

**UNIVERSIDADE DE SÃO PAULO**  
**INSTITUTO DE QUÍMICA**  
Programa de Pós-Graduação em Ciências Biológicas  
(Bioquímica)

Murilo Hoias Teixeira

**Estrutura, dinâmica e reatividade de  
cofatores na cadeia de transporte de  
elétrons**

Versão corrigida

São Paulo

Data do depósito na SPG:  
06/05/2020



Murilo Hoias Teixeira

# **Estrutura, dinâmica e reatividade de cofatores na cadeia de transporte de elétrons**

*Tese apresentada ao Instituto de Química da  
Universidade de São Paulo para obtenção  
do Título de Doutor em Ciências  
(Bioquímica)*

*Orientador: Prof. Dr. Guilherme Menegon Arantes*

São Paulo  
2020

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Ficha Catalográfica elaborada eletronicamente pelo autor, utilizando o programa desenvolvido pela Seção Técnica de Informática do ICMC/USP e adaptado para a Divisão de Biblioteca e Documentação do Conjunto das Químicas da USP

Bibliotecária responsável pela orientação de catalogação da publicação:  
Marlene Aparecida Vieira - CRB - 8/5562

T266e      Teixeira, Murilo Hoias  
             Estrutura, dinâmica e reatividade de cofatores na  
             cadeia de transporte de elétrons / Murilo Hoias  
             Teixeira. - São Paulo, 2020.  
             93 p.

Tese (doutorado) - Instituto de Química da  
Universidade de São Paulo. Departamento de  
Bioquímica.

Orientador: Arantes, Guilherme Menegon

1. cadeia transportadora de elétrons. 2.  
complexo respiratório I. 3. quinonas. 4. agregados  
de ferro-enzofre. 5. dinâmica molecular e energia  
livre. I. T. II. Arantes, Guilherme Menegon,  
orientador.





Dedico este trabalho aos meus pais,  
Lucy e Manoel, ao meu irmão Moisés, à  
minha noiva, Tatiana, e à nossa amada filha, Helena.  
A estrutura que vocês compõe é fundamental.



# Agradecimentos

Presente em todo o processo da minha tese, agradeço ao meu orientador, prof. Dr. Guilherme Menegon Arantes. Trabalhamos juntos desde a minha iniciação científica em 2013 e suas contribuições e conselhos ajudaram a moldar o cientista e pessoa que me tornei. O cientista incansável que ele é, me inspirou a continuar independentemente de todos os desafios que um pesquisador enfrenta diariamente. Desses sete anos e meio de orientação, levo uma excelente experiência profissional e um amigo.

Agradeço também ao prof. Dr. Martin John Field por abrir as portas do seu laboratório e me receber em Grenoble por 11 meses. Sempre gentil, me ajudou desde o trabalho que desenvolvi em sua plataforma de simulação até em um acidente de bicicleta que sofri pelas ruas de Grenoble.

Essa tese dificilmente seria concluída sem a Tatiana Comporte Stabelini. Hoje minha noiva, me deu palavras de carinho, um refúgio do stress do mundo e um outro ponto de vista sobre o meu trabalho. Não satisfeita, me dará a honra de criar uma família juntos, com a chegada da nossa filha, Helena. Obrigado, muito obrigado.

Agradeço o grupo de alunos e ex-alunos do meu laboratório. Ariane, Raphael, Vanesa, Felipe, Sofia, André, João e Rodrigo foram excelentes companhias neste processo, compartilhando conselhos, visões, bandejões, experiências e arquivos. Um obrigado especial ao Felipe e Ariane pela proximidade e disponibilidade para me ajudar em todo e qualquer momento.

Agradeço aos meus ex-colegas em Grenoble, Romane, Bruno, Jordan, Elisa, Andrey, Astrid, Hender, Frank, Nicolas, Peter, Nico, Aljoša, Reiss, Marie, Dominique e Timothy pelos conselhos, risadas e tornar a minha estadia na França mais quente e suave apesar do inverno gelado. Agradeço minhas amigas de moradia Anne e Júlia pela parceria e ótimos momentos.

Não tenho palavras para agradecer minha mãe, Lucy, meu pai, Manoel e meu irmão, Moisés. Em nenhum momento eles deixaram de estender a mão e me empurrar para frente nessa jornada. Dedico-lhes todo o amor e carinho e espero compartilhar, não só esta tese, mas ainda mais uma longa e feliz vida ao lado deles.

Sou agraciado por um mar de amigos dentro e fora da academia e gostaria de deixar aqui uma palavra de carinho e agradecimento a todos. Eles foram essenciais para me dar fôlego e diferentes perspectivas para continuar este trabalho.

Agradeço ao programa de pós-graduação em ciências (bioquímica), cujo aulas e seminários contribuíram para minha formação. Destaco o incentivo e financiamento oferecido pelo programa, através do “Prêmio viagem”, para que eu e outros alunos apresentassem seus trabalhos em congressos e escolas internacionais. Também agradeço ao grupo de teoria do *Institut Laue-Langevin* pelo custeio completo da minha participação na *X International Society for Theoretical Chemical Physics* (ISTCP) em Trømsø, Noruega, durante minha estadia em Grenoble.

Agradeço ao Instituto de Química da Universidade de São Paulo por ser uma segunda casa por quase uma década entre bacharelado e doutorado. Lembrarei deste período de aprendizagem e formação com muito carinho. Agradeço também o *Institut*

*Laue-Langevin* pelo suporte e estrutura durante minha estadia em Grenoble.

Finalmente, agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo financiamento do meu doutorado direto e recursos para que eu frequentasse congressos científicos e escolas nacionais e internacionais (processo 142112/2015-1). Agradeço também à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) por financiar parte do meu doutorado e meu doutorado sanduíche (processos 1492455 e 88881.187923/2018-01, respectivamente). Agradeço aos recursos computacionais utilizados ao longo deste trabalho que foram financiados pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) ou disponibilizados pelo Laboratório Nacional de Computação Científica (LNCC) através do supercomputador Santos Dumont (SDumont).

