Chem. Phys. Lett. 162 (3) (1989) 165–169.

# Article IV

BBA - Bioenergetics 1859 (2018) 326–332

Contents lists available at ScienceDirect

BBA - Bioenergetics

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

# Global collective motions in the mammalian and bacterial respiratory complex I

Andrea Di Luca, Ville R.I. Kaila*

*Department Chemie, Technische Universität München, Lichtenbergstr. 4, Garching, Germany*

ARTICLE INFO

*Keywords:*
Bioenergetics
Mitochondrial enzymes
NADH:ubiquinone oxidoreductase
Enzyme dynamics

ABSTRACT

The respiratory complex I is an enzyme responsible for the conversion of chemical energy into an electrochemical proton motive force across the membrane. Despite extensive studies, the mechanism by which the activity of this enormous, *ca.* 1 MDa, redox-coupled proton pump is regulated still remains unclear. Recent structural studies (Zhu et al., Nature 2016; Fiedorczuk et al., Nature 2016) resolved complex I in different conformations connected to the *active*-to-*deactive* (A/D) transition that regulate complex I activity in several species. Based on anisotropic network models (ANM) and principal component analysis (PCA), we identify here transitions between experimentally resolved structures of the mammalian complex I as low-frequency collective motions of the enzyme, highlighting similarities and differences between the bacterial and mammalian enzymes. Despite the reduced complexity of the smaller bacterial enzyme, our results suggest that the global dynamics of complex I is overall conserved. We further probe how the supernumerary subunits could be involved in the modulation of the A/D-transition, and show that in particular the 42 kDa and B13 subunits affect the global motions of the mammalian enzyme.

## 1. Introduction

Complex I (NADH: ubiquinone oxidoreductase) serves as the primary entry point for electrons in respiratory chains [1–3]. By pumping protons, complex I contributes to the proton motive force (*pmf*) across the membrane, which is in turn used for synthesis of adenosine triphosphate (ATP) and active transport [4,5]. The eukaryotic complex I (Fig. 1) is composed of up to 45 subunits [6–9], 14 of which are conserved, and constitute the inner core of the enzyme [10]. These conserved core subunits suggest that there are mechanistic similarities among different species, as observed in many other respiratory enzymes [11].

The hydrophilic domain of complex I catalyzes electron transfer (eT) between nicotinamide adenine dinucleotide (NADH), flavin mononucleotide (FMN), a chain of 8–9 iron sulfur centers (ISC), and quinone (Q) (Fig. 1), which is located in a binding pocket *ca.* 30 Å above the membrane surface [12–15]. The free energy released in the Q reduction process is further employed for pumping four protons across the mitochondrial inner membrane [16–20], whereas the reduced quinol ($QH_2$) shuttles the electrons further to respiratory complex III. The membrane domain of complex I is responsible for the proton pumping activity [21–24], and it comprises seven of the 14 core subunits, arranged in a *ca.* 180 Å-long array. Three of these subunits have an antiporter-like structure, and have evolved from Mrp-type (multiresistance and pH adaption) $Na^+/H^+$-antiporters. These antiporter-like subunits are likely to be responsible for pumping one proton each. From the Q-binding site, located at the interface between ND1, PSST, and the 49 kDa subunit (*B. taurus* nomenclature), a continuous chain of conserved polar and charged residues span the entire membrane domain. This amino acid array participates together with water molecules in catalyzing the pT reactions up to *ca.* 200 Å away from the Q-reduction site [13,23]. The membrane domain also comprises a horizontal (HL) helix spanning from ND2 to ND5 (NuoN to NuoL in *E. coli*) that was initially suggested to undergo conformational changes during enzyme turnover. Recent crosslinking, biochemical, and structural studies [25–27], however, suggest that the HL-helix is an element important for structural stability of the enzyme. Over the years, structural [28,29], biochemical [30–33], and computational studies [34–37] have been used to derive possible coupling mechanism between the terminal eT and pT based on conformational and/or electrostatic couplings, and involving one or two stroke mechanisms.

Recently resolved structures of complex I from *Yarrowia lipolytica* [24], *Bos taurus* [38], and *Ovis aries* [39] show that supernumerary subunits associate with the eukaryotic complex I, and form an outer shell around the central core enzyme (Fig. 1). The function of these supernumerary subunits is still unclear, but some possible roles have

* Corresponding author.
*E-mail address:* ville.kaila@ch.tum.de (V.R.I. Kaila).

https://doi.org/10.1016/j.bbabio.2018.02.001
Received 26 October 2017; Received in revised form 30 January 2018; Accepted 2 February 2018
Available online 05 February 2018
0005-2728/ 

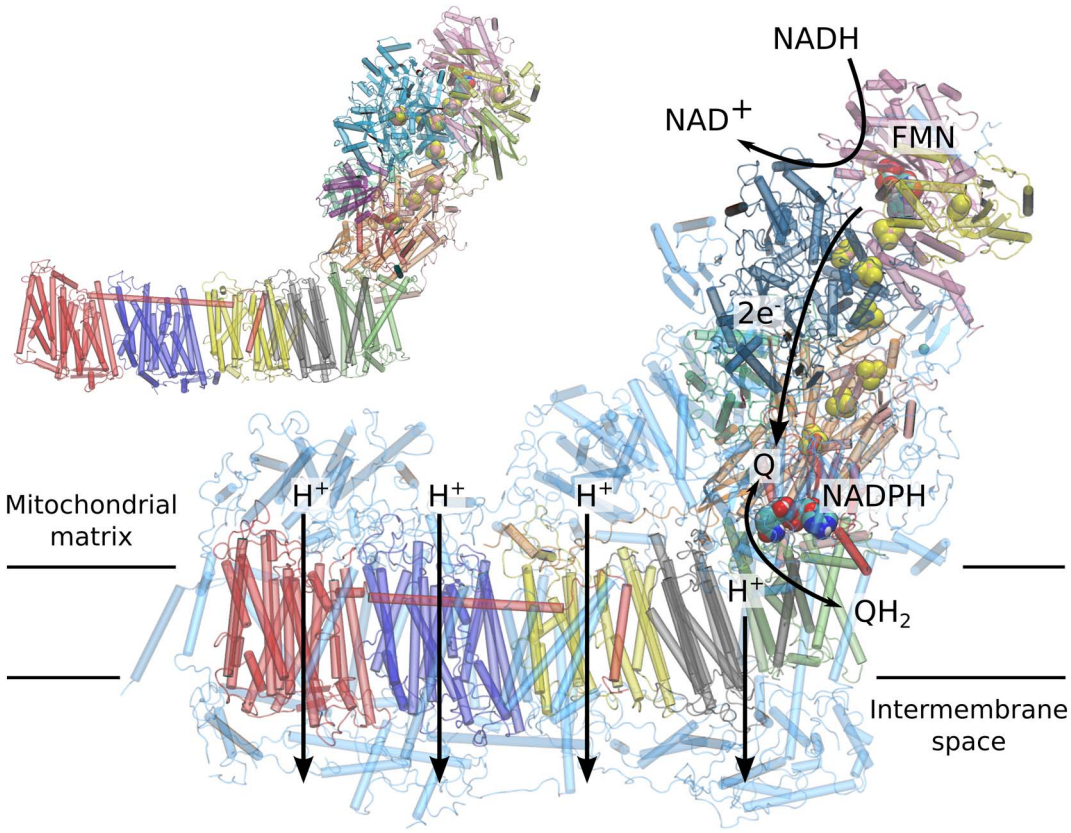


**Fig. 1.** Structure of the mammalian complex I from *B. taurus* class 2 (PDB ID: 5LC5). The supernumerary subunits are shown in light blue. *Inset*: structure of the bacterial complex I from *T. thermophilus* (PDB ID: 4HEA).

been suggested. For example, the 13-kDa subunit, which contains a zinc binding site, is important for complex I-assembly, and affects binding of other subunits upon deletion [40], whereas mutations causing the loss of the acyl carrier protein SDAP-α are involved in regulation of the oxidoreductase activity of complex I [41].

In addition to the complexity that arises from the tightly coupled eT and pT reactions, the mammalian complex I can also undergo an *active*-to-*deactive* (A/D) transition [42–44], and similar to other respiratory enzymes, *e.g.*, cytochrome *c* oxidase [45,46], re-activation of the deactive enzyme can be achieved by turnover [44]. Despite the A/D-transition was first observed a half century ago [47–49], and several efforts have been made to understand its molecular origins (*cf.* [43] and refs. therein), the mechanism of the A/D-transition still remains elusive. Biochemical studies [50–55] identified that ND1, ND3, and the 39 kDa (NDUFA9) subunit move during the A/D transition, and that a loop in ND3, that is solvent exposed in the deactive form, undergoes structural rearrangement. Recent structural insight in the A/D transition was also obtained from the near-atomic resolution structures of the eukaryotic complex I [24,38,39]. In particular, the *Y. lipolytica* complex was resolved in the *deactive* state [24], whereas the mammalian ovine and bovine structures were assigned to both *active* and *deactive* states after separation of different classes of cryo-EM data [38,39]. Subtle movements of all subunits were identified within the cryo-EM classes, but the proposed overall rearrangement for the A/D transitions somewhat differs between the two resolved cryo-EM structures of complex I. In particular, Fiedorczuk et al. [39] proposed an open/close rearrangement for the ovine enzyme, whereas Zhu et al. [38] found a rotation of the hydrophilic and membrane domain in the opposite direction in the bovine enzyme. Although the structural data is central for understanding the dynamics and mechanism of complex I, it should be emphasized that the assignment of structures to specific activity states has been only suggested, but not directly demonstrated. Further studies are thus needed to link together structural features with specific activity states of complex I.

Based on these recent experimentally resolved structures, we study here structural transitions between the different conformational states of the mammalian and bacterial complex I by normal mode and principal component analysis, which can be employed to identify global motions in proteins [56–69]. Comparison of the vibrational modes between the bacterial and mammalian enzyme highlights similarities as well as crucial differences in the global protein dynamics, and clarify the role of key subunits in the A/D transition.

## 2. Methods

### 2.1. Anisotropic network model (ANM) and principal component analysis (PCA)

The anisotropic network model (ANM) [56,57] is a popular elastic network model (ENM) that can predict protein collective motions based on the assumption that the native contact topology determines the global dynamics of the protein [58]. Given a protein structure, the ANMs are usually built on a coarse-grained level, *e.g.*, by using Cα atoms as nodes, and connecting all residue pairs within a certain distance threshold, $R_{cut}$ with harmonic springs of given force constants, $k_i$. The potential energy function of such network can be written as,

$$E_{network} = \frac{1}{2}\sum_{r_{ij}^0 < R_{cut}} k_i (r_{ij} - r_{ij}^0)^2, \quad (1)$$

where $r_{ij}$ and $r_{ij}^0$ represent the distances between Cα atoms in target and input structures, respectively. The model assumes, per definition, that the input structure is located at the minimum of the potential energy landscape, which allows performing a normal mode analysis (NMA) without geometry optimization that could distort the protein structure.

After elimination of translational and rotational degrees of freedom, diagonalization of the network Hamiltonian results in a set of eigenvectors and eigenvalues that represent the normal modes and their respective energies of the system. In its simplest form, the force constants between residue pairs are usually all set equal.

ANM and normal mode analysis (NMA)-based methods [59–61] are well-studied techniques that can give insight into global protein motion. Several studies have shown that parametric changes in the calculations do not considerably affect the results [62–64], suggesting that global motions of biomolecules are encoded in their geometry, *i.e.*, within the network connectivity. It has also been shown that calculated collective modes are able to capture transitions between biological relevant states [65,66].

In addition to NMA, principal component analysis (PCA) [59] can also be used to reconstruct information about global protein dynamics. PCA can be used to project out information about dynamics from an ensemble of structures, *e.g.*, from NMR derived structures [67] or from molecular dynamics (MD) simulations [68]. Given an ensemble of proteins structures, the PCA yields an orthogonal linear transformation of the coordinates that describe major fluctuation directions. The analysis is performed on the covariance matrix defined as,

$$C = \sum_k \Delta r^{(k)} \Delta r^{(k)T}, \tag{2}$$

in which $\Delta r^{(k)} = r^{(k)} - r^0$ is defined for each conformation $k$, and where $r^{(k)}$ is the position vectors of the $k$th conformation. $r^0$ are the equilibrium positions vectors, which are usually taken as the average of all atomic positions after structure/trajectory alignment. Diagonalization of the covariance matrix $C$ yields a set of eigenvectors and eigenvalues that describe global motions and their relative variance, respectively.

### 2.2. Overlap and collectivity of modes

To identify whether an experimentally resolved large-scale motion is similar to the outcome of the model, the overlap between different displacements can be calculated from the generalized cosine of two vectors [69],

$$O_j = \left| \sum_i^{3N} a_{ij} \Delta r_i \right| / \left[ \sum_i^{3N} a_{ij}^2 \sum_i^{3N} \Delta r_i^2 \right]^{1/2}, \tag{3}$$

where $N$ is the number of atoms, $a_{ij}$ are the normal modes coefficients, and $\Delta r_i$ are the displacements vectors between two structures. The overlap assumes values between 0 and 1, corresponding to no correlation (orthogonal vectors) and perfect correlation (parallel or anti-parallel vectors), respectively. Eq. (3) can also be used to calculate the overlap between normal modes, for which the geometrical displacement vector is replaced with a second normal mode.

Soft modes, *i.e.*, structural displacements along low energy modes, have usually a major contribution in protein global motions. To assess the degree of involvement of such network nodes, $j$, their *collectivity* can be evaluated from the "information entropy" of the eigenvectors of the system [70],

$$\kappa_j = \frac{1}{N} \exp\left\{ - \sum_{i=1}^{N} u_{i,j}^2 \log(u_{i,j}^2) \right\}, \tag{4}$$

where $N$ is the number of atoms, ${u_{i,j}}^2 = ({a_{i,x}}^2 + {a_{i,y}}^2 + {a_{i,z}}^2)_j/m_i$ is a measure of the displacement of atom $i$ in its $j$th mode, $\Sigma_{i=1}{}^N {u_{i,j}}^2 = 1$, and $m_i$ are the atom masses. $\kappa_j$ thus defines whether the network nodes move simultaneously, *i.e.*, collectively (see *e.g.* [70]).

### 2.3. Computational details of ANM and PCA

ANM analysis was performed on complex I structures from *B. taurus* [38] (classes 1, 2, and 3, PDB IDs: 5LC5, 5LDW, 5LDX), *Ovis aries* [39] (PDB ID: 5LNK), and *T. thermophilus* (PDB ID: 4HEA) [13] using $R_{cut}$ = 15.0 Å and $k$ (Eq. (1)) set equal for all interactions. In order to compute the overlap of the normal modes between the different species and structural classes, we discarded regions with different number of amino acids from the analysis. For the two structures of the mammalian complex I, we used both core and supernumerary subunits, whereas for comparing the mammalian and bacterial complex I structures, we analyzed only the 14 core subunits of the enzymes. The ANM and PC analysis were performed using ProDy [71]. The PC calculation was based on the last 500 ns of the MD trajectory of *T. thermophilus* complex I sampled every 0.1 ns (see below).

### 2.4. Classical molecular dynamics

To validate the ANM results, we performed classical atomistic molecular dynamics (MD) simulations of the bacterial complex I. To this end, we embedded the X-ray structure of complex I from *T. thermophilus* (PDB ID: 4HEA) [13] in a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) membrane, solvated with TIP3P water, and added sodium chloride to give an ionic strength of *ca.* 100 mM. The complete system comprised *ca.* 830,000 atoms. Ubiquinone ($Q_{10}$) was inserted in the Q-cavity, identified using the HOLE software [72], in its oxidized state with the head group interacting with His-38 and Tyr-87, and modeling the N2 iron-sulfur center in its reduced state. The simulations were performed using NAMD2 [73] together with CHARMM27 force field parameters [74,75], and DFT derived parameters for cofactors (iron-sulfur centers, $Q_{10}$, FMN). The MD simulation was extended for 600 ns using a 2 fs integration step, and an *NPT* ensemble ($T$ = 310 K, $p$ = 1 atm). Long-range electrostatics were treated with the Particle Mesh Ewald (PME) method [76].

## 3. Results and discussion

### 3.1. Global motions in complex I and resembles between structural transitions

The ANM analysis suggests that the lowest-frequency global mode of the bacterial and mammalian complex I comprises a scissoring motion of the hydrophilic and membrane domains, in which subunit ND1 (Nqo8 in *T. thermophilus*) acts as hinge (Fig. 2A, SI Movie 1) and connects the two regions, but shows a less pronounced motion (SI Fig. S1). The scissoring modes in both the bacterial and mammalian structures are in overall good agreement, but we observe subtle differences between the two enzymes that can be traced to the presence of the supernumerary subunit. In the mammalian enzyme, the 42 kDa and B13 subunits, located at the interface between the hydrophilic and membrane domains, add further stiffness in the motion (see below). This leads to a small rotation of the hydrophilic domain, which brings the 39 kDa subunit closer to the membrane domain than the B13 subunit (Fig. 2A, SI Movie 1).

The ANM analysis shows that the second lowest-frequency normal mode is a twisting motion that involves opposite movements of the hydrophilic and membrane domains (Fig. 2B, SI Movie 2), where ND1 serves as a junction point (SI Fig. S1). We find that also the twisting motion is less pronounced in the mammalian enzyme in the region around the 42 kDa subunit. Similar to the first two modes (Table 1), the higher frequency motions can also be considered as global motions, as they have a good degree of collectivity (Eq. (4)) [70]. However, these modes show a less clear directionality, and were therefore not analyzed further.

Analysis of the root-mean-square-fluctuations (RMSF) of the scissoring and twisting principal components for the *T. thermophilus* enzyme show a lower degree of rearrangement in the hinge region around the ND1/Nqo8 subunit (SI Fig. S1). The covariance matrices obtained from NMA and PCA analysis reveal highly correlated regions during the motion (Fig. 4). In particular, the motion of the entire hinge region that extends from Nqo4 to Nqo14 are well correlated, whereas the motion of subunits Nqo13 and Nqo12 correlate best with the nearby subunits. Interestingly, the mammalian subunit 42 kDa, which has been suggested to modulate the activity of complex I, show a high correlation

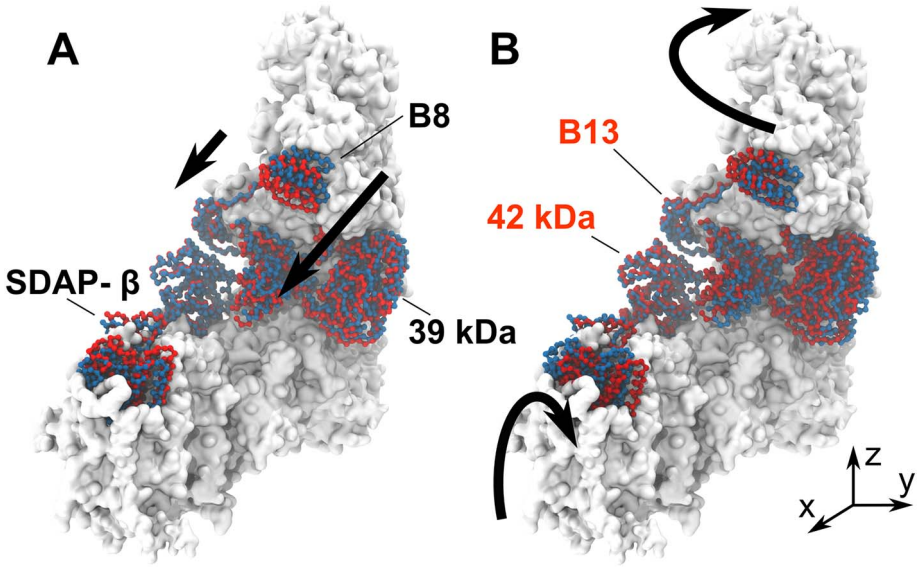


**Fig. 2.** Low-frequency modes of mammalian complex I. A) *Scissoring* (*bending*) of the two domains and B) *twisting* of hydrophilic and membrane domain. Relative displacement of supernumerary subunits is shown in red and blue. The 42 kDa and B13 subunits, located at the interface between the hydrophilic and membrane domain could important role in regulation of the enzyme dynamics (see Table 3).

with the Nqo14 motion (see below). In the ANM and PC analysis, we observe only small motions in the Q-binding site region, but dynamically disordered regions, such as loops, are usually not well captured by the harmonic approximation employed in the mode analysis approach.