“Ce qui est simple est toujours faux. Ce qui ne l’est pas est inutilisable.”

**Valéry, Paul (1942)**



# Resumo

Teixeira, M. H. **Estrutura, dinâmica e reatividade de cofatores na cadeia de transporte de elétrons** 2020. 91p. Tese - Programa de Pós-Graduação em Bioquímica. Instituto de Química, Universidade de São Paulo, São Paulo.

Simulação computacional é uma ferramenta poderosa para estudos em diversos campos da ciência. Neste trabalho utilizamos modelagem molecular para investigar microscopicamente o funcionamento de cofatores da cadeia transportadora de elétrons (ETC) em bactérias e mitocôndrias. As quinonas e agregados de ferro-enxofre (Fe-S) estudados aqui são respectivamente os maiores responsáveis pelo transporte de elétrons inter- e intra-complexos da cadeia respiratória. Entender o comportamento de quinonas dentro de membranas permite fazer previsões sobre as suas interações com proteínas. Simulações de dinâmica molecular clássica com campos de força aditivos realizadas aqui mostraram que a localização de diversas quinonas em uma membrana lipídica com composição similar à membrana interna de mitocôndrias é constante para diferentes quinonas naturais e compatível com a posição dos respectivos sítios de ligação em domínios membranares dos complexos da ETC. O flip-flop de ubiquinona na membrana mitocondrial é facilitado pela sua maior porosidade, fruto de lipídeos insaturados que permitem a intrusão de algumas moléculas de água junto à cabeça polar da quinona. A difusão transversal facilitada impede que este evento limite a taxa de renovação da ETC. O complexo respiratório I da ETC apresenta o sítio de redução de quinonas mais enterrado entre seus pares. O substrato ubiquinona deve transitar de um ambiente membranar por um longo e estreito túnel até o domínio hidrofílico proteico, onde o processo de redução para a forma quinol é catalisado. Para entender este processo, foram realizadas simulações atômicas com amostragem aumentada por potencial guarda-chuva do processo de ligação da quinona no complexo I da bactéria *Thermos thermophilus*. Nestas simulações, um segundo sítio estável, um intermediário no processo de ligação, foi identificado. Este pode estar relacionado ao acoplamento entre o transporte de elétrons e o bombeamento de prótons catalisados pelo complexo I. Recentemente, este intermediário foi isolado e comprovado experimentalmente. Também detectamos uma abundante hidratação interna do túnel de ligação que dificilmente é detectada em experimentos devido à mobilidade e tamanho das moléculas de água. O balanço desta hidratação é responsável pela energética de ligação de quinonas de cauda longa e curta observada nos perfis de energia livre calculados. Outros cofatores, agregados de Fe-S, têm papel fundamental no transporte de elétrons entre sítios distantes dentro de uma mesma proteína. Além disso, agregados de Fe-S oferecem suporte para o enovelamento e estruturação de enzimas envolvidas em diversas vias metabólicas. Aqui, um modelo isolado para hidrólise do agregado mononuclear de Fe-S foi utilizado para explorar os possíveis mecanismos desta reação. O mecanismo concertado de complexação e desprotonação de uma molécula de água com a saída de um grupo tiol apresenta as menores barreiras. Em pH baixo, este mecanismo apresenta barreiras ainda menores e, portanto, pode ser o mecanismo catalítico utilizado por enzimas na síntese e degradação de agregados de Fe-S. As baixas barreiras observadas em meio ácido sugerem que agregados de Fe-S devem estar enterrados em proteínas para evitar sua degradação.

Palavras-chave: cadeia transportadora de elétrons, complexo respiratório I, quinonas, agregados de ferro-enxofre, dinâmica molecular, energia livre



# Abstract

Teixeira, M. H. **Structure, dynamics and reactivity of cofactors in the electron transport chain** 2020. 91p. PhD Thesis - Graduate Program in Biochemistry. Instituto de Química, Universidade de São Paulo, São Paulo.

Computer simulations are powerful tools in several science fields. Here, we apply molecular modeling to microscopically investigate how cofactors of the electron transport chain (ETC) work in bacteria and mitochondria. Quinones and iron-sulfur (Fe-S) clusters studied here are the major players respectively on inter- and intra-electron transport on respiratory chain complexes. Understanding of how quinones behave inside membranes allows predictions of their interactions with proteins. Classical molecular dynamics simulations with additive force fields performed here have shown that localization of different quinones in a lipid membrane with similar composition to inner mitochondrial membranes is constant to several natural quinones and compatible with respective quinone binding sites located in hydrophobic domains of ETC complexes. Ubiquinone flip-flop in the mitochondrial membrane is easier due to a higher porosity, created by unsaturated lipids that allow intrusion of few water molecules together with the the quinone polar head. A facilitated transversal diffusion prevents this event to limit the turnover rate of ETC. The respiratory complex I presents the most buried quinone reduction site of all ETC complexes. The substrate ubiquinone has to transit from a membrane environment through a narrow and long tunnel up to the hydrophilic protein domain, where the reduction process to quinol form is catalyzed. To understand this process, atomistic umbrella sampling simulations of the quinone binding process in the complex I of the bacteria *Thermos thermophilus* were performed. On these simulations, a second stable site, an intermediary in the binding process, was identified. This may be related to the coupling between the electron transport and proton pumping processes catalyzed by complex I. Recently, this intermediary was isolated and verified experimentally. We also found extensive internal hydration inside the binding tunnel which is hardly detected experimentally due to the size and mobility of water molecules. The balance of such hydration is responsible for the binding energetics of long and small tail quinones observed on the calculated free energy profiles. Another class of cofactors, Fe-S clusters, has a fundamental role in the electron transport between distant sites present inside the same protein. Moreover, Fe-S clusters offer structural support to folding and shape of enzymes involved in several metabolic pathways. Here, an isolated model for hydrolysis of a mononuclear Fe-S cluster was chosen to explore possible mechanisms for this reaction. The concerted mechanism of complexation and deprotonation of a water molecule with the leave of a thiol group presents the smaller energy barriers. At low pH, this mechanism presents even smaller barriers and, therefore, may be the catalytic mechanism employed by enzymes in biosynthesis and degradation of Fe-S clusters. The small barriers found in acidic conditions suggest that functional Fe-S clusters should be buried in proteins to avoid its degradation.

Keywords: electron transport chain, respiratory complex I, quinones, iron-sulfur clusters, molecular dynamics, free energy



# List of abbreviations and symbols

ATP	Adenosine tri-phosphate
cyt	Cytochrome
CV	Collective variable
DFT	Density functional theory
E	Potential energy
ETC	Electron transport chain
$\vec{F}$	Force
F	Free energy
Fe-S	Iron-sulfur
IMS	Intermembrane space
OxPhos	Oxidative phosphorylation
$k_B$	Boltzmann constant
MM	Molecular mechanics
NADH	Nicotine adenine dinucleotide
PDB	Protein data bank
PMF	Potential of mean force
POPC	1-Palmitoyl-2-oleoylphosphatidylcholine
Q	Quinone
QH <sub>2</sub>	Quinol
QC/MM	Quantum chemical and molecular mechanical hybrid potential
QC	Quantum chemistry
RMSD	Root mean square deviation
T	Temperature
US	Umbrella sampling
Z	Partition function
WHAM	Weighted histogram analysis method



# Contents

<b>1</b>	<b>INTRODUCTION . . . . .</b>	<b>19</b>
<b>1.1</b>	<b>Electron Transport Chain . . . . .</b>	<b>20</b>
1.1.1	Complex I, II and III . . . . .	21
1.1.2	Overview of natural quinones and iron-sulfur clusters . . . . .	25
<b>1.2</b>	<b>Aims . . . . .</b>	<b>28</b>
<b>1.3</b>	<b>Computational Methods . . . . .</b>	<b>28</b>
1.3.1	Potential energy . . . . .	29
1.3.1.1	Classical Force-fields . . . . .	30
1.3.1.2	Quantum Chemistry . . . . .	31
1.3.1.3	Classic and Quantum Hybrid descriptor . . . . .	32
1.3.2	Sampling . . . . .	33
1.3.2.1	Molecular Dynamics . . . . .	33
1.3.2.2	Free-energy . . . . .	34
1.3.2.3	Enhanced Sampling . . . . .	35
<b>2</b>	<b>EFFECTS OF LIPID COMPOSITION ON MEMBRANE DISTRI- BUTION AND PERMEABILITY OF NATURAL QUINONES . . . . .</b>	<b>39</b>
<b>3</b>	<b>BALANCED INTERNAL HYDRATION DISCRIMINATES SUBSTRATE BINDING TO RESPIRATORY COMPLEX I . . . . .</b>	<b>51</b>
<b>4</b>	<b>MODELING THE HYDROLYSIS OF IRON-SULFUR CLUSTERS . . . . .</b>	<b>73</b>
<b>5</b>	<b>CONCLUSION . . . . .</b>	<b>85</b>
<b>6</b>	<b>REFERENCES . . . . .</b>	<b>87</b>



# 1 Introduction

The triad theory, modeling and experimentation is the base of modern science [1]. Recently, the relevance of computational modeling has been raising due to steady advances on computational power (better machines [2,3] and optimized algorithms) and new methodologies with cheaper and more accurate approximations [4,5].

Open questions on biochemistry and biophysics have been addressed by computer simulations, bringing insights and explanations for several phenomena such as cellular movement [6,7], biomolecules diffusion and binding [8–10], enzymatic catalysis [11–13], protein folding [14,15] and so on. In 2013, a Nobel prize was credited to Warshel, Karplus and Levitt in recognition of the development of hybrid quantum chemical and molecular mechanics (QC/MM) potentials to describe macromolecular systems like whole enzymes in condensed phase [16].

In this work, these computational methods are employed to study the electron transport chain (ETC). ETC, an ubiquitous and central metabolic pathway, is a constant subject of experimental [17–23] and computational studies [24–26]. Specifically, two essential cofactors of this pathway, quinones (Q) and iron-sulfur (Fe-S) clusters, are studied in this thesis. Both of them take part in energy transduction processes respectively between and inside membrane proteins of the ETC.

This thesis is organized as follows: This is an introductory chapter (1) designed to contextualize the reader with the subject of study (section 1.1), the questions addressed (section 1.2) and how it has been done (computational methods, section 1.3). The three following chapters (2-4) include published results and a last chapter (5) presents final remarks, conclusions and perspectives.

## 1.1 Electron Transport Chain

Living organisms centralize their metabolism in a few metabolites to increase efficiency and flexibility [27]. A key metabolite is adenosine tri-phosphate (ATP). This molecule is the main energy “currency” in cells and control of its concentration is essential to regulate their metabolism [28].

The major route to synthesize ATP in cells is oxidative phosphorylation (OxPhos), from which the ETC is part of [28]. It is ubiquitous, from bacteria to humans, and its malfunction leads to cell death and severe diseases [29–31].

In the inner membrane of mitochondria (fig. 1), the ETC is known as respiratory chain and composed essentially by four protein complexes (commonly named complex I to IV). In this context, ATP synthase is also known as respiratory complex V, even though it is not directly involved in electron transport.

Complexes I and II catalyze the oxidation of nicotine adenine dinucleotide (NADH) and succinate, respectively. These two complexes are the pathway entrances and redirect electrons to the following complexes through mobile quinols ( $\text{QH}_2$ ). Electrons are finally delivered to oxygen molecules by complex IV. Part of the energy released in these oxidation processes is stored in a proton gradient between intermembrane space (IMS) and mitochondria matrix. When complex V dissipates the generated proton gradient, ATP molecules recently synthesized and still strongly attached to its catalytic site can then be released to the mitochondria matrix.

Several protein cofactors play fundamental roles in these proteins. Two of them, quinones and Fe-S clusters are present in multiple points of this pathway and are the main subjects of this thesis. Comments on the first three complexes are given in section 1.1.1 with special attention to complex I. All of these proteins interact directly with quinones and move electrons between their reactive sites through "conductive" chains of Fe-S clusters. Further details on quinones and Fe-S clusters are given on section 1.1.2.