To understand how the experimentally identified structural classes are related to their normal modes, we analyzed the overlap between the ANM-identified modes and the structural changes between classes 1 → 3 and 1 → 2 for the bovine complex I (see Methods). We observe a strong correlation between the scissoring and twisting modes with the transitions between the structural classes 1 → 3 and 1 → 2, respectively (Fig. 3). This suggests that the transitions between the different EM-structural classes could be related to the lowest normal modes of the enzyme. It is thus tempting to draw the conclusion that since the different cryo-EM derived structural classes have been related to the A/D transition [38,39], the global complex I motions identified here by the ANM and PC analysis, could also be related to the A/D-transition path.

Among the three resolved structural classes of the bovine enzyme, class 3 has the lowest resolution, and represent a “broken” form of the enzyme, in which the transverse helix of the ND5 subunit has a higher degree of disorder [38]. Although this imposes some uncertainty for the class 3 structure, we nevertheless find close similarities between the scissoring motion obtained from the ANM analysis in all studied structures. Moreover, the structures have a similar motion *direction* for the 1-*to*-3 transition in the bovine structure, as well as in the open/close conformations of the ovine enzyme. This suggests that movement “towards” the class 3 structure, could be related to the conformational dynamics of complex I, which in turn may lead to a partial disruption of the enzyme.

**Table 1**
Overlap of normal modes between the structure of *B. taurus* class 2 (PDB ID: 5LC5) with *B. taurus* classes 1 and 3, and complex I from *Ovis aries* and *T. thermophilus*. The collectivity of the low frequency normal modes (*Coll.*) was calculated from the normal mode analysis using Eq. (4).

| Structure/mode | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| *B. taurus* Class I | 0.99 | 0.99 | 0.99 | 0.97 |
| *B. taurus* Class III | 0.99 | 0.99 | 0.99 | 0.72 |
| *Ovis aries* | 0.98 | 0.97 | 0.97 | 0.73 |
| *T. thermophilus* | 0.82 | 0.73 | 0.46 | 0.58 |
| Coll. Class 1 | 0.65 | 0.57 | 0.53 | 0.13 |
| Coll. Class 2 | 0.64 | 0.57 | 0.50 | 0.13 |
| Coll. Class 3 | 0.66 | 0.57 | 0.55 | 0.003 |
| Coll. Ovine | 0.65 | 0.58 | 0.55 | 0.57 |
| Coll. Thermo | 0.64 | 0.61 | 0.42 | 0.70 |

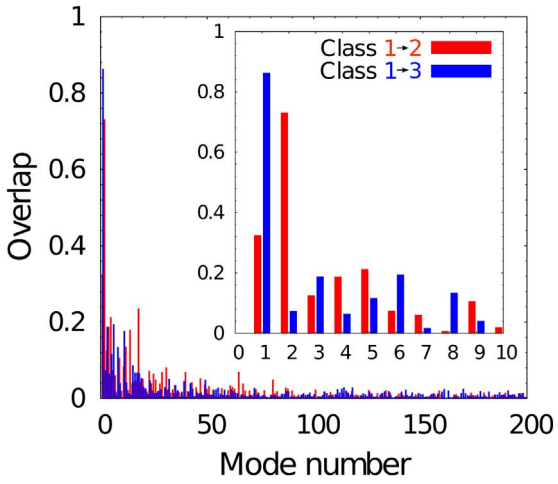


**Fig. 3.** Overlap between class 1–2 and class 1–3 transitions and normal modes from *B. taurus* class 2 structure. *Inset*: the first lowest eigenmodes for the scissoring and bending motions. Only the first and second normal modes show a strong overlap with the transition 1–3 and 1–2, respectively.

### 3.2. ANM analysis

Our ANM analysis suggests that the overlap between the first four normal modes of the different structural classes of bovine complex I is close to 1 (Table 1), suggesting that the motions have high structural and dynamic similarities. These results are consistent with the coarse-grained nature of the ANM technique, and the small relative displacement in the subunits, which is also consistent with the structural data [38,39]. Moreover, these low-energy modes are not affected by the structural details, suggesting that the obtained results are robust. We also compared the normal modes calculated for class 2 between the bovine and ovine structures. Interestingly, the normal modes of these enzymes are also strikingly similar (Table 1), suggesting that they have similar global dynamics. Moreover, our analysis suggests that the overlap between the first two normal modes in the *T. thermophilus* and bovine complex I is around 80% (Table 1), and about 50% for the third and fourth normal modes, whereas we obtain smaller correlations for motions in the higher frequency domains. The ANM analysis, however, does not discriminate between different protein sequences, as it is based on a coarse-grained model, and is thus dominated by the network of nodes, which in turn is determined by the protein structure.

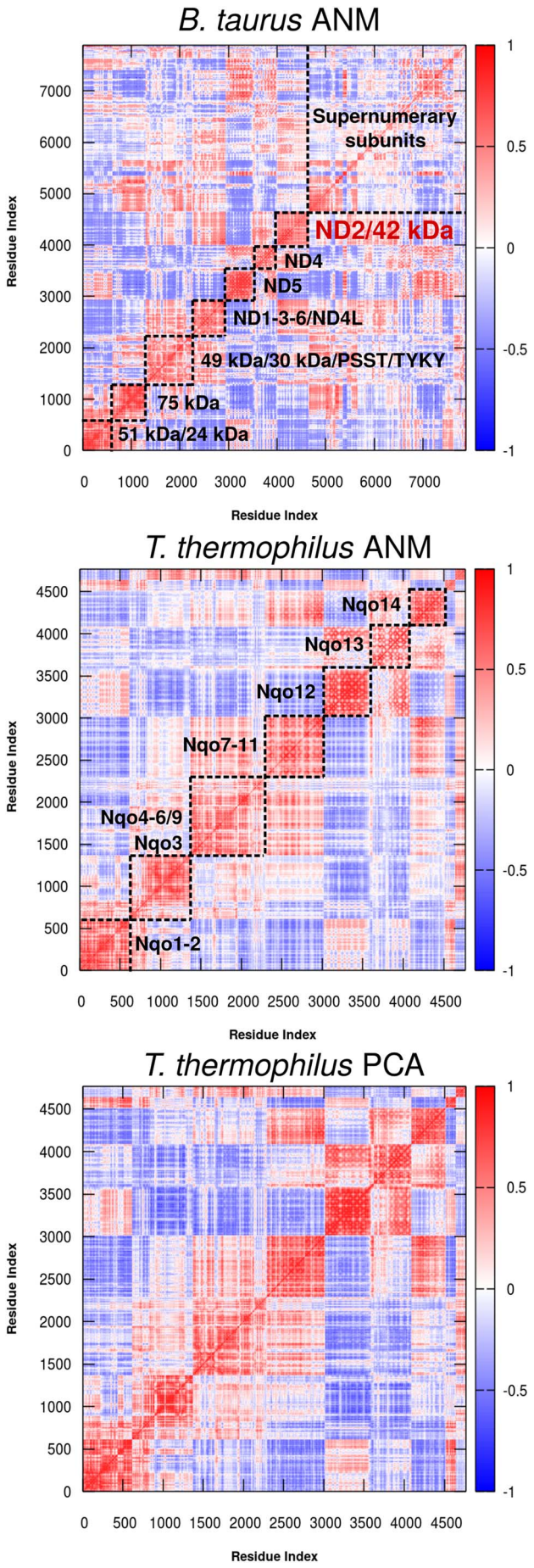


**Fig. 4.** Covariance matrices from ANM and PC analysis for complex I from *T. thermophilus* and *B. taurus*. The ANM covariance matrix is calculated using all normal modes, whereas the PCA covariance matrix is derived from all calculated PCs (see Methods section). The results show high similarities between the PCA and ANM of *T. thermophilus*. Comparison between the modes, suggest that the overall dynamics of complex I core subunits is conserved. The hinge region, comprising subunits Nqo4-11, undergoes a correlated motion, which extends up to Nqo14. The most distal subunit Nqo12 shows a high correlation with the Nqo13 subunit motion. Interestingly, the motion of subunit ND2 (Nqo14) is strongly coupled with the supernumerary subunit 42 kDa. The residue index is the cumulative residue number, with subunit ordering marked in the figure.

Considering a force constant of 1 kcal·mol$^{-1}$·Å$^2$ between network nodes [56], the frequencies for both modes of the mammalian and bacterial enzymes are expected to be in the range of *ca.* 10 cm$^{-1}$.

Overall, we obtain similar normal modes for the bacterial and mammalian enzymes, supporting that the low-frequency motions are structurally encoded. However, we find that the subunit movements are not identical, which suggests that the supernumerary subunits have an important role in determining the dynamics of the mammalian complex I (see below).

### 3.3. Principal component analysis

To confirm the results obtained from the ANM analysis, we also performed classical molecular dynamics (MD) simulations on the complex I from *T. thermophilus* (see Methods). Based on the MD trajectories, we calculated and compared principal components (PCs) with the normal modes obtained from the ANM based on the experimentally-resolved structures. Interestingly, the first PC, which account for about 50% of the fluctuations in *T. thermophilus*, shows a high overlap with the first normal mode of the bovine structure (Table 2), further supporting that the bacterial and mammalian enzymes have similar global dynamics. The second and third PCs of the *T. thermophilus* structure show a less clear overlap, but both modes, nevertheless, correlate well with the second normal mode observed in the bovine structure. This effect might result from the similarities in the dominant normal modes of the bovine structure, and thus support that the principal modes of respiratory complex I are overall similar (Table 2). This, together with the comparison between the normal modes of the mammalian and bacterial enzymes, suggest that the two lowest frequency modes are consistently observed in all calculations.

### 3.4. Supernumerary subunits modulate the A/D transition

The supernumerary subunits of the mammalian complex I form a cage-like structure around the core enzyme, and are likely to be structurally important [38,39]. To probe the role of the supernumerary subunits, we individually removed subunits from the mammalian structure followed by ANM analysis.

We find that simultaneous removal of all supernumerary subunits modifies the scissoring and twisting modes, and lead to global modes that closely resemble the motion of the bacterial complex I (Table 3). However, deletion of single supernumerary subunits suggests that the 42 kDa and B13 subunits (NDUFA10 and NDUFA5 in humans) have a major impact on the conformational transitions (Table 3, Fig. 2, SI

**Table 2**
Overlap between principal components of *T. thermophilus* complex I from 500 ns MD simulation and ANM normal modes for *T. thermophilus* and *B. taurus*.

| | PC1 | PC2 | PC3 | PC4 |
|---|---|---|---|---|
| Contribution (%) | 46.7 | 25.5 | 8.3 | 5.19 |
| ANM *T. thermophilus* Mode 1 | 0.88 | 0.48 | 0.04 | 0.33 |
| ANM *T. thermophilus* Mode 2 | 0.28 | 0.52 | 0.83 | 0.22 |
| ANM *B. taurus* Mode 1 | 0.86 | 0.19 | 0.22 | 0.36 |
| ANM *B. taurus* Mode 2 | 0.1 | 0.74 | 0.61 | 0.07 |

**Table 3**
Overlap of the normal modes between the mammalian (*Bos taurus*) and bacterial (*T. thermophilus*) complex I. Deletion of subunits B13 and 42 kDa produce an effect similar to the deletion of all the supernumerary subunits (highlighted in bold), suggesting a critical role of these supernumerary subunits modulate the collective motions of complex I.

| | Mode 1 | Mode 2 | | Mode 1 | Mode 2 |
|---|---|---|---|---|---|
| Complete *B. taurus* | 0.82 | 0.73 | Without B12 | 0.82 | 0.75 |
| Only 14 core subunits | 0.97 | 0.95 | Without ASHI | 0.83 | 0.72 |
| Without 42 kDa | **0.91** | **0.95** | Without B15 | 0.82 | 0.72 |
| Without 39 kDa | 0.81 | 0.71 | Without B22 | 0.82 | 0.75 |
| Without 18 kDa | 0.82 | 0.73 | Without B18 | 0.82 | 0.72 |
| Without 13 kDa | 0.83 | 0.77 | Without PDSW | 0.83 | 0.75 |
| Without B8 | 0.82 | 0.74 | Without B17.2 | 0.83 | 0.76 |
| Without SDAP-α | 0.82 | 0.73 | Without B14.5a | 0.83 | 0.78 |
| Without SDAP-β | 0.82 | 0.76 | Without 10 kDa | 0.81 | 0.72 |
| Without B13 | **0.93** | **0.92** | Without SGDH | 0.83 | 0.74 |
| Without B14 | 0.84 | 0.75 | Without B17 | 0.82 | 0.74 |
| Without PGIV | 0.82 | 0.71 | Without AGGG | 0.68 | 0.41 |
| Without B14.7 | 0.79 | 0.68 | Without B14.5b | 0.80 | 0.66 |
| Without B16.6 | 0.82 | 0.72 | Without 15 kDa | 0.82 | 0.72 |
| Without MWFE | 0.83 | 0.73 | Without MNLL | 0.82 | 0.73 |
| Without B9 | 0.82 | 0.73 | Without ESSS | 0.83 | 0.76 |
| Without KFYI | 0.82 | 0.73 | | | |

Movies 3 and 4). These subunits are located at the interface between the membrane and hydrophilic domains, and come in close contact with each other. This suggests that the 42 kDa and B13 subunits could account for the major differences in the global dynamics between the bacterial and mammalian complex I by adding further stiffness between the two domains and securing a strong coupling required for the A/D transition. Mutation experiments in this linked region, especially at the interface between the two subunits could thus help to clarify the role of the enzyme function.

The *Yarrowia lipolytica* complex I, which lacks the 42 kDa subunit, also undergo an A/D transition [41]. This behavior further suggests that the B13 subunit (NUFM in *Y. lipolytica*) could be a key subunit in modulating the A/D transition, and that the 42 kDa subunit would have an important role in its regulation. Interestingly, the A/D transition has a much lower activation barrier in complex I from *Y. lipolytica* [41], supporting that the 42 kDa subunit could function as a modulator of some putative "*activity switch*".

Given the similarities in the global motions identified here with the transitions between the experimentally identified structural classes, we suggest that observed modes could be of functional relevance in complex I. One of these motions could be related to the A/D transition, while the other mode could be involved in, *e.g.*, the enzyme turnover or quinone dynamics, and blocked or energetically unfavored, if the enzyme resides in the *deactive* state. This could thus provide a coupling between the A/D-transition and the affinity for the Q substrate. The supernumerary subunits could thus together with their proposed stabilization role [38], also affects the global dynamics complex I by influencing its soft modes.

Considering the similarities in the collective motions between the bacterial and mammalian enzymes, we suggest that complex I has evolved to regulate its activity by an intrinsic motion that is further controlled by the supernumerary subunits. This model does not exclude the possibility that such intrinsic motions could also have a biological role in regulation of the bacterial enzymes. However, further studies are needed to elucidate the biological function of such putative conformational changes.

## 4. Conclusion

We have studied here both bacterial and mammalian structures of the respiratory complex I by ANM and PC analysis, and discussed possible implication for the observed low-frequency collective motions in its global dynamics. The combined results suggest that the lowest frequency normal modes are conserved from the bacterial to mammalian enzymes. Similarities between the calculated motions and transition between the recently resolved experimental structure further suggest that these movements are involved in specific biological activities, such as the A/D transition. We further identified that the 42 kDa and B13 subunits are responsible for the differences in the normal modes between the mammalian and bacterial enzymes, providing a mechanistic picture on how the conformational dynamics of complex I in controlled.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbabio.2018.02.001.

## Transparency document

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

## Acknowledgements

We thank Judy Hirst and Volker Zickermann for insightful discussions. This work was supported by the German Research Foundation. The CPU time provided by the SuperMuc at the Leibniz Rechenzentrum (computing grant: pr48de and pr27xu).

## References

[1] J. Hirst, Mitochondrial complex I, Annu. Rev. Biochem. 82 (1) (2013) 551–575.
[2] U. Brandt, Energy converting NADH: ubiquinone oxidoreductase (complex I), Annu. Rev. Biochem. 75 (1) (2006) 69–92.
[3] L.A. Sazanov, A giant molecular proton pump: structure and mechanism of respiratory complex I, Nat. Rev. Mol. Cell. Biol. 16 (6) (2015) 375–388.
[4] P. Mitchell, Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism, Nature 191 (4784) (1961) 144–148.
[5] M. Yoshida, E. Muneyoki, T. Hisabori, ATP synthase – a marvellous rotary engine of the cell, Nat. Rev. Mol. Cell. Biol. 2 (9) (2001) 669–677.
[6] K.R. Vinothkumar, J. Zhu, J. Hirst, Architecture of mammalian respiratory complex I, Nature 515 (7525) (2014) 80–84.
[7] C. Wirth, U. Brandt, C. Hunte, V. Zickermann, Structure and function of mitochondrial complex I, Biochim. Biophys. Acta 1857 (7) (2016) 902–914.
[8] H. Angerer, K. Zwicker, Z. Wumaier, L. Sokolova, H. Heide, M. Steger, S. Kaiser, E. Nubel, B. Brutschy, M. Radermacher, U. Brandt, V. Zickermann, A scaffold of accessory subunits links the peripheral arm and the distal proton-pumping module of mitochondrial complex I, Biochem. J. 437 (2) (2011) 279–288.
[9] K. Kmita, V. Zickermann, Accessory subunits of mitochondrial complex I, Biochem. Soc. Trans. 41 (5) (2013) 1272–1279.
[10] J. Hirst, J. Carroll, I.M. Fearnley, R.J. Shannon, J.E. Walker, The nuclear encoded subunits of complex I from bovine heart mitochondria, Biochim. Biophys. Acta 1604 (3) (2003) 135–150.
[11] M. Wikström, V. Sharma, V.R.I. Kaila, J.P. Hosler, G. Hummer, New perspectives on proton pumping in cellular respiration, Chem. Rev. 115 (5) (2015) 2196–2221.
[12] C. Hunte, V. Zickermann, U. Brandt, Functional modules and structural basis of conformational coupling in mitochondrial complex I, Science 329 (5990) (2010) 448–451.
[13] R. Baradaran, J.M. Berrisford, G.S. Minhas, L.A. Sazanov, Crystal structure of the entire respiratory complex I, Nature 494 (7438) (2013) 443–448.
[14] T. Ohnishi, Iron-sulfur clusters/semiquinones in complex I, Biochim. Biophys. Acta 1364 (2) (1998) 186–206.
[15] L.A. Sazanov, P. Hinchliffe, Structure of the hydrophilic domain of respiratory complex I from *Thermus thermophilus*, Science 311 (5766) (2006) 1430–1536.
[16] M. Wikström, Two protons are pumped from the mitochondrial matrix per electron transferred between NADH and ubiquinone, FEBS Lett. 169 (2) (1984) 300–304.
[17] A.J.Y. Jones, J.N. Blaza, F. Varghese, J. Hirst, Respiratory complex I in *Bos taurus* and *Paracoccus denitrificans* pumps four protons across the membrane for every NADH oxidised, J. Biol. Chem. 292 (12) (2017) 4987–4995.
[18] A. Galkin, S. Dröse, U. Brandt, The proton pumping stoichiometry of purified mitochondrial complex I reconstituted into proteoliposomes, Biochem. Biophys. Acta 1757 (12) (2006) 1575–1581.
[19] A.S. Galkin, V.G. Grivennikova, A.D. Vinogradov, $\rightarrow H^+/2e^-$ stoichiometry in NADH quinone reductase reactions catalyzed by bovine heart mitochondrial particles, FEBS Lett. 451 (2) (1999) 157–161.
[20] J.G. Fedor, A.J.Y. Jones, A. Di 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 (48) (2017) 12737–12742.
[21] S. Dröse, S. Krack, L. Sokolova, K. Zwicker, H.-D. Barth, N. Morgner, H. Heide, M. Steger, E. Nübel, V. Zickermann, S. Kerscher, B. Brutschy, M. Radermacher, U. Brandt, Functional dissection of the proton pumping modules of mitochondrial complex I, PLoS Biol. 9 (8) (2011) e1001128.
[22] M.L. Verkovskaya, D.A. Bloch, Energy-converting respiratory complex I: on the way