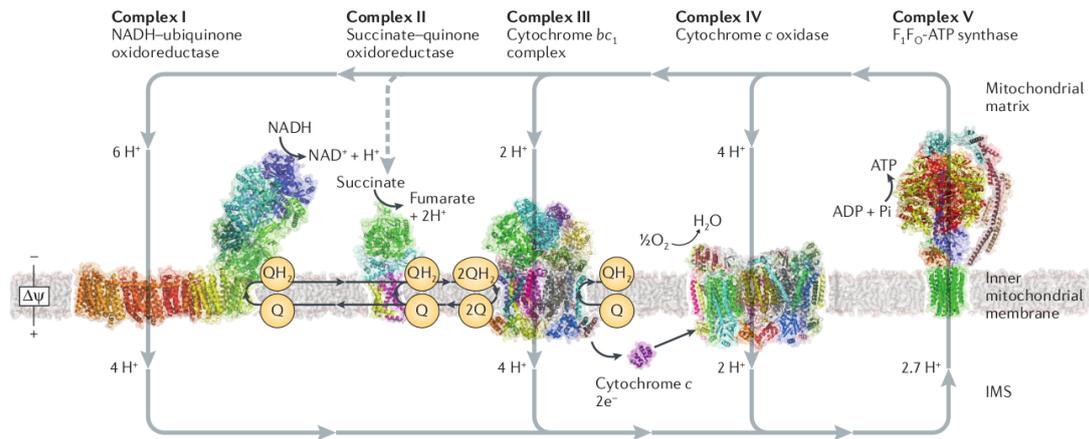


Figure 1 – The electron transport chain is composed by four transmembrane complexes localized in the inner mitochondrial membrane. Complexes I, III and IV generate a proton gradient that drives ATP synthesis in the  $F_1F_0$ -ATP synthase. Complex II provides a second entry point for electrons through succinate oxidation. Electrons extracted in complex I and II are shuttled to complex III by an ubiquinone(Q)/ubiquinol( $QH_2$ ) pool. This figure is extracted from ref [32].

### 1.1.1 Complex I, II and III

The biggest of the five complexes mentioned, the "NADH:ubiquinone oxidoreductase", or simply respiratory complex I, is unique. It is a huge and fully asymmetrical multidomain protein. The prokaryotic protein is usually composed by 14 subunits ( $\sim 500$  kDa mass) which compose the catalytic core and, thus, are highly conserved among all species. The mammalian protein is even more sophisticated. It presents additional 31 subunits (summing up to 1 MDa mass including the core) which may participate in complex assembly, stability and regulation (fig. 2) [32,33].

Not surprisingly, its extent covers both leaflets of the inner mitochondrial membrane and part of mitochondria matrix. The soluble domain possesses both, NADH oxidizing and quinone reductive, reactive sites. These two sites are more than  $80\text{\AA}$  away of each other, and thus too far apart to enable directly electron tunneling. The redox process operate by electrons tunneling indirectly from the NADH oxidizing site to the N2 iron-sulfur cluster through six iron-sulfur clusters. A quinol is then formed by transferring electrons from N2 cluster and protons from neighbor residues to a bound quinone (fig. 3).

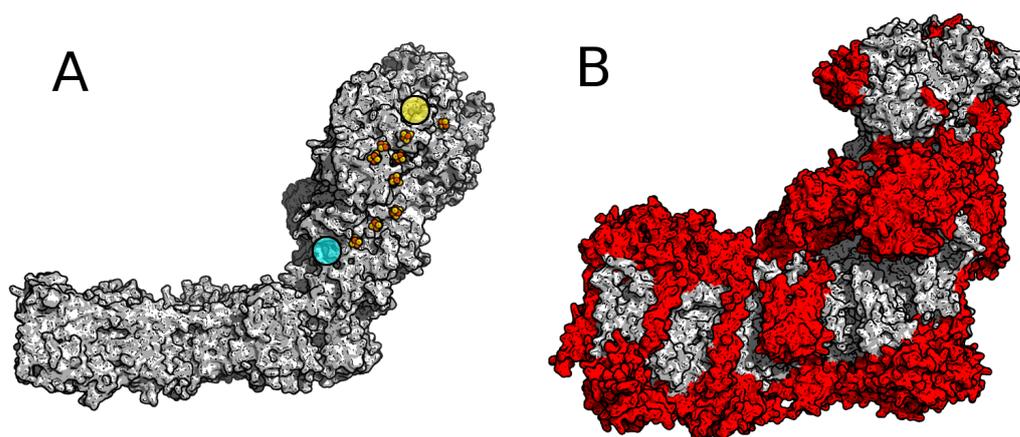


Figure 2 – (A) *Thermus thermophilus* complex I structure (PDB code: 4HEA). (B) *Homo sapiens* complex I structure (PDB code: 5XTD). Protein surfaces are presented in grey and red for catalytic and supernumerary subunits, respectively. On *T. thermophilus* structure, quinone reduction site (cyan shade), NADH oxidation site (yellow shade) and iron-sulfur cluster (spheres) are also indicated.

Quinone binding requires that quinones crawl up from a membrane environment to inside of the soluble domain. This molecular diffusion occurs through a narrow chamber (Q-chamber) connecting the membrane/water interface and the buried N2 reductive site. The tunnel is long enough to simultaneously accommodate 8 out of 10 isoprenoid units of the Q-tail and its entry point is defined by three alpha-helices present in the Nqo8 subunit: AH1, TM1 and TM6 (fig. 3). Point mutation of residues along this chamber are associated to activity loss and human diseases [29].

In chapter 3, we model this binding process on *Thermus thermophilus* complex I in a pure 1-Palmitoyl-2-oleoylphosphatidylcholine (POPC) membrane. The model chosen included all subunits of complex I with an atomistic resolution. In this work, a stable intermediary site in the binding process was indicated. Experimental studies recently published (after our simulations) isolated and verified the site existence (fig. 2B in ref [34]), corroborating the simulation data presented here. Upcoming and still unpublished high resolution structures from collaborators also indicate the presence of internal water molecules inside the Q-chamber (personal communication by Judy Hirst, UK. Also see comments in the “Future Issues” section in ref [35]), in the same location as predicted in chapter 3.