to the molecular mechanism of the proton jump, Int. J. Biochem. Cell Biol. 45 (2) (2013) 491–511.
[23] R.G. Efremov, L.A. Sazanov, Structure of the membrane domain of the respiratory complex I, Nature 465 (7297) (2011) 414–420.
[24] 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 (6217) (2015) 44–49.
[25] G. Belevich, J. Knuuti, M.I. Verkhovsky, M. Wikström, M. Verkhovskaya, Probing the mechanistic role of the long α-helix in subunit L of respiratory complex I from *Escherichia coli* by site-directed mutagenesis, Mol. Microbiol. 82 (5) (2011) 1086–1095.
[26] J. Torres-Bacete, P.K. Sinha, A. Matsuno-Yagi, T. Yagi, Structural contribution of C-terminal segments of NuoL (ND5) and NuoM (ND4) subunits of complex I from *Escherichia coli*, J. Biol. Chem. 286 (39) (2011) 34007–34014.
[27] S. Zhu, S.B. Vik, Constraining the lateral helix or respiratory complex I by cross-linking does not impair enzyme activity or proton translocation, J. Biol. Chem. 290 (34) (2015) 20761–20773.
[28] U. Brandt, A two-state stabilization-change mechanism for proton-pumping complex I, Biochim. Biophys. Acta 1807 (10) (2011) 1364–1369.
[29] R.G. Efremov, L.A. Sazanov, The coupling mechanism of respiratory complex I – a structural and evolutionary perspective, Biochim. Biophys. Acta 1817 (10) (2012) 1785–1795.
[30] M.L. Verkhovskaya, N. Belevich, L. Euro, M. Wikström, Real-time electron transfer in respiratory complex I, Proc. Natl. Acad. Sci. U. S. A. 105 (10) (2008) 3763–3767.
[31] T. Ohnishi, E. Nakamaru-Ogiso, S.T. Ohnishi, A new hypothesis on the simultaneous direct and indirect proton pump mechanisms in NADH–quinone oxidoreductase (complex I), FEBS Lett. 584 (19) (2010) 4131–4137.
[32] J.R. Treberg, M.D. Brand, A model of the proton translocation mechanism of complex I, J. Biol. Chem. 286 (20) (2011) 17579–17584.
[33] M. Verkhovsky, D.A. Bloch, M. Verkhovskaya, Tightly-bound ubiquinone in the *Escherichia coli* respiratory complex I, Biochim. Biophys. Acta 1817 (9) (2012) 1550–1556.
[34] V.R.I. Kaila, M. Wikström, G. Hummer, Electrostatics, hydration and proton transfer dynamics in the membrane domain of respiratory complex I, Proc. Natl. Acad. Sci. U. S. A. 111 (19) (2014) 6988–6993.
[35] 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 (37) (2015) 11571–11576.
[36] 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 (31) (2017) E6314–E6321.
[37] O. Haapanen, V. Sharma, Role of water and protein dynamics in proton pumping by respiratory complex I, Sci. Rep. 7 (7747) (2017) 1–12.
[38] J. Zhu, K.R. Vinothkumar, J. Hirst, Structure of mammalian respiratory complex I, Nature 536 (7616) (2016) 354–358.
[39] 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 (7625) (2016) 406–410.
[40] K. Kmita, C. Wirth, J. Warnau, S. Guerrero-Castillo, C. Hunte, G. Hummer, V.R.I. Kaila, K. Zwicker, U. Brandt, V. Zickermann, Accessory NUMM (NDUFS6) subunit harbors a Zn-binding site and is essential for biogenesis of mitochondrial complex I, Proc. Natl. Acad. Sci. U. S. A. 112 (18) (2015) 5685–5690.
[41] H. Angerer, M. Radermacher, M. Mańkowska, M. Steger, K. Zwicker, H. Heide, I. Wittig, U. Brandt, V. Zickermann, The LYR protein subunit NB4M/NDUFA6 if mitochondrial complex I anchors an acyl carrier protein and is essential for catalytic activity, Proc. Natl. Acad. Sci. U. S. A. 111 (14) (2014) 5207–5212.
[42] E. Maklashina, A.B. Kotlyar, G. Cecchini, Active/de-active transition of respiratory complex I in bacteria, fungi and animals, Biochim. Biophys. Acta 1606 (1) (2003) 95–103.
[43] S. Dröse, A. Stepanova, A. Galkin, Ischemic A/D transition of mitochondrial complex I and its role in ROS generation, Biochim. Biophys. Acta 1857 (7) (2016) 946–957.
[44] M. Babot, A. Birch, P. Labarbuta, A. Galkin, Characterisazion of the active/de-active transition of mitochondrial complex I, Biochim. Biophys. Acta 1837 (7) (2014) 1083–1092.
[45] D. Jancura, V. Berka, M. Antalik, J. Bagelova, R.B. Gennis, G. Palmer, M. Fabian, Spectral and kinetic equivalence of oxidized cytochrome C oxidase as isolated and "activated" by reoxidation, J. Biol. Chem. 281 (41) (2006) 30319–30325.
[46] D. Bloch, I. Belevich, A. Jasaitis, C. Ribacka, A. Puustinen, M.I. Verkhovsky, M. Wikström, The catalytic cycle of cytochrome c oxidase is not the sum of its two halves, Proc. Natl. Acad. Sci. U. S. A. 101 (2) (2004) 529–533.
[47] A.D. Vinogradov, Catalytic properties of the mitochondrial NADH-ubiquinone oxidoreductase (complex I) and the pseudo-reversible active/inactive enzyme transition, Biochem. Biophys. Acta 1364 (2) (1998) 169–185.
[48] S. Minakami, F.J. Schindler, R.W. Estabrook, Hydrogen transfer between reduced diphosphopyridine nucleotide dehydrogenase and the respiratory chain. II. An initial lag in the oxidation of reduced diphosphopyridine nucleotide, J. Biol. Chem. 239 (1964) 2049–2054.
[49] V.N. Luzikov, V.A. Saks, I.V. Berezin, Comparative study of thermal degradation of electron transfer particle and reconstituted respiratory chain. Relation of electron transfer to reactivation of submitochondrial particles, Biochim. Biophys. Acta 223 (1) (1970) 16–30.
[50] A.B. Kotlyar, A.D. Vinogradov, Slow active/inactive transition of the mitochondrial NADH-ubiquinone oxidoreductase, Biochim. Biophys. Acta 1019 (2) (1990) 129–134.
[51] E.V. Gavrikova, A.D. Vinogradov, Active/de-active state transition of the mitochondrial complex I as revealed by specific sulfhydryl group labeling, FEBS Lett. 455 (1–2) (1999) 36–40.
[52] A. Galkin, B. Meyer, I. Wittig, M. Karas, H. Schagger, A. Vinogradov, U. Brandt, Identification of the mitochondrial ND3 subunit as structural component involved in the active/deactive enzyme transition of respiratory complex I, J. Biol. Chem. 283 (30) (2008) 20907–20913.
[53] M. Babot, A. Galkin, Molecular mechanism and physiological role of the active-deactive transition of mitochondrial complex I, J. Biochem. Soc. Trans. 41 (5) (2013) 1325–1330.
[54] M. Ciano, M. Fuszard, H. Heide, C.H. Botting, A. Galkin, Conformation specific crosslinking of mitochondrial complex I, FEBS Lett. 587 (7) (2013) 867–872.
[55] M. Babot, P. Labarbuta, A. Birch, S. Kee, Z. Fuszard, K. Botting, I. Wittig, H. Heide, A. Galkin, ND3, ND1 and 39 kDa subunits are more exposed in the de-active form of bovine mitochondrial complex I, Biochim. Biophys. Acta 1837 (6) (2014) 929–939.
[56] A.R. Atilgan, S.R. Durell, R.L. Jernigan, M.C. Demirel, O. Keskin, I. Bahar, Anisotropy of fluctuation dynamics of proteins with an elastic network model, Biophys. J. 80 (1) (2001) 505–515.
[57] P. Doruker, A.R. Atilgan, I. Bahar, Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: application to alpha-amylase inhibitor, Proteins 40 (3) (2000) 512–524.
[58] I. Bahar, A.J. Rader, Coarse-grained normal mode analysis in structural biology, Curr. Opin. Struct. Biol. 15 (5) (2005) 586–592.
[59] I. Bahar, T.R. Lezon, A. Bakan, I.H. Shrivastava, Normal mode analysis of biomolecular structures: functional mechanisms of membrane proteins, Chem. Rev. 110 (3) (2010) 1463–1497.
[60] L. Skjaerven, S.M. Hollup, N. Reuter, Normal mode analysis for proteins, J. Mol. Struct. THEOCHEM 898 (1) (2009) 42–48.
[61] A.W. Van Wynsberghe, Q. Cui, Interpreting correlated motions using normal mode analysis, Structure 14 (1) (2006) 1647–1653.
[62] M. Lu, J. Ma, The role of shape in determining molecular motions, Biophys. J. 89 (4) (2005) 2395–2401.
[63] S. Nicolay, Y.H. Sanejouand, Functional modes of proteins are among the most robust, Phys. Rev. Lett. 96 (7) (2006) 078104.
[64] F. Tama, C.L. Brooks, Symmetry, form and shape: guiding principles for robustness in macromolecular machines, Ann. Rev. Biophys. Struct. 35 (2006) 115–133.
[65] F. Zhu, G. Hummer, Gating transition of pentameric ligand-gated ion channels, Biophys. J. 97 (9) (2009) 2456–2463.
[66] Q. Cui, G. Li, J. Ma, M. Karplus, A normal mode analysis of structural plasticity in the biomolecular motor F(1)-ATPase, J. Mol. Biol. 340 (2) (2004) 345–372.
[67] P.W. Howe, Principal component analysis of protein structure ensembles calculated using NMR data, J. Biomol. NMR 20 (1) (2001) 61–70.
[68] Q. Luo, E.E. Boczek, Q. Wang, J. Buchner, V.R.I. Kaila, Hsp90 dependence of a kinase is determined by its conformational landscape, Sci. Rep. 7 (43996) (2017) 1–11.
[69] F. Tama, Y.H. Sanejouand, Conformational change of proteins arising from normal mode calculations, Protein Eng. 14 (1) (2001) 1–6.
[70] R. Bruschweiler, Collective protein dynamics and nuclear spin relaxation, J. Chem. Phys. 102 (8) (1995) 3396–3403.
[71] A. Bakan, L.M. Meireles, I. Bahar, Prody: protein dynamics inferred from theory and experiments, Bioinformatics 27 (11) (2011) 1575–1577.
[72] O.S. Smart, J.G. Neduvelil, X. Wang, B.A. Wallace, M.S.P. Sansom, HOLE: a program for the analysis of the pore dimensions of ion channel structural models, J. Mol. Graph. 14 (6) (1996) 354–360.
[73] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, K. Laxmikant, K. Schulten, Scalable molecular dynamics with NAMD, J. Comput. Chem. 26 (16) (2005) 1781–1802.
[74] A.D. MacKerell, D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, All atom empirical potential for molecular modeling and dynamics studies of proteins, J. Phys. Chem. B 102 (18) (1998) 3586–3616.
[75] J.B. Klauda, R.M. Venable, J.A. Freites, J.W. O'Connor, D.J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, MacKerell AD Jr., R.W. Pastor, Update of the CHARMM all-atom additive force field for lipids: validation of six lipid types, J. Phys. Chem. B 114 (23) (2010) 7830–7843.
[76] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N-log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (12) (1993) 10089–10092.

# Article V

# Redox-coupled quinone dynamics in the respiratory complex I

Judith Warnau[a,b,1], Vivek Sharma[c,d,1,2], Ana P. Gamiz-Hernandez[a], Andrea Di Luca[a], Outi Haapanen[c], Ilpo Vattulainen[c,e,f], Mårten Wikström[d], Gerhard Hummer[b,g,2], and Ville R. I. Kaila[a,2]

[a]Department Chemie, Technische Universität München, D-85748 Garching, Germany; [b]Department of Theoretical Biophysics, Max Planck Institute of Biophysics, 60438 Frankfurt am Main, Germany; [c]Department of Physics, University of Helsinki, FI-00014 Helsinki, Finland; [d]Institute of Biotechnology, University of Helsinki, FI-00014 Helsinki, Finland; [e]Laboratory of Physics, Tampere University of Technology, FI-33101 Tampere, Finland; [f]MEMPHYS – Center for Biomembrane Physics, Department of Physics, University of Southern Denmark, 5230 Odense, Denmark; and [g]Institute of Biophysics, Goethe University Frankfurt, 60438 Frankfurt am Main, Germany

Edited by Michael L. Klein, Temple University, Philadelphia, PA, and approved July 13, 2018 (received for review March 29, 2018)

**Complex I couples the free energy released from quinone (Q) reduction to pump protons across the biological membrane in the respiratory chains of mitochondria and many bacteria. The Q reduction site is separated by a large distance from the proton-pumping membrane domain. To address the molecular mechanism of this long-range proton-electron coupling, we perform here full atomistic molecular dynamics simulations, free energy calculations, and continuum electrostatics calculations on complex I from *Thermus thermophilus*. We show that the dynamics of Q is redox-state-dependent, and that quinol, $QH_2$, moves out of its reduction site and into a site in the Q tunnel that is occupied by a Q analog in a crystal structure of *Yarrowia lipolytica*. We also identify a second Q-binding site near the opening of the Q tunnel in the membrane domain, where the Q headgroup forms strong interactions with a cluster of aromatic and charged residues, while the Q tail resides in the lipid membrane. We estimate the effective diffusion coefficient of Q in the tunnel, and in turn the characteristic time for Q to reach the active site and for $QH_2$ to escape to the membrane. Our simulations show that Q moves along the Q tunnel in a redox-state-dependent manner, with distinct binding sites formed by conserved residue clusters. The motion of Q to these binding sites is proposed to be coupled to the proton-pumping machinery in complex I.**

NADH:ubiquinone oxidoreductase | diffusion model | electron transfer | molecular simulations | cell respiration

Complex I (NADH:ubiquinone oxidoreductase) functions as an important member of the aerobic respiratory chains of many organisms. This gigantic (0.5–1 MDa) redox-driven proton pump receives electrons from the oxidation of foodstuffs, via reduced NADH, and transfers them to quinone (Q), a reaction that is coupled to proton translocation across the membrane (1–5). The electrochemical proton gradient thus established drives synthesis of ATP and active transport (6). Structural studies (7–12) show that the electron transfer module of complex I is located exclusively in the hydrophilic domain of the protein, whereas the proton-pumping activity takes place in the antiporter-like membrane subunits (13–15), located as far as *ca*. 200 Å from the Q reduction site (Fig. 1). Despite many mechanistic suggestions (2, 4, 5, 7, 8, 10–12, 16, 17), it remains currently unclear how the coupling between the electron and proton transfer reactions is achieved across such large distances.

In contrast to most Q reductases, the Q-binding site in complex I is located ∼20–30 Å above the membrane plane, in a tight tunnel (11, 12) that ends at around 12 Å from the iron–sulfur center, N2, which serves as the immediate electron donor for Q (Fig. 1, *Inset*) (18). It has been suggested that the reduction of Q is coupled to the proton-pumping machinery in complex I (19–22). More specifically, Euro et al. (14) proposed that it is the initial negative charge on Q deposited upon reduction that triggers proton translocation via an electrostatic/conformational change mechanism. This idea was subsequently adopted and extended in the mechanisms proposed by Sazanov and coworkers (10, 11) Brandt (20), Verkhovskaya and Bloch (23), Kaila (17), and Wikström and Hummer (24, 25). Recently, we showed (26) that the reduction of Q is coupled to a local proton transfer from Tyr-$87_{Nqo4}$ and His-$38_{Nqo4}$, forming $QH_2$. This in turn triggers a charge redistribution cascade that propagates into the membrane-bound NuoH/Nqo8 subunit. It was also suggested that the coupling is mediated by conformational and electrostatic rearrangements, which lead to increased $pK_a$ values of several conserved residues in the Nqo8 subunit, thereby activating the proton pump (26).

We recently proposed that there might be two preferred Q-binding positions within the Q tunnel (4) (see also ref. 24), a low ($Q_L$) and a high ($Q_H$) potential site. The $Q_L$ site is near center N2 that was initially characterized biochemically (27–29) and later confirmed computationally (26, 30). The $Q_H$ site could be located somewhere in the Q tunnel, albeit its precise location as well as its molecular architecture remains undescribed (Fig. 1). To date, however, a bound Q molecule has not been resolved in any of the crystal or recent cryo-EM structures of complex I (11, 12, 31, 32). Two conformations have been reported for Q bound to the conserved Tyr-$87_{Nqo4}$ (30), and recent studies also show that the mammalian complex I can operate with different numbers of isoprene units, $Q_1$–$Q_{10}$ (33). Earlier reports from labeling experiments (34, 35) also support the existence of multiple

## Significance

Complex I is the primary energy-converting enzyme of aerobic respiratory chains. By reducing quinone to quinol, this gigantic enzyme pumps protons across its membrane domain, which in turn powers ATP synthesis and active transport. Despite the recently resolved molecular structures of complex I, the quinone dynamics and its coupling to the pumping function remains unclear. Here we show by large-scale molecular simulations that the quinone reduction leads to ejection of the quinol molecule from the active site into a second binding site near the proton-pumping membrane domain of complex I. The identified region has been linked with human mitochondrial disorders. Our work suggests that the quinone dynamics provides a key coupling element in complex I.

Author contributions: J.W., V.S., G.H., and V.R.I.K. designed research; J.W., V.S., A.P.G.-H., A.D.L., O.H., and V.R.I.K. performed research; J.W., V.S., A.P.G.-H., G.H., and V.R.I.K. contributed new reagents/analytic tools; J.W., V.S., A.P.G.-H., A.D.L., O.H., I.V., M.W., G.H., and V.R.I.K. analyzed data; and J.W., V.S., G.H., and V.R.I.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.



[1]J.W. and V.S. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: vivek.sharma@helsinki.fi, gerhard.hummer@biophys.mpg.de, or ville.kaila@ch.tum.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1805468115/-/DCSupplemental.

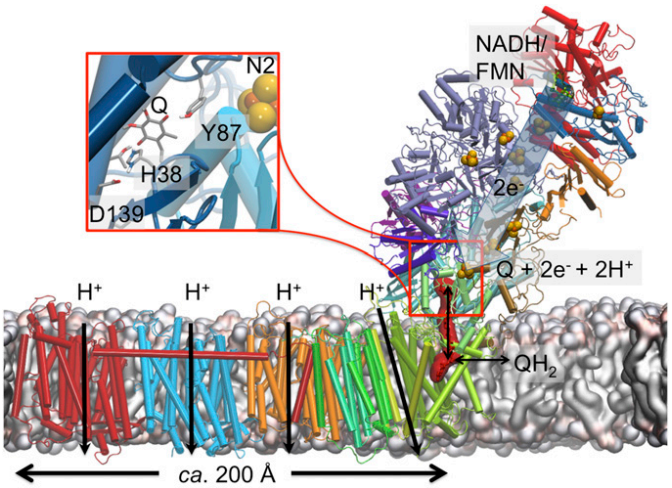


**Fig. 1.** Structure and function of complex I from *T. thermophilus* (PDB ID code 4HEA). Reduction of Q (shown in red surface representation) by electron transfer from NADH/FMN in the hydrophilic arm of complex I drives proton pumping in the membrane domain, up to 200 Å away from the site of Q reduction. (*Inset*) The Q-reduction site near the N2 center, where Q interacts with residues Tyr-$87_{Nqo4}$ and His-$38_{Nqo4}$.

Q-binding sites in complex I. Moreover, it was suggested that the Q molecule could itself act as a piston in the proton-pumping mechanism of complex I, thereby functioning as a dynamic redox transducer that shuttles between the two sites (4). This model has some mechanistic similarities to the two-stroke model by Brandt (20), in which two sequential proton-pumping steps are coupled to affinity variations of Q between high- and low-affinity sites.