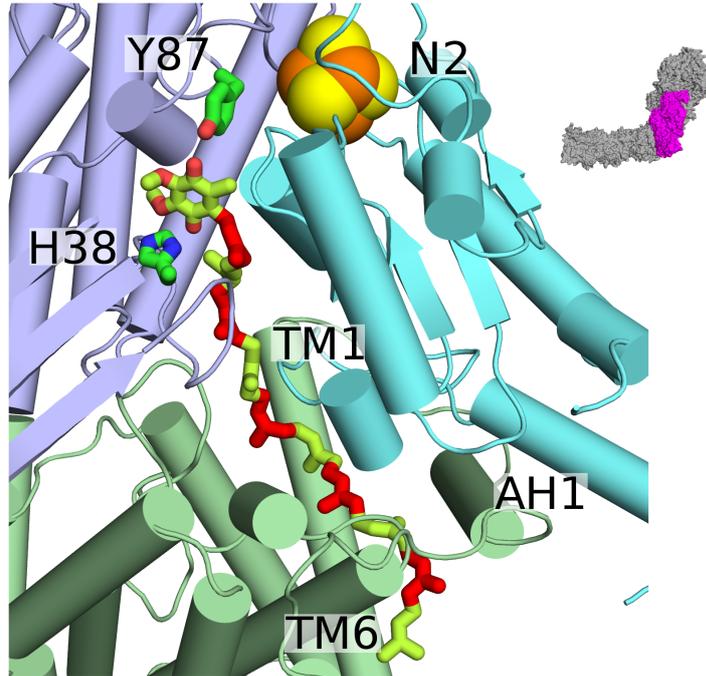


Figure 3 – Close view of subunits Nqo4 in light blue, Nqo6 in cyan and Nqo8 in pale green from *Thermus thermophilus* complex I that form the Q-chamber with modeled ubiquinone-10 colored in green and red alternating for each isoprenoid unit. Elements of secondary structure and catalytic residues (green sticks) discussed in the text are indicated. The miniature in grey highlight in magenta the region containing subunits Nqo4, 6 and 8.

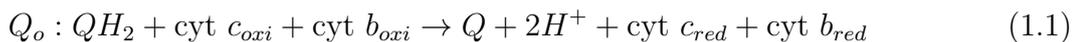
The electron transport process by itself is not unique and conductive iron sulfur chains can be found in other proteins [36–38]. However, generation of the so said proton gradient by this enzyme is. Three membrane subunits of complex I evolved from antiporter proteins and are capable of conducting active and directional proton transport [39]. These three subunits and an unclear fourth proton path are responsible to, somehow, conduct, on average, four proton for each NADH oxidized in the soluble domain (fig. 1). It is still unclear how this process is orchestrated in this huge complex, but clues have been discovered in the recent years [40–42] and some of them link it to Q-chamber occupation [43, 44].

There is a second electron entry point in the respiratory chain. In mitochondria, succinate dehydrogenase, or respiratory complex II, is composed by four subunits separated in two domains, a hydrophilic and a hydrophobic. Its soluble domain is located in the matrix (fig. 1) which allows this enzyme to connect both ETC and Krebs Cycle.

Complex II removes electrons (oxidation) of succinate, forms fumarate, and deliver them (reduction) to quinones, forming quinols. The succinate oxidative and quinone reductive sites are also 37Å far apart and connected by 3 iron-sulfur clusters, which facilitates electron tunneling [45].

All quinones reduced by complex I and II are then processed by the third complex of the ETC, “coenzyme Q:cytochrome c - oxidoreductase”. Also called as cytochrome bc<sub>1</sub> or respiratory complex III, it is a dimer with each subunit composed by several subunits (three in “simple” organisms or up to eleven in more “complex” ones) and four cofactors, one iron-sulfur cluster (2Fe-2S), two heme b and one c<sub>1</sub> in each monomeric unit.

Differently than the proton pump mechanism of respiratory complex I, complex III uses the proposed "Q cycle" to generate a proton gradient [28,46,47]. This transmembrane enzyme has binding sites for both quinols (Q<sub>o</sub> site) and quinones (Q<sub>i</sub> site), which are conveniently located so that their oxidation and reduction reactions release and capture protons in opposite membrane sides (fig. 4). This sequential process requires two quinols to reduce a cytochrome c. In each turnover, no electrons are wasted since a new quinol is formed (eq. 1.1 and 1.2).



where  $Q$ ,  $QH$  and  $QH_2$  are quinone, semiquinone and quinol molecules, respectively. Cytochromes  $b$  and  $c$  are indicated in their oxidized (oxi) and reduced (red) forms. It is still unclear how quinols are oxidized in Q<sub>o</sub>, since no semiquinone intermediates have been observed experimentally which would require an uncommon two-electron oxidation [48].

All quinone reductive and oxidative sites of complexes I to III present their entrances for Q molecules around 1.75 nm away from the membrane center (fig. 4). It is probable that electron carriers and enzymes related evolved together to enhance compatibility by adjusting equilibrium and entrance site positions, respectively. On chapter 2, we investigate the equilibrium positions of different quinones in a mitochondria-like membrane to test this hypothesis.