Here we identify by large-scale atomistic molecular dynamics (MD) simulations and Poisson–Boltzmann (PB) continuum electrostatic calculations, as well as by free energy simulations and Bayesian diffusion models, putative Q-binding sites in complex I and show how the dynamics of Q may be coupled to the proton-pumping process in complex I. Our combined simulation data provide molecular insight into how complex I may employ the Q dynamics to serve a piston function that transduces the redox energy into a proton-pumping activity.

## Results and Discussion

**Q Dynamics near the N2 Iron–Sulfur Center.** To probe the dynamics of Q in its binding site near the N2 center, we performed atomistic MD simulations of complex I with oxidized Q ($Q_{ox}$), and reduced and doubly protonated quinol ($QH_2$). Data from 350-ns MD simulations suggest that the $Q_{ox}$ remains bound at the site near N2 (Fig. 2). In contrast, we find that the reduced quinol species ($QH_2$) moves 8–10 Å away from Tyr-$87_{Nqo4}$, when Tyr-$87_{Nqo4}$ is deprotonated and His-$38_{Nqo4}$ is neutral, corresponding to a state after electron transfer from N2 and proton transfer from His-$38_{Nqo4}$ and Tyr-$87_{Nqo4}$ (26). We also observed similar redox- and protonation-state-dependent Q dynamics in multiple short MD simulations, as well as by using independent simulation setups of complex I (*SI Appendix*, Table S1), suggesting that the obtained overall behavior of $Q_{ox}$ and $QH_2$ is robust (*SI Appendix*, Fig. S1). It has been shown earlier that $QH_2$ formation triggers a conformational change in the active site, in which the anionic Asp-$139_{Nqo4}$ dissociates from the neutral His-$38_{Nqo4}$, and the latter residue approaches the functionally important Tyr-$87_{Nqo4}$ (26, 29). Structural rearrangement in this region was also observed in the crystal structures of complex I (11, 12). These protein motions are likely factors responsible for the observed difference between the dynamics of the two neutral species, $Q_{ox}$ and $QH_2$ (Fig. 2).

To probe the relative effects of conformational and protonation changes on the dynamics of Q near the N2 site, we performed additional *ca*. 150-ns MD simulations of structures in which we replaced $Q_{ox}$ with $QH_2$, and vice versa (see also ref. 26). Starting from the $Q_{ox}$ position, $QH_2$ remained bound to the $Q_{ox}$ site (*SI Appendix*, Fig. S2). Note that in this simulation, Tyr-$87_{Nqo4}$ and His-$38_{Nqo4}$ were protonated, and the His-$38_{Nqo4}$/Asp-$139_{Nqo4}$ ion pair remained intact. Similarly, $Q_{ox}$ starting from the $QH_2$ position, with Tyr-$87_{Nqo4}$ deprotonated and His-$38_{Nqo4}$ neutral, did not relax to the position seen in the $Q_{ox}$ simulation with Tyr-$87_{Nqo4}$ and His-$38_{Nqo4}$ modeled in their protonated states (*SI Appendix*, Fig. S2). Hence, our data are consistent with the hypothesis that conformational changes in the protein linked to deprotonation of the His/Asp pair are partly responsible for the lowered affinity for $QH_2$ near Tyr-$87_{Nqo4}$.

**Free Energy of Redox-State-Dependent Q Dynamics.** To explore the energetics driving the Q dynamics that take place beyond time scales accessible by our unbiased MD simulations, we performed free energy calculations using umbrella sampling (US) (36) in combination with the weighted histogram analysis method (WHAM) (37) (*Methods*). Due to sampling problems of a long-tailed $Q_6$–$Q_{10}$, we employed a short-tailed $Q_1$ in the US/WHAM calculations, which can also function as a substrate of complex I (33).

Fig. 3 shows the resulting potentials of mean force (PMFs) for the $Q_{ox}$ and $QH_2$ motions in the Q tunnel as functions of the Tyr-$87_{Nqo4}$–Q distance. The PMF profiles suggest that the oxidized Q ($Q_{ox}$) in a membrane milieu has to surpass an activation energy barrier of *ca*. 5–8 kcal·mol$^{-1}$ to reach the site near N2 (Fig. 3), where we observe a local plateau in the free energy surface. The PMF profile also indicates that Q binds weakly relative to the membrane Q pool, with a global minimum at around 30 Å from Tyr-$87_{Nqo4}$ and a *ca*. 5 kcal·mol$^{-1}$ barrier to reach the membrane environment (Fig. 3).

The US/PMF profile for $QH_2$ shows a minimum at site 1′, *ca*. 8–10 Å from Tyr-$87_{Nqo4}$, which roughly corresponds to the equilibrium position of the $QH_2$ headgroup observed in the free MD simulations (Fig. 2 and *SI Appendix*, Fig. S1). Between 10 and 30 Å, the PMFs for both Q and $QH_2$ are quite flat, indicating relatively unhindered back-and-forth motion. Interestingly, site 1′ for $QH_2$ coincides with the crystallographically refined position of a Q analog in the structure of mitochondrial complex I from *Yarrowia lipolytica* (12). The free energy profile suggests that upon formation of $QH_2$ at the site close to N2 (site 1), the species relaxes to a new position, site 1′, 8–10 Å away from Tyr-$87_{Nqo4}$. The relaxation to the new position may have functional relevance in preventing reverse electron transfer, as the distance to the N2 center increases to >14 Å, a typical distance threshold for biological electron transfer processes (38).

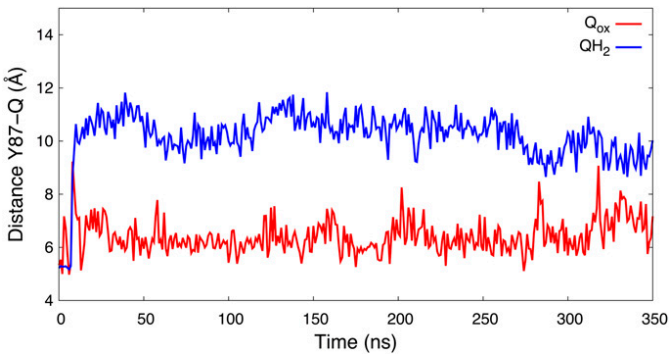


**Fig. 2.** Distance of the Q headgroup from the active site Tyr-$87_{Nqo4}$ for oxidized Q ($Q_{ox}$, in red) and reduced/protonated $QH_2$ (in blue) states obtained from 350-ns MD simulations of each state.

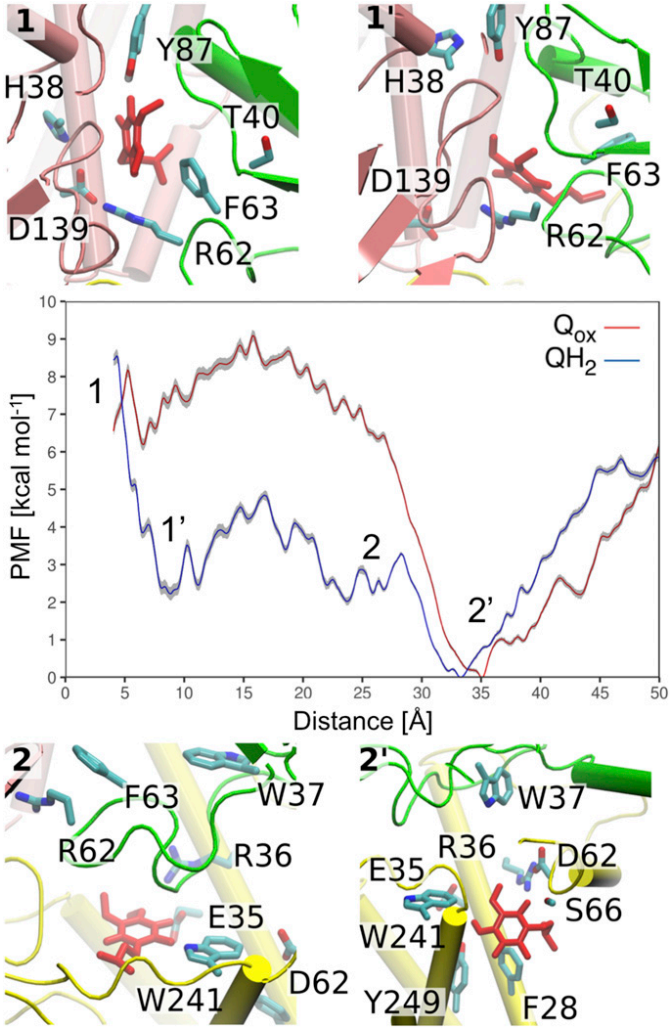


**Fig. 3.** The free energy (PMF) profiles (in kilocalories per mole) obtained from US simulations for oxidized ($Q_{ox}$, red) and reduced ($QH_2$, blue) short-tailed $Q_1$ species. The PMF shows the standard deviation of the statistical error (in gray), which was estimated by bootstrap analysis. (*Insets*) Structural snapshots corresponding to transient binding sites 1, 1′, 2, and 2′ in the PMF profile. See also *SI Appendix*, Fig. S3 for a close-up of site 2′. Overlaps in the sampled reaction coordinates are shown in *SI Appendix*, Fig. S19.

Importantly, the US/PMF calculations also suggest for the $QH_2$ species a second minimum, site 2, at *ca*. 25 Å, which is next to a conserved set of acidic residues that have been suggested to play an important role in redox-coupled proton pumping (17, 26, 39) (discussed below). We also observe a second minimum 2′ at around 33–35 Å away from Tyr-$87_{Nqo4}$ (Fig. 3) within Nqo8 at the opening to the membrane (*SI Appendix*, Fig. S3). To reach site 1 near N2 from site 2′, Q has to surpass an activation energy barrier of *ca*. 5–8 kcal·mol$^{-1}$. To further probe this putative second Q-binding site we explored the dynamics of $QH_2$ and $Q_{ox}$ using equilibrium MD simulation analyzed with a Bayesian diffusion model (40) (discussed below).

**Q Dynamics in the Tunnel and a Putative Second Q-Binding Site.** To probe how $Q_{10}$ moves within the long Q tunnel, we ran multiple equilibrium MD simulations starting from different positions of $Q_{ox}$ and $QH_2$ obtained from nonequilibrium steered MD (SMD) simulations (*Methods* and *SI Appendix*, Figs. S4 and S5). From the equilibrium simulations, we extracted the time series of the Tyr-$87_{Nqo4}$(OH)–Q distance, a reaction coordinate also used in the US simulations (*SI Appendix*, Fig. S5). In a Bayesian analysis (40) (*Methods*), we then used these time series to estimate local diffusion coefficients and free energy profiles for the $Q_{ox}$ and $QH_2$ motions along the tunnel.

The free energy profiles are shown in Fig. 4*A*. For $Q_{ox}$, we find a local minimum (site 1) at *ca*. 4 Å from Tyr-$87_{Nqo4}$ and edge-to-edge distance of *ca*. 14 Å of the Q headgroup to the iron–sulfur cluster N2. This site is stabilized further by a hydrogen bond to protonated His-$38_{Nqo4}$ (Fig. 5*B*), consistent with previous studies (30), and our equilibrium MD simulations. A shoulder in the PMF (site 1′) indicates a metastable site at a distance of ∼10–15 Å to Tyr-$87_{Nqo4}$, where the Q headgroup forms interactions with Phe-$63_{Nqo6}$ (Fig. 5*B* and *SI Appendix*, Fig. S6). We also observe a distinct second binding site (2 and 2′) at a distance of around 25–35 Å from Tyr-$87_{Nqo4}$ (Fig. 4), which coincides with the global minimum in the PMF profiles obtained by US/WHAM (Fig. 3). In site 2, the aromatic side chains of Trp-$37_{Nqo6}$, Trp-$241_{Nqo8}$, and Tyr-$249_{Nqo8}$ (Fig. 5 *B* and *C* and *SI Appendix*, Fig. S6) stabilize the Q headgroup, which is also surrounded by two conserved ion pairs, Arg-$36_{Nqo8}$/Asp-$62_{Nqo8}$ and Arg-$62_{Nqo6}$/Glu-$35_{Nqo8}$ (Fig. 5*B*). In site 2′, the Q headgroup interacts mainly with Trp-$241_{Nqo8}$, at the opposite side from site 2 (*SI Appendix*, Fig. S6), and with the backbone carbonyl group of Ser-$66_{Nqo8}$ (Fig. 5*B* and *SI Appendix*, Fig. S6). Moreover, the Q headgroup is located near the entrance of the Q tunnel, and its hydrophobic tail is in contact with the membrane milieu. Exit of Q to the membrane is hindered by a free energy barrier of >3 kcal·mol$^{-1}$, consistent with the US/WHAM profiles for the short-tailed Q. The transient binding sites obtained from our Bayesian analysis are overall similar to those obtained for the short-tailed $Q_1$ by US (Fig. 3).

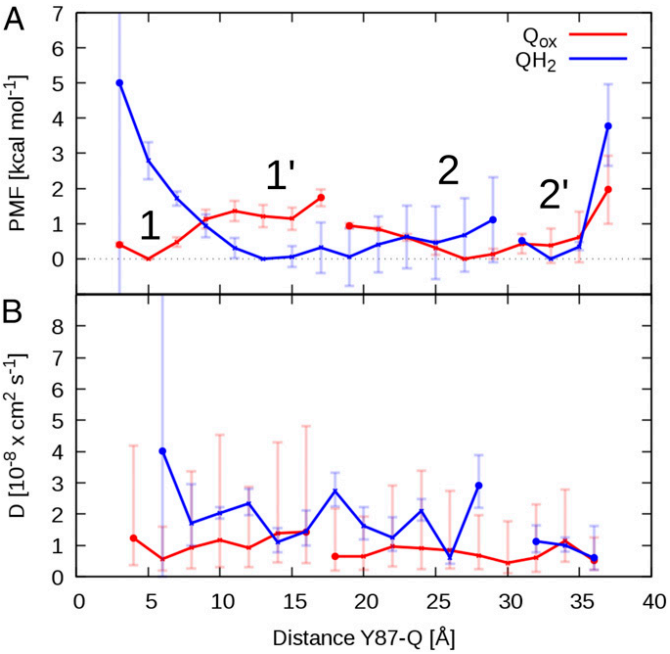


**Fig. 4.** Free energy (PMF) profile and diffusion coefficient of $Q_{ox}$ (red) and $QH_2$ (blue) in the Q tunnel. (*A*) Free energy profiles as a function of Tyr-$87_{Nqo4}$(OH)–Q distance and (*B*) corresponding position-dependent diffusion coefficient profiles extracted from unbiased equilibrium MD simulations from system setup 2 by a Bayesian analysis using a 1D diffusion model. The two PMF profiles are shifted vertically with their global minimum set to zero. In *A* 1, 1′, 2, and 2′ indicate local minima in the Q-binding cavity for the $Q_{ox}$ and $QH_2$ headgroup, respectively. Marked gaps in the profiles indicate unresolved free energy differences in rarely sampled areas of the reaction coordinate (*SI Appendix*, Fig. S5). Error bars in the PMF indicate standard errors of the mean, which were estimated by block averaging (*Methods*).

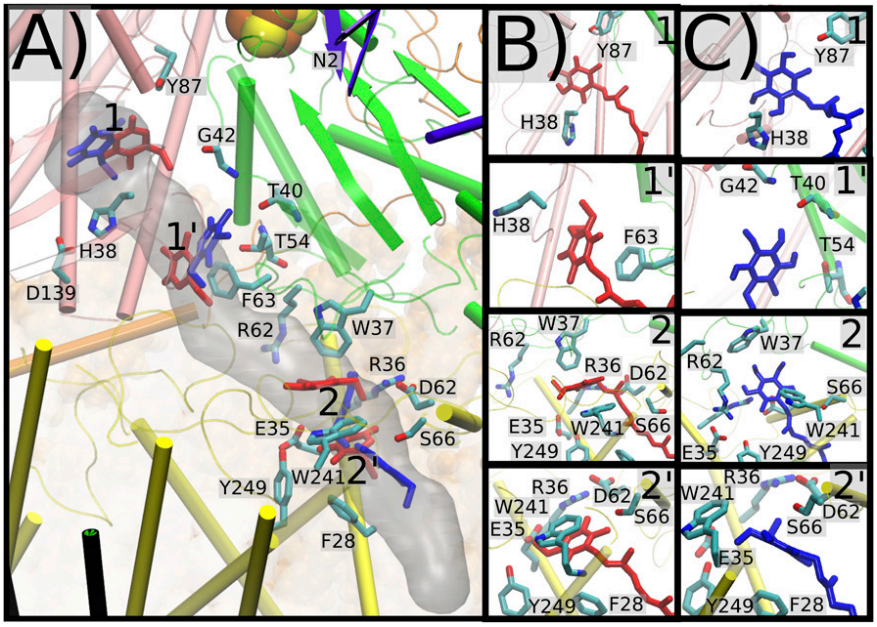


**Fig. 5.** (*A*) Structure of the Q tunnel, with $Q_{ox}$ (in red) and $QH_2$ (in blue) at transient binding sites 1, 1′, 2, and 2′ obtained from the 1D diffusion model (Fig. 4). (*B* and *C*) Close-ups of the Q-binding sites. Sites 1/1′ and 2/2′ correspond to the site close to N2 and the site at the entrance of the Q tunnel, respectively.

We also find four local minima for the $QH_2$ headgroup (Fig. 5*C*), which are near the corresponding local minima for the oxidized $Q_{ox}$ (Fig. 5*B*). Moving away from Tyr-87$_{Nqo4}$ toward the opening of the tunnel, the PMF profile for $QH_2$ smoothly decreases toward the site 1′ at *ca*. 13 Å (Fig. 4*A*). At site 1′, $QH_2$ forms hydrogen bonds with the backbone of Thr-40$_{Nqo6}$, Gly-42$_{Nqo6}$, and Thr-54$_{Nqo6}$ (Fig. 5*C*). The PMF profile for $QH_2$ is flat in the range of 10- to 28-Å distance to Tyr-87$_{Nqo4}$ (see *SI Appendix*, Fig. S5 for the projected reaction coordinate), which includes sites 1′ and 2 (Fig. 4*A*). At site 2, $QH_2$ interacts with the aromatic amino acid residues Trp-241$_{Nqo8}$, Tyr-249$_{Nqo8}$, and Trp-37$_{Nqo6}$ (*SI Appendix*, Fig. S7), similarly as for the oxidized Q. Two ion pairs, Arg-36$_{Nqo8}$/Asp-62$_{Nqo8}$ and Arg-62$_{Nqo6}$/Glu-35$_{Nqo8}$, strongly interact with the quinol headgroup in site 2 (Fig. 5*C* and *SI Appendix*, Fig. S7). The dispersion of the trajectories (*SI Appendix*, Fig. S5) indicates that sites 2 and 2′ for $QH_2$ are separated by a significant free energy barrier, which leads to a sampling gap between 28 and 32 Å of the distance reaction coordinate (see *SI Appendix*, Fig. S5 for the projected reaction coordinate). At site 2′, $QH_2$ interacts with Trp-241$_{Nqo8}$ on the side opposite minimum 2, and with the aromatic side chain of Phe-28$_{Nqo8}$. $QH_2$ is further stabilized by hydrogen bonds with the backbone carbonyl group of Ser-66 $_{Nqo8}$ (Fig. 5*C* and *SI Appendix*, Fig. S7; see also *SI Appendix*, Table S3 and sequence alignments in *SI Appendix*, Figs. S8–S10). Note that site 1, in which $QH_2$ forms a contact with Tyr-87$_{Nqo4}$ (Fig. 5*C*), is only resolved in the PMF profiles with projected reaction coordinate and is not connected to the rest of the PMF (*SI Appendix*, Fig. S5). Interestingly, apart from three MD trajectories, which end up in this site, all other trajectories move to site 1′, consistent with the equilibrium MD simulations (Fig. 2). Similarly, we observe a sampling gap for Q between sites 1′ and 2. Overall, the Q tunnel can be divided into a part connecting sites 1, 1′, and 2 and the tunnel entrance around site 2′, which is separated by a cluster of aromatic residues, including Trp-241$_{Nqo8}$, and salt bridges (Figs. 3 and 5). In the channel connecting sites 1′ and 2, the quinol can diffuse quite freely, whereas the passage between the two sections is at least partially hindered.