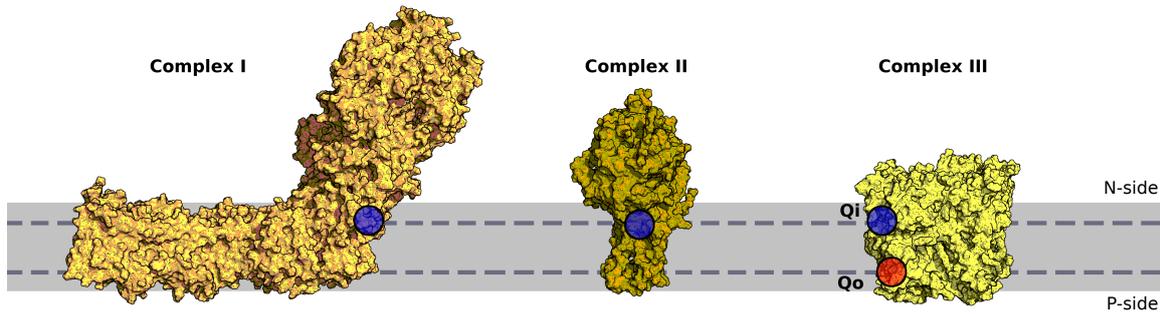


Figure 4 – Entrance sites for Q binding in respiratory complexes III. Sites for quinone reduction shown in blue and for quinol oxidation in red. Bilayer is represented in solid gray and equilibrium or preferential position of the Q-head across the membrane normal (1.75 nm) are shown with a dashed line. Negative (N-side) and positive (P-side) membrane sides refers to interfaces with the mitochondria matrix and IMS, respectively. Structures were taken from PDB code 4HEA [29] for complex I, 2H88 [20] for complex II and 2QJP [49] for complex III.

Electrons are then passed from complex III to the last inter protein electron carrier in ETC, soluble cytochrome *c*, which delivers them to complex IV to finally reduce oxygen to water (fig. 1). Quinones and Fe-S cluster does not take part in energy transduction processes which happen in complexes IV and V. Details on these complexes will not be described here, but are available elsewhere [50,51].

### 1.1.2 Overview of natural quinones and iron-sulfur clusters

The capability of electron transport chain enzymes of moving electrons through long distances is essential for their function. Inside proteins, this role is performed by isolated metal ions such as copper in complex IV or inserted in Fe-S clusters and heme groups. Electron shuttle between complexes is done by quinone/quinol or cytochrome *c* equipped with a heme group.

Quinones are ubiquitous in nature acting as electron carriers in membranes, helping photosynthesis and respiration. They are constituted of a hydrophilic head and a hydrophobic tail. The head group (Q-head) is an oxidized aromatic group presenting at least two aromatically conjugated carbonyl oxygens. There are different Q-heads such as

menaquinone and plastoquinone which are less polar than ubiquinone (fig. 5). The Q-head is related to differences in membrane partitioning [52] and binding affinity to respiratory and photosynthesis complexes [53, 54].

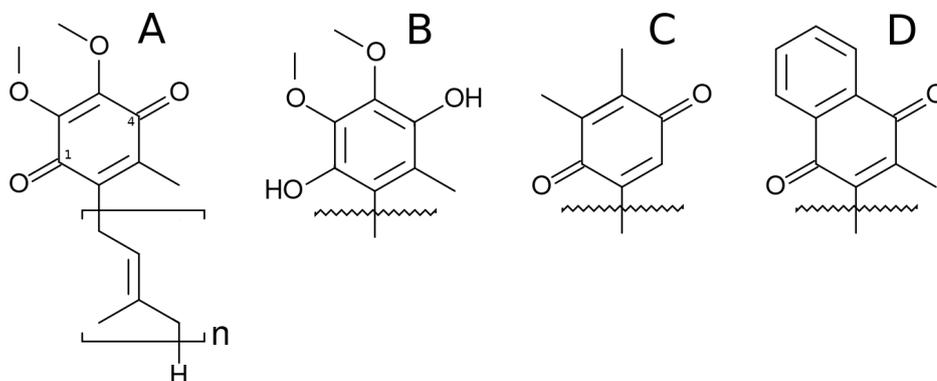


Figure 5 – Quinone structures. Ubiquinone (A), Ubiquinol (B), Plastoquinone (C) and Menaquinone (D) are indicated. Quinone tail with  $n$  isoprenoid units is represented only on ubiquinone for simplicity. Reactive oxygens are bonded to carbons 1 and 4 (A). This figure is extracted from chapter 2.

The hydrophobic tail (Q-tail) is composed *in vivo* by 6 to 10 isoprenoid units. Tail size is associated to overall substrate affinity [53–55], lateral [10, 56, 57] and transversal [10, 58] diffusion rates and partitioning constants [10, 57] in membranes. In complex I, assays with smaller tail quinones present reduced activity and even an auto-inhibitory effect when high membrane concentrations were used [55, 59–61].

Quinone specific localization in membranes has been discussed in the literature [10, 62–71]. Different preferential position in membranes may lead to different reactivity or binding affinity and kinetics with transmembrane enzymes [10, 62]. Previous works have shown contradictory information, pointing for an equilibrium position of Q-head in the water-lipid interface [10, 68–71] or the center [62–67] of biomembranes. Simulations described on chapter 2 indicate a membrane interface equilibrium position (fig. 4) for all quinones tested.

Lateral and transversal (i. e. flip-flop) diffusion rate of quinones in membrane have been pointed as limiting steps of ETC and photosynthesis [10, 18, 62, 72–74]. Thus, understanding quinone mobility and specific interactions with biological partners is key to

precisely describe the molecular mechanisms of ETC. Transversal diffusion of ubiquinones as a limiting step of ETC is also investigated on chapter 2, where flip-flop barriers and rates are calculated and compared for two membranes with different lipid tail unsaturation levels.

Iron-sulfur clusters are present as natural cofactors in a wide range of proteins [75–77]. Formed by iron cations (Fe II or III) and sulfide anions ( $S^{-2}$ ), they may exist in many shapes, Fe:S proportions and with several different ligands [28, 78]. Due to this rich chemistry, it has been a hot topic of research for several decades [75–77, 79].

Usually, Fe-S clusters are composed of 1 (*e.g.* rubredoxin) to 4 (*e.g.* ferredoxin, respiratory complexes, etc.) iron atoms (fig. 6). Higher nuclearity may also exist in nature by bridging smaller cluster by sulfides or thiolates [80, 81]. Special clusters present metal doping, including Mo or Ni atoms, in the structure [82].