The PMF profiles extracted from the Bayesian models are somewhat shallower than the PMFs calculated using US/WHAM simulations, but they nevertheless show a qualitatively similar behavior. For $QH_2$, both models indicate that the quinol moves away from Tyr-87$_{Nqo4}$ to a distance of about 10 Å, to a site that coincides with the location of a Q analog in the complex I structure of *Y. lipolytica* (12), and from 10 to 25 Å both PMFs are quite flat. The steep drop in the US/WHAM PMF at about 30 Å coincides with the sampling gap in the unrestrained $QH_2$ simulations. Also for $Q_{ox}$, the PMFs exhibit similar features and local minima, with one notable exception: the US/WHAM PMF for Q shows a sharp (*ca*. 7 kcal·mol$^{-1}$) drop at 30 Å, where the diffusion model is quite flat and the unrestrained simulations (*SI Appendix*, Fig. S5) do not indicate any distinct features. As discussed above, we attribute these differences, on one hand, to local structural changes in the protein around the Q tunnel, and, on the other hand, to differences in the Q-tail length, which in the case of $Q_{10}$ is almost entirely in the membrane at a Tyr-87$_{Nqo4}$–Q distance of *ca*. 30 Å. In addition to differences in the simulation protocol, we expect significant statistical uncertainties in both profiles, in reflection of the large system and complex Q motion. Our statistical analyses and estimated error bars (*SI Appendix*, Fig. S5) suggest that the overlap in the sampled reaction coordinate during the MD simulations is good, which forms the basis for employing the diffusion model, but the relative barriers are somewhat sensitive to the employed parameters in the diffusion model (*SI Appendix*, Fig. S5 *E* and *F*). Importantly, however, the diffusion model calculations predict, consistent with the results obtained from the US/WHAM simulations, that a second Q-binding site is located 25–35 Å from Tyr-87$_{Nqo4}$.

In Fig. 4*B*, we show the position-dependent diffusion coefficients for $Q_{ox}$ and $QH_2$. We obtained values in the range of $D$ = 1–4 × 10$^{-8}$ cm$^2$·s$^{-1}$ for the back-and-forth diffusion of Q in the Q tunnel of complex I. Despite large statistical uncertainties, these values are about one order of magnitude slower than the diffusion coefficients of a lipid in a typical fluid membrane environment and of ubiquinone in vesicles (41). As a rough estimate of the Q exit time in the absence of significant barriers, we obtain $\tau_{exit} \sim (30\ \text{Å})^2/D \sim 0.01$ ms. An Arrhenius correction for a barrier of 3 kcal·mol$^{-1}$ would slow down this time to about 1 ms, which is in the range of the experimental turnover rate of complex I (4).

Three main findings emerge from the MD simulations: (*i*) that the reduced and protonated $QH_2$ moves away from Tyr-87$_{Nqo4}$ to a site occupied by a Q analog in a structure of *Y. lipolytica* complex I (12); (*ii*) that a distinct second binding site emerges close to the opening of the Q tunnel to the membrane, where the Q headgroup forms tight interactions with a cluster of aromatic and charged amino acids; and (*iii*) that the motion of $Q_{10}$ along the tunnel is relatively facile, with a diffusion coefficient lower by about a factor 10 compared with free dynamics in a membrane and relatively modest free energy barriers.

**Q Dynamics-Triggered Redox and Protonation Shifts.** We counted 183 ion pairs in subunits Nqo4 and Nqo6-Nqo10 of complex I, which may be of functional relevance for the proton-pumping process. To probe the coupling of the Q dynamics and the conformation of these ion pairs, we select seven pairs (*SI Appendix*, Table S2), which are in the vicinity of the Q-binding cavity or close to the antiporter-like subunits Nqo8/Nqo7 of the membrane domain of complex I (*SI Appendix*, Fig. S11). Interestingly, we find that the conformational state of these ion pairs depends on the Q-binding position (*SI Appendix*, Figs. S11 and S12): For example, the Asp-139$_{Nqo4}$/His-38$_{Nqo4}$ ion pair is closed in all simulations with $Q_{ox}$, and more open in the $QH_2$ state (*SI Appendix*, Fig. S11), consistent with earlier findings (26). We observe the largest difference in the opening dynamics between $Q_{ox}$ and $QH_2$, when $Q_{ox}$ is bound at the first binding site (1/1′) and $QH_2$ is bound in the second binding site (2/2′). Many of the studied ion pairs remain close in simulations with $Q_{ox}$ (*SI Appendix*, Fig. S11), whereas in the simulations with $QH_2$ these ion pairs open up, with the exception of the Asp-62$_{Nqo8}$/Arg-36$_{Nqo8}$ ion pair (*SI Appendix*, Fig. S11). Our analysis also indicates that these ion pairs dissociate when the Q moves from site 2 to 2′ in the SMD pulling simulation (*SI Appendix*, Fig. S13). The Asp-62$_{Nqo8}$/Arg-36$_{Nqo8}$ ion pair is of special interest, as it is has been found to affect the assembly and function of complex I (42).

To further probe how these dynamical changes affect the redox potential of Q in the putative second binding site near Trp-241$_{Nqo8}$, and protonation events that may couple to occupation of this site, we performed PB continuum electrostatics calculations of Q along snapshots of structures obtained from the MD simulations. We find that the movement of Q, from the site 1 to site 2, is linked to an increase in the redox potential by *ca.* 200 mV ($Q_{ox}/SQ^{\bullet/-}$ couple; *SI Appendix*, Fig. S14), which arises from differences in local protein surroundings, especially by interaction or proximity to positively charged residues (Arg-62$_{Nqo6}$, Arg-36$_{Nqo8}$, and Lys-69$_{Nqo8}$), and dissociation from the N2 center (30). Moreover, approximate electrostatic binding free energies further suggest that the motion of the $Q_{ox}/QH_2$ toward the second binding site is coupled with an energy release of *ca.* 5 kcal·mol$^{-1}$ (*SI Appendix*, Fig. S15), which could account in part for the increase in redox potential and the overall shape of the PMF profiles. Our PB calculations also suggest that Q binding at site 2 may trigger protonation changes of nearby residues His-233$_{Nqo8}$ and Asp-72$_{Nqo7}$ (*SI Appendix*, Fig. S16), an event that could be involved in the proton-pumping process (discussed below).

While some negatively (positively) charged residues remained deprotonated (protonated) through the simulation trajectories, the PB calculations suggest that certain specific residues might undergo protonation changes that are linked with the Q position. To this end, Asp-72$_{Nqo7}$ and Glu-74$_{Nqo7}$ prefer to be deprotonated when $Q_{ox}$ is in the binding site 1 and protonated when $Q_{ox}$ is approaching these residues. In contrast, Glu-130$_{Nqo8}$ is initially protonated when $Q_{ox}$ is in binding site 1, but it deprotonates when the $Q_{ox}$ approaches binding site 2/2′. We find that His-38$_{Nqo4}$ is predicted to be protonated with $Q_{ox}$ and neutral (δ- or ε- tautomer on His) with $QH_2$, but interestingly the PB calculations suggest that Tyr-87$_{Nqo4}$ would prefer to reprotonate also in the simulation trajectory with $QH_2$/His/TyrO$^-$. The reason for the latter finding might be that $QH_2$ moves rapidly to the second binding site, which is expected to increase the proton affinity of Tyr-87$_{Nqo4}$. The reprotonation of Tyr-87$_{Nqo4}$ might, however, be kinetically limited by formation of a water contact to the N-side of the membrane or by conformational changes of nearby residues (39). To this end, His-34$_{Nqo4}$ is located at the interface between the active site and the N-side, and the residue could thus act as an intermediate proton donor during reprotonation of the active site. The PB calculations suggest that His-34$_{Nqo4}$ remains deprotonated when Q is in site 1, but it protonates when the $Q_{ox}$ moves away or is reduced and protonated into $QH_2$. Moreover, consistent with our earlier work (26), Glu-35$_{Nqo8}$, Glu-248$_{Nqo8}$, Glu-223$_{Nqo8}$, and Glu-235$_{Nqo8}$ are deprotonated when $Q_{ox}$ residues in binding site 1, but they show a strong upshift in their p$K_a$ values when the $Q_{ox}$ moves approaches bindings site 2/2′ (*SI Appendix*, Fig. S16). Some of these acidic residues were recently modeled in their protonated states by studying multiple redox/protonation states of a Q molecule in the middle of the Q tunnel (39), in line with the Q-shuttle proposal (4), and were found to be conformationally flexible depending upon their protonation states. The observed p$K_a$ shifts correlate with the ion-pair dynamics (*SI Appendix*, Fig. S12). For example, the increase in the p$K_a$ of Glu-225$_{Nqo8}$ correlates well with the distance to Lys-40$_{Nqo7}$ and Arg-73$_{Nqo8}$ (*SI Appendix*, Fig. S12), and similar correlations are observed for Glu-35$_{Nqo8}$ and Glu-235$_{Nqo8}$ and their nearby ion pairs. Importantly, the predicted Q-binding sites and surrounding ion pairs can form valuable input for future site-directed mutagenesis and labeling experiments (*SI Appendix*, Table S4).

**Mechanistic Implications.** Experimental data (43) suggest that the complex I from *Escherichia coli* comprises one tightly bound Q with a ratio of 1.3 (± 0.1) per FMN molecule. However, it remains entirely unknown where the Q molecule is bound in the complex I structure. Based on the data from our PMF calculations, we suggest that Q can bind at or near the site of reduction, and at a site close to the entrance of the Q tunnel, formed by two highly tilted and one horizontal helix of the Nqo8 subunit (*SI Appendix*, Fig. S3). This second site, located at a distance of around 30 Å away from Tyr-87$_{Nqo4}$, is lined with aromatic and charged groups that offer favorable interactions for the $Q_{ox}$ and $QH_2$ headgroups. The equilibrium redox titration shows that the tightly bound Q gets reduced to quinol at potentials <−300 mV (43), which according to our calculations would correspond to the Q bound at a site *ca.* 4 Å from Tyr-87$_{Nqo4}$. At the second binding site near the tunnel opening, the calculated redox potential is substantially higher, shifted up by the interactions with nearby positively charged residues. Interestingly, such a shift in redox potential coupled to Q motion has been proposed (4) as an explanation of redox titration experiments (43).

Upon electron transfer from N2 to the Q bound at site 1, leading to the formation of semiquinone (SQ), our calculations indicate that the significant work linked to dissociating SQ could kinetically trap the latter (*SI Appendix*, Fig. S17). A similar kinetic trapping of an SQ has also been suggested in cytochrome $bc_1$ (44). Arrival of the second electron from the iron–sulfur chain leads to formation of the two-electron reduced $QH_2$ species by coupled proton transfer from His-38$_{Nqo4}$ and Tyr-87$_{Nqo4}$, followed by structural rearrangement (26). This process reduces the binding free energy of the quinol by 2–3 kcal·mol$^{-1}$, suggesting that the species is released to Q-binding site 1′. Our calculations suggest that the diffusion of quinol toward the latter site is exergonic, a process that may thus comprise a primary energy transduction step in complex I that the enzyme employs for driving the proton-pumping machinery (17). Diffusion of the $QH_2$ from site 1′ away from N2 also appears to be barrierless until it reaches site 2, found based on our free energy calculations. Our electrostatic calculations suggest that the Q bound at the second site at the entrance of the Q tunnel may alter the

protonation probability of nearby titratable residues (*SI Appendix*, Figs. S16 and S18). To this end, previous MD simulations (16, 45) suggest that the protonation states of buried residues in the membrane domain of complex I may control the formation proton-conducting water wires between the bulk.

In the standard single-Q model, Q molecules exchange with the membrane pool. Since the N2 cluster is the primary electron donor to Q, a newly arriving Q species bound initially at Q-binding site 2/2′ will need to diffuse toward the active site 1 close to N2, become reduced and travel back to site 2/2′, and finally diffuse into the membrane. We roughly estimated a minimum round-trip time in the range of tens of microseconds, in the absence of free energy barriers, which would move into the millisecond regime for barriers as low as 3 kcal·mol$^{-1}$. Indeed, the barriers in our simulations appear to be higher than this, but we caution against overinterpretation of these values, which may reflect in part the difficulties in equilibrating and sampling the motions of such a large and complex protein.

Nevertheless, it is worth noting that even under the most ideal circumstances of a large Q pool in the membrane and completely unrestricted motion of Q in the *ca*. 35-Å tunnel with a diffusion coefficient $D$, that is 1/10 of free diffusion, and a barrier of only 3 kcal·mol$^{-1}$, the round-trip time is already around 1 ms. Any further increase in the barrier height would thus limit the overall turnover. As a possible alternative, binding of a second Q has been occasionally discussed in the literature (4) (see also refs. 17 and 46). Interpreting our simulation results also in such a framework, one Q molecule would continuously shuttle between two binding sites 1 and 2, and function as a redox transducer. From site 2 (or 2′) the electrons would have to be transferred to a secondary Q in the membrane or in site 2′. To contribute to the Q pool, the quinol at the second Q-binding site, *ca*. 30 Å away from Tyr-87$_{Nqo4}$, has to further reduce a membrane-bound Q if the former is restricted to the piston motion between the two sites, as possibly supported by the obtained high free energy barrier for $QH_2$ to exit the membrane. The oxidation of quinol at this site by a membrane-bound Q requires the latter to be within *ca*. 14 Å of the former to allow for an efficient electron transfer according to biological electron transfer theories (38). A putative membrane bound site, $Q_M$ (M for membrane) on the surface of the Nqo8 subunit would have to fulfill this criterion. This site would need to have a protonic connectivity with the aqueous N-phase or to the $QH_2$ species at second Q-binding site, which is required to stabilize the negatively charged species ($Q^{\bullet/-}$ or $QH^-$) that forms upon reduction, and the coupled oxidation of quinol. The putative double-Q-piston model is consistent with some previous experimental and computational suggestions, for example the redox-state-dependent conformational fluctuation of E-quartet glutamates in subunit Nqo8 (26), conformational changes of residues in the two crystal structures, in particular Glu-213$_{Nqo8}$ (11, 12), and data from labeling studies (35), that allow us to construct a basic architecture of this putative third Q-binding site.

The possibility of multiple Q-binding sites in complex I is indirectly also supported by the recent reinterpretation of EPR data of SQ species in complex I from *E. coli* (47). In addition to already well-known fast- and slow-relaxing SQ signals ($SQ_{Nf}$ and $SQ_{Ns}$, respectively) (19, 28) that fit well with the two proposed sites, with distances of ∼10 Å and 35 Å from the N2 center, respectively, a third very-slow-relaxing SQ signal ($SQ_{Nvs}$) has been proposed (47). It was suggested that the latter signal originates from a Q bound at a membrane–protein interface (47). In the context of our model, the latter location would correspond to the $Q_M$ site, whereas $SQ_{Nf}$ and $SQ_{Ns}$ would be analogous to the sites *ca*. 4–10 Å and *ca*. 25–35 Å from Tyr-87$_{Nqo4}$, respectively.

After reoxidation of $QH_2$ bound at the second site, its protons are either released to the N- or P-sides of the membrane, transferred to the $Q_M$ molecule, or a combination of these scenarios. Moreover, reprotonation of the residues that served as initial proton donors in the Q reduction step (Tyr-87$_{Nqo4}$ and His-38$_{Nqo4}$) is required for complex I to restore its ground state for the next reaction cycle.

## Conclusions

Our molecular simulations on complex I in states that occur immediately before and after the Q reduction suggest that a single Q molecule shuttles between the hydrophilic and membrane domains of complex I within a tight tunnel. This remarkable diffusion process spans a distance of *ca*. 30 Å and is likely to be important for establishing a strong coupling between the spatially distant proton and electron transfer activities in complex I. Based on our molecular simulations, we have described here the molecular structure of a putative second Q-binding site, hoping to stimulate new site-directed mutagenesis and labeling experiments (*SI Appendix*, Table S4). The studied Q piston motion is suggested to comprise an elementary energy transduction step in complex I that is responsible for activating the long-range controlled proton-pumping machinery.

## Methods

**MD Simulations.** The crystal structure of complex I from *Thermus thermophilus* was taken from the Protein Data Bank (PDB ID code 4HEA) (11). Following the protocol described in our earlier work (26), we constructed an atomistic model system of complex I immersed in a lipid–solvent environment. The model system consisted of 809,314–823,699 atoms, including the entire complex I, POPC lipids, TIP3P water molecules, and $Na^+$ and $Cl^-$ ions, mimicking a 150 mM salt concentration. We performed simulations in various redox/protonation states of Q using $Q_1$, $Q_6$, and $Q_{10}$ (see text and *SI Appendix*, Table S1). Force-field parameters for the ubiquinone substrates and iron–sulfur centers were derived from density functional theory calculations from our previous work (26, 30, 33). Our previous calculations (30) suggest that the derived parameters reproduce experimental redox potentials. The simulations were performed with NAMD2 (48), using the CHARMM27 and CHARMM36 force fields for protein, lipids, water, and ions (49, 50) at constant temperature ($T$) and pressure ($P$), with $T$ = 310 K and $P$ = 1 atm. The time step was 1–2 fs, long-range electrostatics was treated with particle mesh Ewald method, and all hydrogen bonds are restrained by the ShakeH algorithm as implemented in NAMD. The equilibrium properties of our simulation models were monitored with different indicators (e.g., the lipid equilibration is shown in *SI Appendix*, Fig. S20). Simulation trajectories were analyzed using Visual Molecular Dynamics (51). All simulation setups are summarized in *SI Appendix*, Table S1.

**SMD Simulations.** We performed two independent SMD simulations (52, 53) of bound Q and quinol in the first binding site in hydrogen-bonding distance to Tyr-87$_{Nqo4}$, with a constant pulling velocity of 0.5 Å·ns$^{-1}$ and a force constant of 100 kcal·mol$^{-1}$·Å$^{-2}$, applied to the center of mass of Q in the direction of the exit of the binding pocket. To keep the protein and the lipid membrane stable in the box, we fixed the Cα atoms in the transmembrane helices of Nqo10 with a force constant of 2 kcal·mol$^{-1}$·Å$^{-2}$. No temperature or pressure controls were used in the SMD simulations, which were performed using NAMD2 (48). We also performed a second set of SMD simulations with higher pulling velocities, by pulling the last carbon atom (C34) of the $Q_6$ tail at a constant velocity of 5 Å·ns$^{-1}$ in various redox/protonation states of Q ($Q_{ox}$, $QH_2$, and SQ). After multiple tests, the value of the force constant ($k$) was chosen to be 8 kcal·mol$^{-1}$·Å$^{-2}$. The snapshots along the pulling trajectories were used to perform equilibrium US simulations (discussed below). Due to the slow convergence (54), we did not calculate the PMF from the SMD simulations. However, our preliminary calculations show that it is energetically costly to pull a negatively charged SQ out of the Q tunnel in comparison with neutral $Q_{ox}$ and $QH_2$ species (*SI Appendix*, Fig. S17).

**PMF from Diffusion Model.** Based on snapshots extracted every 2 ns from the SMD trajectory of Q and $QH_2$, we initiated multiple unrestrained equilibrium MD simulations starting from different positions of the $Q_{ox}$ and $QH_2$. For $Q_{ox}$, we performed initially 38 such MD simulations, each 11 ns long, and to improve the sampling in rarely visited regions of the reaction coordinate we initiated 23 additional 11-ns MD simulations. We completed the dataset of Q by adding 4 × 11-ns MD simulations in which the Q headgroup was initially in hydrogen-bonded contact with His-38$_{Nqo4}$ and Tyr-87$_{Nqo4}$, using a starting structure obtained from previous work (26). For the quinol simulations, we initially performed 38 × 11-ns MD simulations and to improve the sampling in

rarely visited regions, we initiated 14 additional 11-ns MD simulations. We completed the dataset of $QH_2$ by adding 5 × 11-ns MD simulation in which the Q headgroup was initially in hydrogen-bonded contact with His-$38_{Nqo4}$ and Tyr-$87_{Nqo4}$, starting from the same structure as the $QH_2$ SMD, but with $QH_2$ replaced by Q. From the complete 65 × 11-ns and 57 × 11-ns equilibrium MD simulations of $Q_{ox}$ and $QH_2$, respectively, the last 10 ns were used for the analysis. In addition to the Tyr-Q distance reaction coordinate, we also projected the shortest distance between Tyr-$87_{Nqo4}$ and a carbonyl oxygen of the Q/ol headgroup on the SMD pulling vector (*SI Appendix*, Fig. S4). Results for these PMF calculations are shown in Fig. 4 and *SI Appendix*, Fig. S5.

**One-Dimensional Diffusion Model.** The rate constant for the overall Q release by diffusion along the Q tunnel is exponentially sensitive to the free energy barrier heights and linearly proportional to the diffusion coefficient. The position-dependent diffusion coefficient $D(q)$ along the reaction coordinate $q$ describes the local dynamics on the free energy surface $F(q)$. Several 1D coordinates $q$ were considered to monitor the motion of Q through the Q-binding cavity, each discretized into bins $q_i$ ($0 \le i \le N$) of width $\Delta q = |q_{i+1} - q_i| = 2$ Å. Error bars in the PMF indicate standard errors of the mean, which were estimated by block averaging, dividing the trajectories into three blocks of equal length.

From the observed bin transitions in the MD trajectories, we estimated $D(q)$ and $F(q)$ self-consistently by using a 1D diffusion model as described in ref. 40. The MD simulation reports on the local "propagators" along the reaction coordinate $q$. One can compare the observed bin transitions in the MD trajectories with those expected from the diffusive dynamics. A likelihood function is constructed that gives the probability of observing exactly the motion along $q$ seen in the simulations.

The likelihood $L$ between transitions given the rate model is

$$\log L = \sum_{i,j} N_{ji} \log(p(j, \Delta t \vee i, 0)), \qquad [1]$$

where $N_{ji}$ is the number of observed transitions from state $i$ to $j$ with the lag time $\Delta t$. $N_{ji}$ was computed from MD trajectories with coordinates saved every 10 ps. The lag times of 1.5, 2, and 2.5 ns were used. $p(j, \Delta t \vee i, 0)$ is the conditional probability that the system is in the state $j$ at the time $t + \Delta t$ given that it is in the state $i$ at time $t$. The log ($L$) is maximized with a Monte Carlo search by varying $F_i$ and $D_{i \pm 1/2}$. All kinetic constants $k$ of the local propagator in the diffusion model enter into the rate matrix $K$.

The rate constant from state $i$ to state $i \pm 1$ is

$$k_{i \to i \pm 1} = \frac{D_{i \pm \frac{1}{2}}}{\Delta q^2} \exp\left[\frac{-F_{i \pm 1} - F_i}{2 k_B T}\right]. \qquad [2]$$

$F_i$ is the free energy at the center of bin $q_i$, and $D_{i \pm \frac{1}{2}} = D(q_i + q_{i \pm 1}/2)$ is the diffusion coefficient between bin $i$ and $i \pm 1$. Because the local propagator is moving on a 1D reaction coordinate, direct transitions occur only between neighboring bins. The rate matrix $K$ thus adopts the following form:

$$K_{ij} = \begin{cases} k_{i \leftarrow j} & \text{if } |i-j| = 1 \\ -k_{j \to j+1} - k_{j \to j-1} & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}. \qquad [3]$$

**US Simulations.** US simulations (36) were performed using the colvar module in NAMD2 (48). Smaller model systems of complex I consisting of *ca*. 180,000 atoms were constructed from the SMD simulation trajectories. Each model system consisted of Nqo4–Nqo10 and Nqo15 complex I subunits, immersed in lipid–solvent surroundings. US simulations were performed by constraining the distance between Tyr-$87_{Nqo4}$ and the headgroup of Q by a half-harmonic potential of 5 kcal·mol$^{-1}$·Å$^{-2}$. The distance sampled was in the range of [4 Å, 50 Å] with a 0.5-Å spacing, simulating in total 61 umbrella windows. At $d$ = 4 Å, the Q headgroup is hydrogen bonding with Tyr-$87_{Nqo4}$, whereas at a distance of 50 Å it is almost out in the lipidic milieu, such that the entire Q-tunnel region was explored.

The US simulations were performed for two states of Q ($Q_{ox}$ and $QH_2$), using a short-tailed $Q_1$. A sufficient overlap between the neighboring histograms was observed for each Q-state simulation (*SI Appendix*, Fig. S19), and convergence was analyzed by calculating the time evolution of PMF profiles. The PMF was calculated using WHAM (37), as implemented in ref. 55, with a convergence threshold of 0.00001 kcal·mol$^{-1}$, and after discarding the first 5 ns of data. The bootstrap error analysis, as implemented in ref. 55, showed a statistical uncertainty of ±0.1 kcal·mol$^{-1}$ (see also Fig. 3).

Our analysis revealed that closely spaced US windows (with shorter simulations timescales) allow better coverage of the reaction coordinate and faster convergence of PMF profiles. However, despite these and using a smaller ligand ($Q_1$), the very high complexity of the protein–substrate system results in minor oscillations in PMF profiles, thereby representing limited sampling.

**PB Continuum Electrostatics.** Redox potentials and p$K_a$ values were computed based on electrostatic potentials obtained by solving the linearized PB equation using APBS (56) and Monte Carlo titration sampling (57, 58). The protein was described using atomic partial charges, embedded in an inhomogeneous dielectric continuum with dielectric constants of $\varepsilon_p = 4$ for the protein and $\varepsilon_w = 80$ for the water, a probe radius of 1.4 Å, and ionic strength of 100 mM potassium chloride. The redox potential was computed as a difference of electrostatic free energy shifts between a model compound in water and the model compound in the protein. Charges for Q and iron–sulfur clusters as well as redox potentials of model compounds in water were obtained from ref. 30. Due to limited experimental data available for complex I, benchmarking the accuracy of the PB calculations is outside the scope of the present work. However, previous studies (59) suggest that PB calculations can reproduce experimental data in proteins with a mean error of *ca*. 1 pK unit (59) or 60 mV (60, 61). Redox potentials of iron–sulfur cluster are computationally challenging (62), but we found (30) that some experimental redox potentials for complex I are indeed reproduced within a 100- to 200-mV error bar. However, we expect the complexity of the simulation system is likely to introduce larger overall errors but nevertheless to qualitatively capture electrostatic effects linked to Q dynamics. Results from the PB calculations are shown in *SI Appendix*, Figs. S14–S16.

**ACKNOWLEDGMENTS.** V.S. thanks Dr. Giray Enkavi for helpful discussions, and J.W. thanks Dr. Jürgen Köfinger and Dr. Ahmadreza Mehdipour for helpful discussions. We thank the Center for Scientific Computing –IT Center for Science Ltd and the SuperMuc at the Leibniz Rechenzentrum (pr27xu) for computing time. We also acknowledge that part of the results of this research have been achieved using the PRACE-3IP project (FP7 RI-312763) resource Lindgren based in Sweden at Kungliga Tekniska Högskolan. This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program, Grant 715311 (to V.R.I.K.), the Max Planck Society (J.W. and G.H.), the Magnus Ehrnrooth Foundation (V.S.), and grants from Sigrid Jusélius Foundation (to M.W. and I.V.), the Academy of Finland (I.V. and V.S.), the University of Helsinki (V.S.), the Center of Excellence in Biomembrane Research, Academy of Finland (I.V. and V.S.), the German Research Foundation [V.R.I.K.; and G.H. through Grant CRC (collaborative research center) 807], the CHEMS doctoral school funding of the University of Helsinki (to O.H.), and an ERC Advanced Research Grant (to I.V.).

1. Wikström M (1984) Two protons are pumped from the mitochondrial matrix per electron transferred between NADH and ubiquinone. *FEBS Lett* 169: 300–304.
2. Brandt U (2006) Energy converting NADH:quinone oxidoreductase (complex I). *Annu Rev Biochem* 75:69–92.
3. Hirst J (2013) Mitochondrial complex I. *Annu Rev Biochem* 82:551–575.
4. Wikström M, Sharma V, Kaila VRI, Hosler JP, Hummer G (2015) New perspectives on proton pumping in cellular respiration. *Chem Rev* 115:2196–2221.
5. Sazanov LA (2015) A giant molecular proton pump: Structure and mechanism of respiratory complex I. *Nat Rev Mol Cell Biol* 16:375–388.
6. Yoshida M, Muneyuki E, Hisabori T (2001) ATP synthase–A marvellous rotary engine of the cell. *Nat Rev Mol Cell Biol* 2:669–677.
7. Hinchliffe P, Sazanov LA (2005) Organization of iron-sulfur clusters in respiratory complex I. *Science* 309:771–774.
8. Sazanov LA, Hinchliffe P (2006) Structure of the hydrophilic domain of respiratory complex I from Thermus thermophilus. *Science* 311:1430–1436.
9. Hunte C, Zickermann V, Brandt U (2010) Functional modules and structural basis of conformational coupling in mitochondrial complex I. *Science* 329: 448–451.
10. Efremov RG, Sazanov LA (2011) Structure of the membrane domain of respiratory complex I. *Nature* 476:414–420.
11. Baradaran R, Berrisford JM, Minhas GS, Sazanov LA (2013) Crystal structure of the entire respiratory complex I. *Nature* 494:443–448.
12. Zickermann V, et al. (2015) Structural biology. Mechanistic insight from the crystal structure of mitochondrial complex I. *Science* 347:44–49.
13. Amarneh B, Vik SB (2003) Mutagenesis of subunit N of the Escherichia coli complex I. Identification of the initiation codon and the sensitivity of mutants to decylubiquinone. *Biochemistry* 42:4800–4808.
14. Euro L, Belevich G, Verkhovsky MI, Wikström M, Verkhovskaya M (2008) Conserved lysine residues of the membrane subunit NuoM are involved in energy conversion by the proton-pumping NADH:ubiquinone oxidoreductase (Complex I). *Biochim Biophys Acta* 1777:1166–1172.

15. Nakamaru-Ogiso E, et al. (2010) The membrane subunit NuoL(ND5) is involved in the indirect proton pumping mechanism of Escherichia coli complex I. *J Biol Chem* 285: 39070–39078.
16. Kaila VRI, Wikström M, Hummer G (2014) Electrostatics, hydration, and proton transfer dynamics in the membrane domain of respiratory complex I. *Proc Natl Acad Sci USA* 111:6988–6993.
17. Kaila VRI (2018) Long-range proton-coupled electron transfer in biological energy conversion: Towards mechanistic understanding of respiratory complex I. *J R Soc Interface* 15:20170916.
18. Verkhovskaya ML, Belevich N, Euro L, Wikström M, Verkhovsky MI (2008) Real-time electron transfer in respiratory complex I. *Proc Natl Acad Sci USA* 105:3763–3767.
19. Ohnishi T (1998) Iron-sulfur clusters/semiquinones in complex I. *Biochim Biophys Acta* 1364:186–206.
20. Brandt U (2011) A two-state stabilization-change mechanism for proton-pumping complex I. *Biochim Biophys Acta* 1807:1364–1369.
21. Treberg JR, Brand MD (2011) A model of the proton translocation mechanism of complex I. *J Biol Chem* 286:17579–17584.
22. Nakamaru-Ogiso E, Narayanan M, Sakyiama JA (2014) Roles of semiquinone species in proton pumping mechanism by complex I. *J Bioenerg Biomembr* 46:269–277.
23. Verkhovskaya M, Bloch DA (2013) Energy-converting respiratory complex I: On the way to the molecular mechanism of the proton pump. *Int J Biochem Cell Biol* 45: 491–511.
24. Wikström M, Hummer G (2012) Stoichiometry of proton translocation by respiratory complex I and its mechanistic implications. *Proc Natl Acad Sci USA* 109:4431–4436.
25. Hummer G, Wikström M (2016) Molecular simulation and modeling of complex I. *Biochim Biophys Acta* 1857:915–921.
26. Sharma V, et al. (2015) Redox-induced activation of the proton pump in the respiratory complex I. *Proc Natl Acad Sci USA* 112:11571–11576.
27. Zickermann V, et al. (2003) Functional implications from an unexpected position of the 49-kDa subunit of NADH:ubiquinone oxidoreductase. *J Biol Chem* 278: 29072–29078.
28. Yano T, Dunham WR, Ohnishi T (2005) Characterization of the delta muH+-sensitive ubisemiquinone species (SQ(Nf)) and the interaction with cluster N2: New insight into the energy-coupled electron transfer in complex I. *Biochemistry* 44:1744–1754.
29. Tocilescu MA, et al. (2010) The role of a conserved tyrosine in the 49-kDa subunit of complex I for ubiquinone binding and reduction. *Biochim Biophys Acta* 1797:625–632.
30. Gamiz-Hernandez AP, Jussupow A, Johansson MP, Kaila VRI (2017) Terminal electron-proton transfer dynamics in the quinone reduction of respiratory complex I. *J Am Chem Soc* 139:16282–16288.
31. Zhu J, Vinothkumar KR, Hirst J (2016) Structure of mammalian respiratory complex I. *Nature* 536:354–358.
32. Fiedorczuk K, et al. (2016) Atomic structure of the entire mammalian mitochondrial complex I. *Nature* 538:406–410.
33. Fedor JG, Jones AJY, Di Luca A, Kaila VRI, Hirst J (2017) Correlating kinetic and structural data on ubiquinone binding and reduction by respiratory complex I. *Proc Natl Acad Sci USA* 114:12737–12742.
34. Nakanishi S, Abe M, Yamamoto S, Murai M, Miyoshi H (2011) Bis-THF motif of acetogenin binds to the third matrix-side loop of ND1 subunit in mitochondrial NADH-ubiquinone oxidoreductase. *Biochim Biophys Acta* 1807:1170–1176.
35. Murai M, Mashimo Y, Hirst J, Miyoshi H (2011) Exploring interactions between the 49 kDa and ND1 subunits in mitochondrial NADH-ubiquinone oxidoreductase (complex I) by photoaffinity labeling. *Biochemistry* 50:6901–6908.
36. Torrie GM, Valleau JP (1977) Nonphysical sampling distributions in Monte Carlo free-energy estimation–Umbrella sampling. *J Comput Phys* 23:187–199.
37. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J Comput Chem* 13:1011–1021.
38. Page CC, Moser CC, Chen X, Dutton PL (1999) Natural engineering principles of electron tunnelling in biological oxidation-reduction. *Nature* 402:47–52.
39. Haapanen O, Sharma V (2017) Role of water and protein dynamics in proton pumping by respiratory complex I. *Sci Rep* 7:7747.
40. Hummer G (2005) Position dependent diffusion coefficients and free energies from Bayesian analysis of equilibrium and replica molecular dynamics simulations. *New J Phys* 7:34.
41. Filippov A, Orädd G, Lindblom G (2003) Influence of cholesterol and water content on phospholipid lateral diffusion in bilayers. *Langmuir* 19:6397–6400.
42. Sinha PK, et al. (2009) Critical roles of subunit NuoH (ND1) in the assembly of peripheral subunits with the membrane domain of Escherichia coli NDH-1. *J Biol Chem* 284:9814–9823.
43. Verkhovskaya M, Wikström M (2014) Oxidoreduction properties of bound ubiquinone in complex I from Escherichia coli. *Biochim Biophys Acta* 1837:246–250.
44. Madeo J, Mihajlovic M, Lazaridis T, Gunner MR (2011) Slow dissociation of a charged ligand: Analysis of the primary quinone Q(A) site of photosynthetic bacterial reaction centers. *J Am Chem Soc* 133:17375–17385.
45. Di Luca A, Gamiz-Hernandez AP, Kaila VRI (2017) Symmetry-related proton transfer pathways in respiratory complex I. *Proc Natl Acad Sci USA* 114:E6314–E6321.
46. Haapanen O, Sharma V (2018) A modeling and simulation perspective on the mechanism and function of respiratory complex I. *Biochim Biophys Acta* 1859: 510–523.
47. Narayanan M, Leung SA, Inaba Y, Elguindy MM, Nakamaru-Ogiso E (2015) Semiquinone intermediates are involved in the energy coupling mechanism of E. coli complex I. *Biochim Biophys Acta* 1847:681–689.
48. Phillips JC, et al. (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26: 1781–1802.
49. MacKerell AD, et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102:3586–3616.
50. Klauda JB, et al. (2010) Update of the CHARMM all-atom additive force field for lipids: Validation on six lipid types. *J Phys Chem B* 114:7830–7843.
51. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. *J Mol Graph* 14:33–38, 27–28.
52. Grubmüller H, Heymann B, Tavan P (1996) Ligand binding: Molecular mechanics calculation of the streptavidin-biotin rupture force. *Science* 271:997–999.
53. Izrailev S, Stepaniants S, Balsera M, Oono Y, Schulten K (1997) Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys J* 72:1568–1581.
54. Moradi M, Tajkhorshid E (2013) Driven metadynamics: Reconstructing equilibrium free energies from driven adaptive-bias simulations. *J Phys Chem Lett* 4:1882–1887.
55. Grossfield A (2014) An implementation of WHAM: The Weighted Histogram Analysis Method, Version 2.0.9. Available at membrane.urmc.rochester.edu/content/wham. Accessed June 1, 2017.
56. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 98: 10037–10041.
57. Rabenstein B, Ullmann GM, Knapp EW (1998) Calculation of protonation patterns in proteins with structural relaxation and molecular ensembles–Application to the photosynthetic reaction center. *Eur Biophys J* 27:626–637.
58. Kieseritzky G, Knapp EW (2008) Improved pK(a) prediction: Combining empirical and semimicroscopic methods. *J Comput Chem* 29:2575–2581.
59. Meyer T, Knapp EW (2015) pKa values in proteins determined by electrostatics applied to molecular dynamics trajectories. *J Chem Theory Comput* 11:2827–2840.
60. Ishikita H, Knapp EW (2005) Control of quinone redox potentials in photosystem II: Electron transfer and photoprotection. *J Am Chem Soc* 127:14714–14720.
61. Gamiz-Hernandez AP, Kieseritzky G, Ishikita H, Knapp EW (2011) Rubredoxin function: Redox behavior from electrostatics. *J Chem Theory Comput* 7:742–752.
62. Moesca J-M, Chen JL, Noodleman L, Bashford D, Case DA (1994) Denstity-functional/Poisson-Boltzmann calculations of redox potentials for iron-sulfur clusters. *J Am Chem Soc* 116:11898–11914.

# Article VI

BIOCHEMISTRY

# How cardiolipin modulates the dynamics of respiratory complex I

**Alexander Jussupow*, Andrea Di Luca*, Ville R. I. Kaila†**



**Cardiolipin modulates the activity of membrane-bound respiratory enzymes that catalyze biological energy transduction. The respiratory complex I functions as the primary redox-driven proton pump in mitochondrial and bacterial respiratory chains, and its activity is strongly enhanced by cardiolipin. However, despite recent advances in the structural biology of complex I, cardiolipin-specific interaction mechanisms currently remain unknown. On the basis of millisecond molecular simulations, we suggest that cardiolipin binds to proton-pumping subunits of complex I and induces global conformational changes that modulate the accessibility of the quinone substrate to the enzyme. Our findings provide key information on the coupling between complex I dynamics and activity and suggest how biological membranes modulate the structure and activity of proteins.**

## INTRODUCTION

Complex I (NADH:ubiquinone oxidoreductase) functions as a redox-driven proton pump in aerobic respiratory chains (*1*–*3*). With a molecular mass of ca. 1 MDa distributed among up to 45 subunits in eukaryotes (*4*–*7*), complex I is the largest, most intricate, and least understood enzyme of the respiratory chain. The over 100-Å-long hydrophilic domain of complex I catalyzes electron transfer between nicotinamide adenine dinucleotide (NADH) and quinone, and the released free energy is used for proton pumping across the 200-Å-wide membrane domain (Fig. 1) (*8*–*10*). The established proton motive force powers energy-requiring processes in the cell in the form of active transport and synthesis of adenosine triphosphate (ATP) (*11*, *12*). The electron and proton transfer processes are fully coupled despite their large spatial separation. Although the overall mechanism of this process still remains unclear, molecular principles recently started to emerge, suggesting that the signal propagates by combined conformational, electrostatic, and hydration changes across the membrane domain (*13*–*16*).