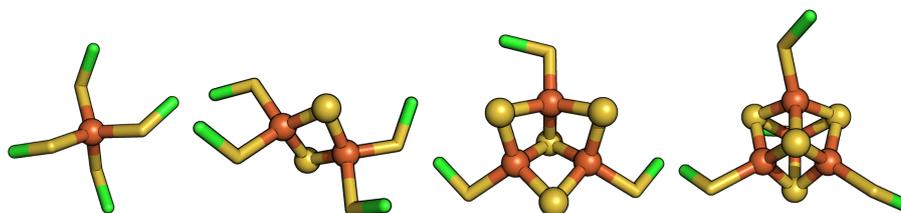


Figure 6 – Iron-sulfur clusters with crescent nuclearity (left to right). Iron, sulfur and carbons atoms are colored in orange, yellow and green, respectively. Bonds are indicated as sticks and atoms forming the clusters as spheres. Hydrogens or protein scaffold atoms linked to carbons are not represented for simplicity.

Historically, Fe-S clusters are known by their oxidoreductive nature. Its redox potential is highly variable by metal constitution [83, 84], ligands [85–87] and pH when the near environment is titrable [86]. Redox tuning is essential to create a directional electron flow in the ETC [28, 78], where the reduction potential of cofactors rises towards the final complexes until oxygen, highly reducible.

Despite their redox activity, these clusters may play a structural role in protein assembly and folding [88]. It provides a stable joint point for proteins like rubredoxin [89] or endonuclease III [90]. General stability and resistance to degradation are fundamental to perform such roles. On chapter 4, a Fe-S cluster degradation mechanism through hy-

drolysis is investigated under neutral and acid conditions. Alternative biological activity as iron level sensors (FNR in *E.coli* [91]), iron reserves (iron regulatory protein 1 [92]) and more are also possible [79].

## 1.2 Aims

The main aim of this thesis is to understand molecular mechanisms involving ETC cofactors on energy transduction.

Specifically, on chapter 2 the aim is to address how Q-head constitution affects quinone partition, localization and transversal diffusion in multi-component membranes mimicking mitochondria inner membrane. An additional aim of this chapter is to understand the influence of membrane lipid unsaturation levels in such quinone properties and possibly link these results to the overall ETC turnover rate regulation.

On chapter 3, the main purpose is to identify water and specific residue interactions that may regulate quinone binding in complex I on *Thermus thermophilus*. Furthermore, the chapter also intends to comprise how internal chamber hydration alters binding thermodynamics and kinetics of ubiquinones with different Q-tail lengths.

Lastly, chapter 4 targets in describing what is the specific chemical reaction steps on iron-sulfur cluster hydrolysis in neutral and acidic conditions. This may indicate why Fe-S clusters should be buried in proteins to prevent degradation and how they are synthesized *in vivo*.

## 1.3 Computational Methods

The range of computer simulations in science grows bigger and broad, including several areas like mathematics, biology, astronomy and so on [93]. Biochemistry and biophysics are no different and questions raised in section 1.2 will be addressed in this thesis using molecular modeling.

A model for a particular system is built in a way it reproduces the features required

to describe a target phenomenon. Every model is an approximation of reality and solid knowledge of the approximations applied outline the possible conclusions one can draw from its results. In molecular simulations, models are defined by their atomic composition, an initial set of coordinates and a descriptor for interatomic interactions [94]. Time propagation or minimum energy geometry search techniques are used to study dynamic and equilibrium properties of the system modeled [94, 95].

Here, a brief introduction to computational methods applied in this thesis are given. The section 1.3.1 introduces molecular mechanics, quantum chemistry and hybrid methods used to describe particle interactions. Sampling procedures and a conceptual definition of free-energy are given in section 1.3.2.

### 1.3.1 Potential energy

Potential energy is the energy associated with a specific configuration. It can describe the relative stability between different configurations of a system. Higher energy configurations are less stable than lower ones. The best interaction descriptor to get exact potential energy values is quantum mechanics. However, it is also very expensive computationally and imposes a limitation on model size (hundreds of atoms) and number of independent calculations that can be made.

Approximations on potential energy description, such as classical molecular mechanics, overcome this problem by collapsing electron densities in the nuclei and defining bond and non-bonding interactions between atoms. At the end, this replaces an intricate interdependent system of equations by a set of simple polynomial and sinusoidal linear equations. More subtle and, consequently, more expensive approximations are also possible. For example, the wave function or the density functional theory (DFT) methods, whose are capable of dealing with electron distributions are broadly used for small systems. Molecular mechanics (MM), quantum chemistry (QC) and hybrid QC/MM approaches will be described in sections 1.3.1.1, 1.3.1.2 and 1.3.1.3, respectively.

### 1.3.1.1 Classical Force-fields

Classical molecular mechanics force-fields are useful to sample conformations of macromolecules with low computational cost. These force-fields are a set of functions and computational protocols applied to describe the energy of configurations of a system. The choice of functions or protocols may be different for each force-field. All functions described below are the ones applied by the CHARMM36 [96] force-field which is the one chosen to describe MM interactions in chapters 2 to 4 of this thesis. Energy contributions between ligand atoms describing covalent bonds and angles between bonds may be estimated by harmonic functions like [94]:

$$v_b = \sum_{bond/ang} \frac{1}{2} k_b (b - b_0)^2 \quad (1.3)$$

where  $v_b$  is the bond or angle energy,  $k_b$  is a force constant,  $b$  is the bond length or bond angle between three connected atoms and  $b_0$  is the equilibrium bond length or bond angle between three connected atoms. The energies of proper and improper dihedrals (distortions out of a planar configuration) may be approximated by periodic functions [97]:

$$v_d = \sum_{dih} \sum_n V_d [1 + \cos(n\phi - \delta)] \quad (1.4)$$

where  $v_d$  is the dihedral or distortion energy,  $V_d$  is a force constant,  $n$  is the angle periodicity,  $\delta$  is the angle phase and  $\phi$  is the dihedral angle.