In addition to the functional elements of the 14 core subunits (*17*), the activity of complex I strongly depends on cardiolipin (*18*–*20*), an anionic lipid that accounts for 20% of mitochondrial membranes. The activity of all mitochondrial energy-transducing enzymes depends on this lipid (*21*, *22*), but unlike, e.g., $F_oF_1$–ATPase, where the lipid-protein interactions are transient (*23*, *24*), complex I interacts with at least 10 tightly bound cardiolipin molecules that strongly modulate the enzymatic activity (*25*). Recent cryo–electron microscopy (cryo-EM) structures (*6*, *7*) suggest that some cardiolipins are likely to bind between the proton-pumping membrane subunits, but the functional role of these lipids remains puzzling, considering that they are located up to 100 Å away from the quinone reduction site. The quinone site is located ca. 20 Å above the membrane plane, at the top of a narrow tunnel with an opening to the lipid membrane (Fig. 1).

To elucidate a molecular mechanism by which cardiolipin modulates the activity of complex I, we study here the dynamics of the bacterial enzyme in cardiolipin-containing membranes on millisecond time scales using a combination of coarse-grained and atomistic molecular dynamics (MD) simulation approaches (Fig. 1, inset).

Department Chemie, Technische Universität München (TUM), Lichtenbergstraße 4, D-85747 Garching, Germany.
*These authors contributed equally to this work.
†Corresponding author. Email: ville.kaila@ch.tum.de

## RESULTS

### Millisecond simulations identify cardiolipin binding sites

To probe the effect of cardiolipin on the structure and dynamics of complex I, we performed coarse-grained molecular dynamics (CG-MD) simulations on millisecond time scales with the bacterial complex I embedded in a lipid membrane containing 20% cardiolipin and compared these to simulations of complex I in POPC (palmitoyl-oleoyl phosphatidylcholine):POPE (palmitoyl-oleoyl phosphatidylethanolamine) membrane without cardiolipin (see Methods; table S1). The coarse-grained simulations suggest that complex I forms contacts with ca. 150 lipid molecules during the simulations (Fig. 2, A and B). However, in stark contrast to POPE and POPC (see Methods) that transiently interact with the protein, cardiolipin molecules form much stronger interactions with the membrane domain of complex I (Fig. 2C and movie S1). The Nqo8 subunit, which provides a hinge between the membrane domain and the hydrophilic domain, strongly interacts with 10 to 15 cardiolipin molecules (Fig. 2, A to C) that are also supported by atomistic MD simulation of the complex I hinge region (see Methods; fig. S1). We also observe prominent binding sites in the CG-MD models at the interface between the antiporter-like subunits, Nqo12/Nqo13 and Nqo13/Nqo14 (Fig. 2A and fig. S2). These sites are close to the broken transmembrane (TM) helix elements, TM12, which are involved in establishing proton channels across the membrane (*15*). The identified binding sites superimpose with cardiolipin molecules recently resolved from cryo-EM structure of mammalian complex I (Fig. 2A) (*6*, *7*), suggesting that the cardiolipin interaction sites are conserved across different species, as in the other respiratory enzymes (*22*).

The cardiolipin binding is induced by positively charged surface regions distributed around the membrane domain of complex I (Fig. 2D) and conserved across different complex I isoforms (fig. S3). In addition, phenylalanine residues interact with cardiolipin, particularly in the Nqo8 subunit (Fig. 2, D and E). Notably, phenylalanine residues provide important cardiolipin binding sites also in cytochrome c oxidase (*21*).

### Cardiolipin binding affects complex I structure and dynamics

To analyze how cardiolipin affects the global dynamics of complex I, we projected out thermal fluctuations from the CG-MD simulation trajectories using principal components (PC) analysis (*26*). The analysis indicates that complex I undergoes large-scale bending (PC1) and twisting (PC2) motion around the hydrophilic and membrane

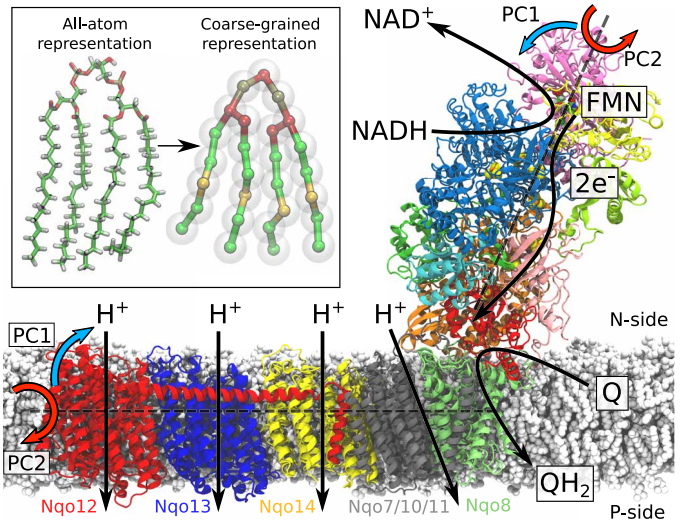


**Fig. 1. Structure and function of the respiratory complex I (PDB ID: 4HEA).** Electrons are transferred from the NADH/FMN (flavin mononucleotide) site to quinone (Q) via a tunneling wire comprising eight iron-sulfur centers. The Q reduction takes place at the interface between the hydrophilic and membrane domains and couples to the pumping of four protons across the membrane. The global bending (PC1, blue arrows) and twisting motions (PC2, red arrows) are shown along two principal axes (see text). Inset: Comparison of atomistic and coarse-grained molecular simulation models of a cardiolipin molecule.

domains with Nqo8 acting as a hinge (Fig. 1), consistent with global motions identified from previous network models and atomistic simulations of the bacterial complex I in a POPC membrane (*27*). The PC analysis suggests that the coupled twisting-bending motions (PC3) around the hydrophilic and membrane domains is strongly modulated by cardiolipin (Fig. 3A and fig. S4). This global motion displaces half of the membrane and hydrophilic domains (Fig. 3, A and B), and its origin can be traced to cardiolipin bound at the Nqo13/Nqo14 interface and at Nqo8 that induce a 0.4-nm structural shift between TM1 of Nqo8 and helix 1 of Nqo9 (Fig. 3C). These structural rearrangements may have important functional implications as they modulate the accessibility of the quinone to its binding site (see below).

## Cardiolipin modulates quinone access and channel formation dynamics

To probe how the quinone dynamics is affected by cardiolipin, we studied the global motions of complex I with and without a quinone molecule inside the quinone tunnel (see Methods; table S1). The quinone finds stable binding sites in the CG-MD simulations at the top ($d$ ~ 0.5 to 0.7 nm), slightly below ($d$ ~ 0.7 to 0.9 nm), and at the bottom part ($d$ ~ 3.2 to 3.7 nm from $Tyr^{87}$) of the cavity (Fig. 4, A and B). These binding sites are consistent with minima observed in free energy profiles of the quinone diffusion along the tunnel derived from a recent atomistic study of complex I in a POPC membrane on nanosecond time scales (*28*). We find that the enzyme undergoes a structural rearrangement in the CG-MD simulations along the bending mode when the quinone molecule moves from the lower edge of the tunnel opening ($d$ ~ 3.2 to 4 nm in Fig. 4, A to C) to the upper binding site ($d$ ~ 0.6 to 0.8 nm in Fig. 4, A to C). When the quinone passes a strongly bent kink region ($d$ ~ 2 nm; Fig. 4B) (*29*) in the middle of the tunnel, we observe a structural rearrangement along the twisting mode that facilitates the substrate motions along its tunnel, indicated by a transition around $d$ ~ 2 nm in Fig. 4A. In contrast, when the quinone is located at the lower edge of the tunnel around $d$ ~ 3.2 to 4 nm, complex I samples conformations similar to those observed in simulations of the unbound (*apo*) state (Fig. 4A), suggesting that the initial quinone binding to complex I does not trigger large-scale conformational changes in the protein structure.

The activity of the mammalian complex I is regulated by the so-called active (A)–to–deactive (D) transition (*30*) that is also characterized by structural changes along these twisting and bending modes (*5*, *7*, *27*). Since we observe a strong correlation between the quinone position in the tunnel and these global motions (Fig. 4 and fig. S5), our findings suggest that quinone dynamics could also be modulated by the structural changes required for the A-D transition in the eukaryotic complex I.

In the absence of cardiolipin in the membrane, the quinone rarely visits the upper binding site, but it finds more stable binding poses at the edge of the tunnel ($d$ ~ 2.7 to 3.6 nm; Fig. 4A). We find that complex I also undergoes similar global twisting and bending motions around the hydrophilic and membrane domains as in the cardiolipin-bound state (fig. S4A). However, in stark contrast to the latter, complex I now samples a configurational space that is very similar to that in the different quinone-bound states as well as that in the unbound (*apo*) state of the enzyme, with large fluctuations along the twisting and bending modes throughout the complete distance range ($d$ ~ 0.6 to 4.2 nm; Fig. 4A).

This suggests that the twisting and bending motions are now decoupled from the quinone diffusion, indicating that cardiolipin alters the free energy landscape of the enzyme (Fig. 4A). In other words, the bound cardiolipin molecules seem to direct complex I to sample conformations that favor the quinone motion along its cavity and stabilize quinone binding. These findings thus provide a possible molecular explanation to how cardiolipin enhances complex I activity.

The bound cardiolipin enhances the stability of the quinone tunnel (Fig. 4D), in addition to two prominent cavities that are large enough for water molecules to enter from the negatively charged side of the membrane (N-side) to the quinone tunnel. Coarse-grained simulations usually have a tendency to overestimate protein-protein interactions (*31*), and they may therefore underestimate the probability of finding such channels.

One of these putative channels comprises several conserved residues (table S2) and leads from the N-side to the conserved $His^{38}$ and $Tyr^{87}$ residues of Nqo4 (Fig. 4C and movie S2), which function as proton donors upon reduction of quinone to quinol ($QH_2$) (*32*). Reprotonation of these residues is thus a prerequisite for complex I to initiate the next catalytic cycle. This cavity is close to a loop region in Nqo7 that also undergoes conformational changes during the A-D transition in the mammalian complex I (*5*). It is important to note that a functional redox-driven pump requires tight gating of the protonation reactions to avoid leaks. A continuous water connectivity between the quinone reduction site and the N-side could thus compromise the proton-pumping machinery. To this end, we find that the stability of this putative channel depends on the position of the quinone in the tunnel (Fig. 4D), suggesting that opening of the channel could be controlled by the redox chemistry of the quinone site.

The second channel forms at the Nqo8/Nqo9 interface near TM1 (Fig. 4D). This channel also starts at the N-side of the membrane, but it leads to a unique kink region in the middle of the quinone tunnel. This region coincides with the putative second quinone binding site around $d$ ~ 3.2 to 4.0 nm (Fig. 4C), to which the quinone

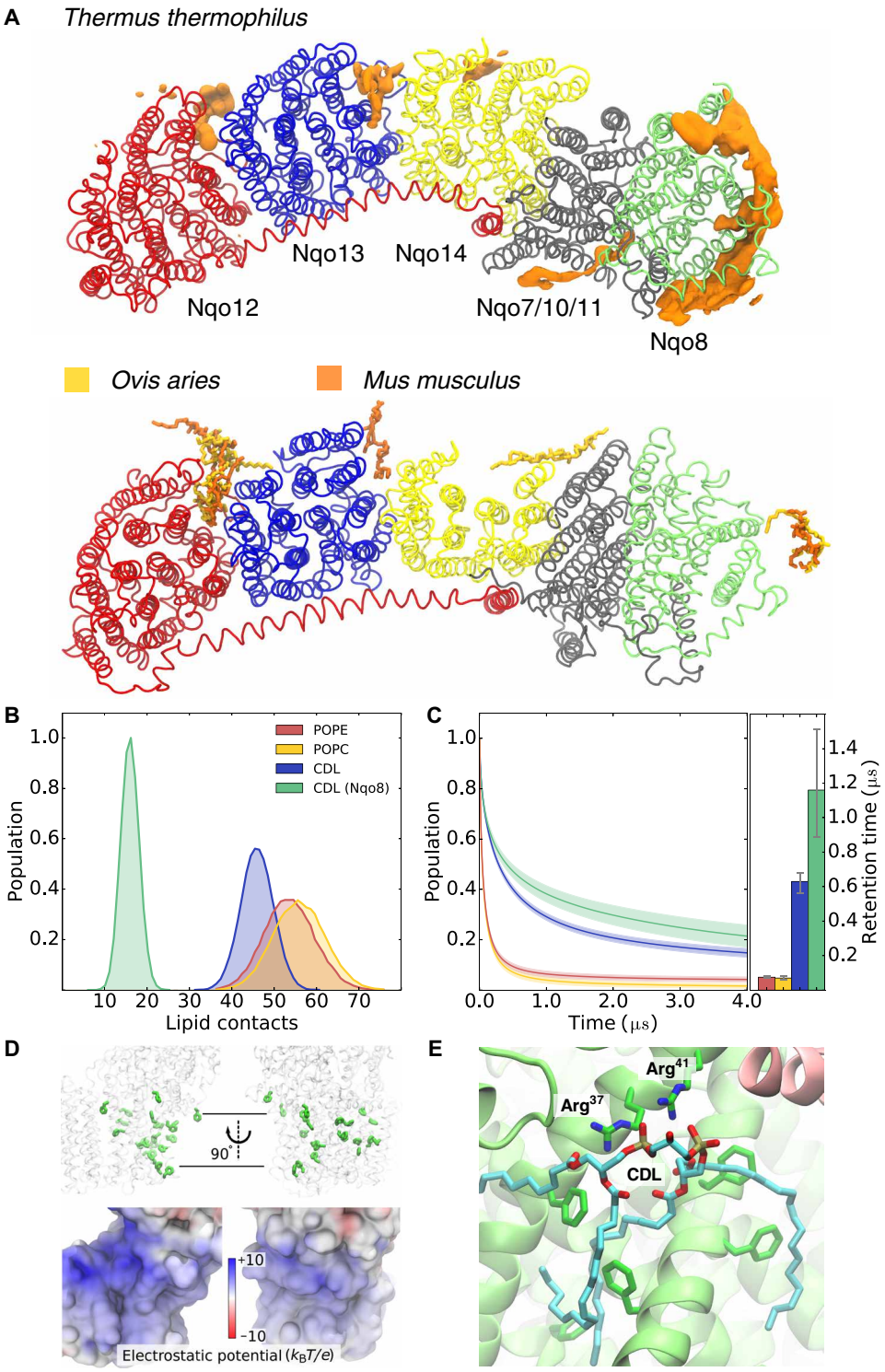


**Fig. 2. Cardiolipin binding sites in complex I.** (**A**) Top: Cardiolipin (CDL) binding sites identified by CG-MD simulations of complex I from *Thermus thermophilus*. Bottom: Experimentally refined CDL molecules around cryo-EM structures of complex I from *Ovis aries* (*6*) (PDB ID: 5LNK) and *Mus musculus* (*7*) (PDB ID: 6G2J). Note that TM helices 1 to 3 of subunit Nqo14 (subunit in yellow, *T. thermophilus*) are not present in the homologous mammalian ND2 subunits. Both enzymes are viewed from the N-side of the membrane. The supernumerary subunits are omitted for visual clarity. (**B**) Statistics of lipid contacts with the membrane domain of complex I obtained from CG-MD simulations. (**C**) Autocorrelation function of the lipid binding dynamics. The average retention time for CDL is about six times higher than that for POPE or POPC, and ca. 12 times higher than that for Nqo8-bound CDL. (**D**) Top: Membrane-exposed phenylalanine residues in Nqo8. Bottom: Electrostatic potential map of complex I (front view). The color scale ranges from +10 $k_BT/e$ (blue, ca. +260 mV) to −10 $k_BT/e$ (red, ca. −260 mV). (**E**) Putative CDL binding sites in Nqo8, obtained from atomistic simulations, highlighting interactions with $Arg^{37}$, $Arg^{41}$, and phenylalanine residues.

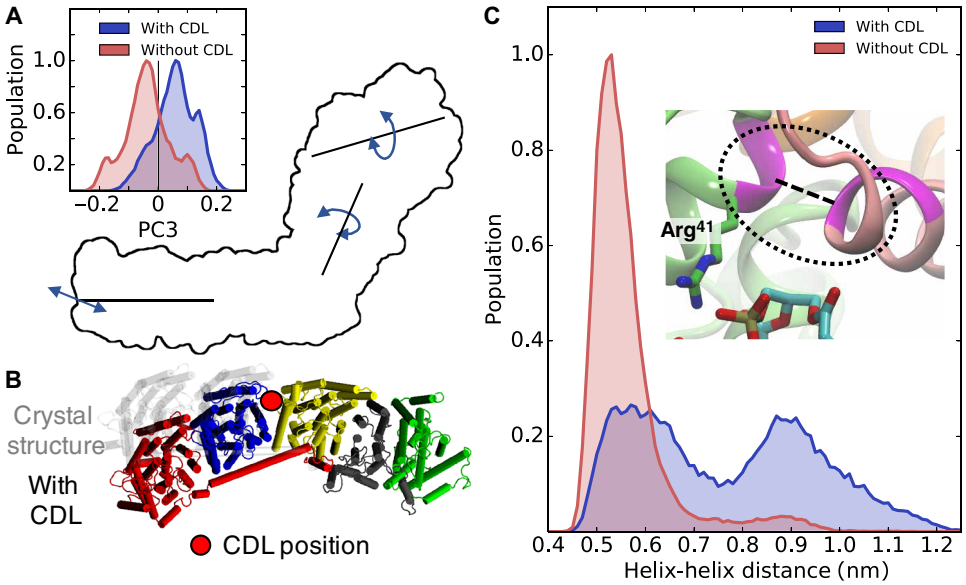


**Fig. 3. Effect of CDL on the dynamics of complex I.** (**A**) The coupled twisting-bending motion (PC3) is strongly influenced by CDL binding to complex I. The crystal structure is taken as reference conformation for the PC. (**B**) Structural changes in the membrane domain (view from the N-side) induced by cardiolipin binding at the Nqo13/Nqo14 interface. (**C**) Cardiolipin induces conformational changes at the interface between Nqo8 and the hydrophilic domain. The figure shows the distance between helix TM1 of Nqo8 and helix 1 of Nqo9.

molecule docks before exiting to the membrane [see above, and cf. also (*28*)]. The kink site could be relevant for the activation of the proton-pumping machinery in Nqo8, as this site involves many conserved residues that have been linked with human mitochondrial disorders. In simulations without cardiolipin, this channel remains strictly closed, whereas both cardiolipin and quinone bound to the active site strongly enhance opening of this channel (ca. 20 to 30% open channels with quinone and cardiolipin relative to 0.4 to 3% without cardiolipin; Fig. 4D).

## DISCUSSION

We have identified here, on the basis of millisecond molecular simulations, how cardiolipin modulates the structure and dynamics of respiratory complex I. Our simulations indicate that cardiolipin binds at specific sites in the membrane domain of complex I, in particular around the Nqo8 subunit and at the interface between antiporter-like subunits Nqo12/Nqo13 and Nqo13/Nqo14. Some of the predicted cardiolipin binding sites are supported by recent cryo-EM structures of the mammalian enzyme (*6*, *7*), whereas others remain to be experimentally validated. Our findings suggest that in addition to the conserved structural core features of complex I, the lipid binding sites are also conserved across different species.

Cardiolipin binding arises from a combination of electrostatic effects between the cardiolipin headgroup and the positively charged protein residues, as well as by dispersive interactions that are dominated by phenylalanine residues. The lipid binding modulates intersubunit contacts, which are important for the protonation signal propagation across the membrane domain (*16*). This suggests that cardiolipin binding could indirectly also affect the energetics of the proton pumping process. Cardiolipin molecules bound at Nqo8 might also stabilize the highly tilted TM helices of the subunit and affect the energetics of conformational changes linked with the protonation dynamics in this region.

In addition to these local conformational changes, our simulations suggested that complex I also undergoes global conformational changes along the bending and twisting modes that are strongly influenced by cardiolipin. These global motions were recently linked with the active-to-deactive transition (*27*). Our simulations indicate that bacterial complex I undergoes displacement along a similar structural motion when the quinone molecule moves along its tunnel. Recent studies suggest that the bacterial enzyme could also undergo a fast resting-to-active transition (*33*–*35*). In the absence of cardiolipin, the twisting and bending motions are decoupled from the quinone position along its tunnel, resulting in a lower accessibility of the quinone to its binding site. These findings might provide a molecular explanation to the regulatory role of cardiolipin in complex I: The lipid interactions modify the free energy landscape of complex I along the bending and twisting motions, which in turn modulate the quinone dynamics. Although detailed free energy calculations and atomistic simulations will be important to elucidate structural details of this effect, differences between the simulations with and without cardiolipin are systematic and robust and support that cardiolipin plays a central role in the regulation of the complex I dynamics.

We found that cardiolipin also strongly modulates the formation of two channels from the N-side of the membrane, one around the Nqo7 loop region that could provide protonation pathways for the quinone reduction site, and another channel at the Nqo8/Nqo9 interface that leads to a kink site of the quinone cavity, which could be functionally relevant for activating the pumping machinery (*28*). The functional role of these putative channels can be probed by site-directed mutagenesis experiments, to which our simulations provide important possible candidates (table S2).

In conclusion, our simulations identified how the lipid membrane may modulate global enzyme dynamics that in turn regulate the accessibility of substrates to the active site. These results provide a molecular understanding of complex I function and the role of cardiolipin in the regulation of this respiratory enzyme.

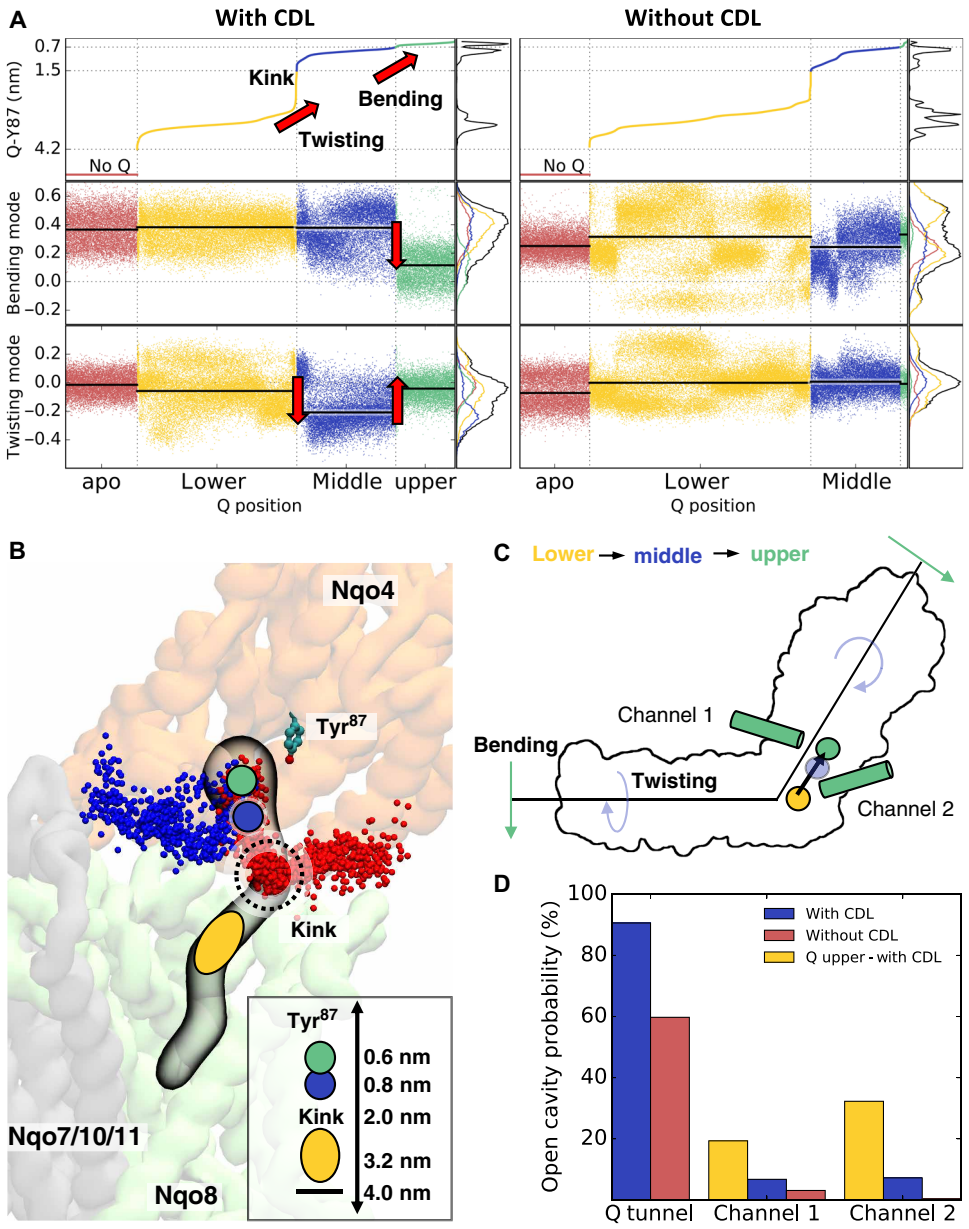


**Fig. 4. CDL modulates quinone and complex I dynamics.** (**A**) Projection of complex I dynamics with and without CDL along the bending (PC1) and twisting (PC2) modes sorted by the quinone (Q) position along the tunnel. The quinone position is determined by the distance between the quinone headgroup and Tyr$^{87}$. The quinone motion along the cavity is coupled to the complex I motion along the bending and twisting modes. (**B**) Quinone binding sites and channels connecting the quinone cavity with the N-side of the membrane. The upper binding sites ($d$ ~ 0.6 to 0.8 nm from Tyr$^{87}$) are shown as green and blue circles, the kink region is marked at $d$ ~ 2.0 nm, and the lower binding sites ($d$ ~ 3.2 to 4.0 nm) are shown as a yellow oval. Channel 1 (blue) is located near the Nqo7 loop region, and channel 2 (in red) is located at the interface between Nqo8 and Nqo9. Both channels are more stable in simulations with CDL and when quinone is bound at the upper site ($d$ ~ 0.6 to 0.8 nm) (see also movie S2). Inset: Distance between Tyr$^{87}$ and the kink region and different quinone binding sites. (**C**) Schematic representation of structural changes in complex I that couple to the quinone motion. (**D**) Probability of finding open cavities connected to the quinone binding site.

## METHODS

### General methodological overview

CG-MD simulations were used to study the interaction between cardiolipin and complex I up to millisecond time scales. The CG-MD trajectories were analyzed by PC analysis (*26*) and protein cavity search algorithms (*36*) to connect the large-scale motion with the quinone dynamics and the formation of channels leading to the quinone cavity. Cardiolipin binding around the Nqo8 region was validated by atomistic MD simulations to refine the molecular details of the cardiolipin-protein interactions. All CG-MD results were compared with global motions obtained from previous atomistic simulations and network models (*27*), as well as to atomistic simulations of the quinone dynamics (*28*).

### CG-MD simulations

Coarse-grained simulation models of complex I were built from the x-ray crystal structure from *Thermus thermophilus* (PDB ID: 4HEA) (*13*).

The protein was embedded in cardiolipin:POPE:POPC and POPE:POPC membranes with lipid ratios of 1:2:2 and 1:1, respectively, solvated using MARTINI water (*37*) and neutralized. The cardiolipin was modeled in its singly pronated state (Fig. 1B). CG-MD simulations were performed at $T = 300$ K using an $NpT$ ensemble with a 20-fs time step, a thermostat with a coupling constant of $\tau_t = 1.0$ ps (*38*), a semi-isotropic Parrinello-Rahman barostat (*39*) with a coupling constant of $\tau_p = 12.0$ ps, and a compressibility of $\chi = 3.0 \times 10^{-4}$ bar$^{-1}$. Nonbonded interactions were treated with a dielectric constant of 15 and using a cutoff distance of 1.1 nm. The CG-MD simulations were performed in total for ca. 1.2 ms, with each individual trajectory at least 50 μs long. Simulation details are summarized in table S1. All CG-MD simulations were performed using Gromacs 2016.2 (*40*) and the MARTINI 2.2 force field (*37*, *41*), and quinone parameters were obtained from the literature (*42*). Visual molecular dynamics (VMD) (*43*), Gromacs select tool, and PLUMED (*44*) were used for analysis and/or visualization, and the Backward script (*45*) was used to convert coarse-grained structures back to an atomistic representation. PyEMMA (*46*) was used to perform the PC analysis (*26*) on the coarse-grained simulation trajectories on the backbone atom level. Caver (*36*) was used with a probe radius of 2.25 Å to identify the channels that are shown in fig. S4.

### Atomistic molecular simulation

Molecular models of the membrane domain of complex I were built from the x-ray crystal structure from *T. thermophilus* (PDB ID: 4HEA) (*13*). The system comprised subunits Nqo4/Nqo6/Nqo7, Nqo8 to Nqo11, Nqo14, and the C-terminal end of the HL helix (residues 569 to 605). The model was embedded in a lipid membrane with a cardiolipin:POPC:POPE ratio of 1:2:2 built using the CHARMM-GUI module (*47*), and solvated with TIP3P (*48*) water molecules. Sodium and chloride ions were added to neutralize the system with an ionic strength of ca. 100 mM. The total system comprised ca. 332,100 atoms. After minimization, relaxation, and equilibration using harmonic forces of 2 kcal mol$^{-1}$ Å$^{-2}$ on the protein backbone atoms, MD simulations without restraints were carried out for 0.5 μs at $T = 310$ K using a 2-fs time step, and the long-range electrostatics were treated using the Particle Mesh Ewald approach (*49*). The system was treated using the CHARMM36 force field (*50*) and the density functional theory–derived parameters for the redox-active cofactors. Classical MD simulations were performed with NAMD2 (*51*), and VMD (*43*) was used for visualization and analysis. The electrostatic potential calculations were performed using the Adaptive Poisson-Boltzmann Solver (*52*).

## SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/5/3/eaav1850/DC1

Fig. S1. Comparison of cardiolipin molecules bound to Nqo8 from coarse-grained and atomistic models.
Fig. S2. The structure of the membrane domain of complex I from *T. thermophilus*.
Fig. S3. Electrostatic potential surface of the bacterial and mammalian complex I.
Fig. S4. Global motion of complex I with and without cardiolipin.
Fig. S5. Projection of complex I dynamics with and without cardiolipin along the bending (PC1) and twisting (PC2) modes.
Table S1. Overview of all coarse-grained MD simulations.
Table S2. Residues within 3 Å of putative channels 1 and 2.
Movie S1. Cardiolipin binding to complex I from a coarse-grained simulation trajectory.
Movie S2. Cardiolipin-induced channel formation dynamics.

## REFERENCES AND NOTES


1. U. Brandt, Energy converting NADH: Quinone oxidoreductase (complex I). *Annu. Rev. Biochem.* **75**, 69–92 (2006).
2. J. Hirst, Mitochondrial complex I. *Annu. Rev. Biochem.* **82**, 551–575 (2013).
3. L. A. Sazanov, A giant molecular proton pump: Structure and mechanism of respiratory complex I. *Nat. Rev. Mol. Cell Biol.* **16**, 375–388 (2015).
4. 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**, 44–49 (2015).
5. J. Zhu, K. R. Vinothkumar, J. Hirst, Structure of mammalian respiratory complex I. *Nature* **536**, 354–358 (2016).
6. 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**, 406–410 (2016).
7. A. 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**, 548–556 (2018).
8. A. J. Y. Jones, J. N. Blaza, F. Varghese, J. Hirst, Respiratory complex I in *Bos taurus* and *Paracoccus denitrificans* pumps four protons across the membrane for every NADH oxidized. *J. Biol. Chem.* **292**, 4987–4995 (2017).
9. M. Wikström, G. Hummer, Stoichiometry of proton translocation by respiratory complex I and its mechanistic implications. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 4431–4436 (2012).
10. A. S. Galkin, V. G. Grivennikova, A. D. Vinogradov, $\rightarrow H^+/2e$ stoichiometry in NADH-quinone reductase reactions catalyzed by bovine heart submitochondrial particles. *FEBS Lett.* **451**, 157–161 (1999).
11. P. Mitchell, Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* **191**, 144–148 (1961).
12. M. Yoshida, E. Muneyuki, T. Hisabori, ATP synthase — A marvellous rotary engine of the cell. *Nat. Rev. Mol. Cell Biol.* **2**, 669–677 (2001).
13. R. Baradaran, J. M. Berrisford, G. S. Minhas, L. A. Sazanov, Crystal structure of the entire respiratory complex I. *Nature* **494**, 443–448 (2013).
14. V. R. I. Kaila, Long-range proton-coupled electron transfer in biological energy conversion: Towards mechanistic understanding of respiratory complex I. *J. R. Soc. Interface* **15**, 20170916 (2018).
15. 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**, E6314–E6321 (2017).
16. A. Di Luca, M. E. Mühlbauer, P. Saura, V. R. I. Kaila, How inter-subunit contacts in the membrane domain of complex I affect proton transfer energetics. *Biochim. Biophys. Acta Bioenerg.* **1859**, 734–741 (2018).
17. J. Hirst, J. Carroll, I. M. Fearnley, R. J. Shannon, J. E. Walker, The nuclear encoded subunits of complex I from bovine heart mitochondria. *Biochim. Biophys. Acta Bioenerg.* **1604**, 135–150 (2003).
18. M. Fry, D. E. Green, Cardiolipin requirement for electron transfer in complex I and III of the mitochondrial respiratory chain. *J. Biol. Chem.* **256**, 1874–1880 (1981).
19. S. Dröse, K. Zwicker, U. Brandt, Full recovery of the NADH: Ubiquinone activity of complex I (NADH:ubiquinone oxidoreductase) from *Yarrowia lipolytica* by the addition of phospholipids. *Biochim. Biophys. Acta Bioenerg.* **1556**, 65–72 (2002).
20. M. Vos, A. Geens, C. Böhm, L. Deaulmerie, J. Swerts, M. Rossi, K. Craessaerts, E. P. Leites, P. Seibler, A. Rakovic, T. Lohnau, B. de Strooper, S.-M. Fendt, V. A. Morais, C. Klein, P. Verstreken, Cardiolipin promotes electron transport between ubiquinone and complex I to rescue *PINK1* deficiency. *J. Cell Biol.* **216**, 695–708 (2017).
21. C. Arnarez, S. J. Marrink, X. Periole, Identification of cardiolipin binding sites on cytochrome *c* oxidase at the entrance of proton channels. *Sci. Rep.* **3**, 1263 (2013).
22. C. Arnarez, J.-P. Mazat, J. Elezgaray, S.-J. Marrink, X. Periole, Evidence for cardiolipin binding sites on the membrane-exposed surface of the cytochrome *bc*1. *J. Am. Chem. Soc.* **135**, 3112–3120 (2013).
23. A. R. Mehdipour, G. Hummer, Cardiolipin puts the seal on ATP synthase. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 8568–8570 (2016).
24. A. L. Duncan, A. J. Robinson, J. E. Walker, Cardiolipin binds selectively but transiently to conserved lysine residues in the rotor of metazoan ATP synthases. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 8687–8692 (2016).
25. M. S. Sharpley, R. J. Shannon, F. Draghi, J. Hirst, Interactions between phospholipids and NADH: Ubiquinone oxidoreductase (complex I) from bovine mitochondria. *Biochemistry* **45**, 241–248 (2006).
26. C. C. David, D. J. Jacobs, Principal component analysis: A method for determining the essential dynamics of proteins. *Methods Mol. Biol.* **1084**, 193–226 (2014).
27. A. Di Luca, V. R. I. Kaila, Global collective motions in the mammalian and bacterial respiratory complex I. *Biochim. Biophys. Acta Bioenerg.* **1859**, 326–332 (2018).
28. J. Warnau, V. Sharma, A. P. Gamiz-Hernandez, A. Di Luca, O. Haapanen, I. Vattulainen, M. Wikström, G. Hummer, V. R. I. Kaila, Redox-Coupled Quinone Dynamics in the Respiratory Complex I. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E8413–E8420 (2018).

29. J. G. Fedor, A. J. Y. Jones, A. Di 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**, 12737–12742 (2017).
30. M. Babot, A. Birch, P. Labarbuta, A. Galkin, Characterisation of the active/de-active transition of mitochondrial complex I. *Biochim. Biophys. Acta Bioenerg.* **1837**, 1083–1092 (2014).
31. M. Javanainen, H. Martinez-Seara, I. Vattulainen, Excessive aggregation of membrane proteins in the Martini model. *PLOS ONE* **12**, e0187936 (2017).
32. 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**, 11571–11576 (2015).
33. N. Belevich, C. von Ballmoos, M. Verkhovskaya, Activation of proton translocation by respiratory complex I. *Biochemistry* **56**, 5691–5697 (2017).
34. N. Belevich, M. Verkhovskaya, Resting state of respiratory complex I from *Escherichia coli*. *FEBS Lett.* **590**, 1570–1575 (2016).
35. N. Belevich, G. Belevich, Z. Chen, S. C. Sinha, M. Verkhovskaya, Activation of respiratory Complex I from Escherichia coli studied by fluorescent probes. *Heliyon* **3**, e00224 (2017).
36. E. Chovancova, A. Pavelka, P. Benes, O. Strnad, J. Brezovsky, B. Kozlikova, A. Gora, V. Sustr, M. Klvana, P. Medek, L. Biedermannova, J. Sochor, J. Damborsky, CAVER 3.0: A tool for the analysis of transport pathways in dynamic protein structures. *PLOS Comput. Biol.* **8**, e1002708 (2012).
37. S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, A. H. de Vries, The MARTINI force field: Coarse grained model for biomolecular simulations. *J. Phys. Chem. B* **111**, 7812–7824 (2007).
38. G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).
39. M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
40. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
41. D. H. de Jong, G. Singh, W. F. D. Bennett, C. Arnarez, T. A. Wassenaar, L. V. Schäfer, X. Periole, D. P. Tieleman, S. J. Marrink, Improved parameters for the martini coarse-grained protein force field. *J. Chem. Theory Comput.* **9**, 687–697 (2013).
42. D. H. de Jong, N. Liguori, T. van den Berg, C. Arnarez, X. Periole, S. J. Marrink, Atomistic and coarse grain topologies for the cofactors associated with the photosystem II core complex. *J. Phys. Chem. B* **119**, 7791–7803 (2015).
43. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
44. G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, G. Bussi, PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **185**, 604–613 (2014).
45. T. A. Wassenaar, K. Pluhackova, R. A. Böckmann, S. J. Marrink, D. P. Tieleman, Going backward: A flexible geometric approach to reverse transformation from coarse grained to atomistic models. *J. Chem. Theory Comput.* **10**, 676–690 (2014).
46. M. K. Scherer, B. Trendelkamp-Schroer, F. Paul, G. Pérez-Hernández, M. Hoffmann, N. Plattner, C. Wehmeyer, J.-H. Prinz, F. Noé, PyEMMA 2: A software package for estimation, validation, and analysis of Markov models. *J. Chem. Theory Comput.* **11**, 5525–5542 (2015).
47. S. Jo, T. Kim, V. G. Iyer, W. Im, CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).
48. 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**, 926–935 (1983).
49. T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An *N*·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
50. R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig, A. D. MacKerell Jr., Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain $\chi_1$ and $\chi_2$ dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).
51. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, K. Schulten, Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
52. N. A. Baker, D. Sept, S. Joseph, M. J. Holst, J. A. McCammon, Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10037–10041 (2001).


**Acknowledgments:** We are thankful for the computing time provided by SuperMuc at the Leibniz Rechenzentrum (pr27xu). **Funding:** This work received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program/grant agreement 715311. This work was also supported by the German Research Foundation. **Author contributions:** A.J., A.D.L., and V.R.I.K. designed the project; A.J. performed the coarse-grained simulations; A.D.L. performed the atomistic simulations; A.J., A.D.L., and V.R.I.K. analyzed the calculations; and V.R.I.K. wrote the manuscript with contributions from all the other authors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 18 September 2018
Accepted 30 January 2019
Published 20 March 2019
10.1126/sciadv.aav1850

**Citation:** A. Jussupow, A. Di Luca, V. R. I. Kaila, How cardiolipin modulates the dynamics of respiratory complex I. *Sci. Adv.* **5**, eaav1850 (2019).