Non-bonding contributions usually include electrostatic and van der Waals interactions. These are described in the force-fields by the Coulomb law ( $v_{elect}$ ) and Lennard-Jones potential ( $v_{vdW}$ ) terms [94], respectively:

$$v_{elect} = k_{elect} \sum_{i < j} \frac{q_i q_j}{r_{ij}} \quad (1.5)$$

where  $k_{elect}$  is a constant related to the dielectric constant of the medium,  $q_i$  and  $q_j$  are partial charges attributed to atoms  $i$  and  $j$  to represent its net charge and  $r_{ij}$  is a distance between the atoms,

$$v_{vdW} = k_{vdW} \sum_{i < j} \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \quad (1.6)$$

where  $k_{vdW}$  is a scaling factor and the pair  $A_{ij}$  and  $B_{ij}$  depends on the atom types  $i$  and  $j$  and may be defined by combination rules that usually use arithmetic and geometric averages of atom type parameters. The term  $1/r_{ij}^{12}$  refers to the overlapping energy between near electron densities, which causes repulsion between atoms, while the term  $1/r_{ij}^6$  refers to a dispersion energy resulting from charge distribution fluctuations of both atoms which pull them together [95]. Note that these potential energy approximations neglect several contributions such as polarization or charge transfer which are naturally included in more rigorous treatments based on quantum mechanics.

### 1.3.1.2 Quantum Chemistry

A more exact framework to describe interactions between particles is quantum mechanics. It is the appropriate descriptor of electrons behavior in molecules and, thus, is capable to accurately describe the formation and cleavage of atom bonds. In many-electron wave function problems, the major target is to solve the non-relativistic time-independent Schrodinger equation:

$$\hat{H}\Psi = E\Psi \quad (1.7)$$

where  $E$  is the potential energy,  $\Psi$  the wave function and  $\hat{H}$  the Hamiltonian which describes interactions of the system.

The  $\Psi$  wavefunction solution for this eigenvalue problem is commonly a combination of several electronic configurations ( $\Psi_i$ )

$$\Psi = \sum_i c_i \Psi_i \quad (1.8)$$

where  $c_i$  weighs the contribution of each  $\Psi_i$  configuration. Simpler methods assume that  $\Psi$  is monoconfigurational, restricting the problem to find a single most stable electronic configuration. Additionally to ease the problem, the Born-Oppenheimer (BO) approximation separates the relative motions of electrons and nuclei in the system. The BO approximation allows one to solve the Eq. 1.7 for a fixed arrangement of nuclei and, accordingly to the Hellmann-Feynman theorem [98], forces can be calculated and the system nuclei spatial configuration propagated in a similar way to classical mechanics.

The simplest way to solve Eq. 1.7 within the BO approximation is using the mean field approximation. In a mean-field, electrons do not interact directly with each other and, instead, each electron interacts with an average density created by the remaining electrons on the system. All these approximations together define the Hartree-Fock (HF) method. The HF approach accurately describes the physics of electronic structures presenting negligible electron correlation [99]. If necessary, it is possible to partially recover this correlation contribution with more expensive methods such as the Møller-Plesset perturbation theory [100], multiconfigurational wave-functions [5] or the Coupled-Cluster theory [101].

Another relatively cheaper approach is the Density Functional Theory. This theory takes advantage that in a system, both wave function (highly dimensional) and the corresponding electron density (three dimensional), are capable of describing all ground-state properties [102]. The major limitation of this approach is that there is no analytic formulation to describe the full relationship between density and other system properties, such as the potential energy. The lacking of an exact energy functional formulation, leads to several flavors of DFT functionals, each one with different approximations and/or parametrizations, presenting also different pros and cons.

### 1.3.1.3 Classic and Quantum Hybrid descriptor

The description of potential energy as a hybrid of quantum chemistry and molecular mechanics (QC/MM) [103] allows the study of a wide range of phenomena in condensed phase like, for example, reactivity of metalloproteins. To apply a QC/MM potential in biological systems, only regions of an enzyme where the reactive process are localized, like substrate and residues within the active site, are treated quantum mechanically, while all non-reactive atoms are described by molecular mechanics force-fields. By doing this, the physics in the region where the electronic structure changes (*e.g.* bond cleavage) is preserved without losing the environmental influence of the remaining system. The full Hamiltonian is effectively written as:

$$\hat{H}_{eff} = \hat{H}_{QC} + V_{MM} + \hat{H}_{QC/MM} \quad (1.9)$$

where an effective energy function ( $\hat{H}_{eff}$ ) is defined by adding a contribution of a quantum region ( $\hat{H}_{QC}$ ), a molecular mechanics described region ( $V_{MM}$ ) and an interaction contribution between these two regions ( $\hat{H}_{QC/MM}$ ) usually defined by:

$$\hat{H}_{QC/MM} = - \sum_i^{n_e} \sum_A^{M_m} \frac{q_A}{r_{iA}} + \sum_q^{M_q} \sum_A^{M_m} \frac{z_q q_A}{r_{qA}} + \sum_q^{M_q} \sum_A^{M_m} \hat{H}_{QC/MM}^{vdW} \quad (1.10)$$

which includes electrostatic interaction terms between  $n_e$  electrons( $i$ ) and  $M_m$  classical partial charges ( $q_A$ ), between these same charges and the  $M_q$  quantic nuclei ( $z_q$ ), and a van der Waals term between classic and quantic centers, usually described by a Lennard-Jones potential. Variations for describing the  $\hat{H}_{QC/MM}$  term are possible and described elsewhere [104].

### 1.3.2 Sampling

Molecular modelling of phenomena requires that interaction energies and the configurational space of a model are carefully described [95]. In this thesis, molecular dynamics (MD) simulation is broadly used to propagate spatial coordinates and its principles are briefly outlined on section 1.3.2.1. A conceptual definition of free energy and a sampling technique applied to achieve proper configurational sampling convergence are detailed on sections 1.3.2.2 and 1.3.2.3, respectively.

#### 1.3.2.1 Molecular Dynamics

Independently of the potential energy descriptor chosen, propagation of the system coordinates through time can be made by following Newton's equation of motion:

$$m_i \frac{\partial^2 R_i}{\partial t^2} = \vec{F}_i \quad (1.11)$$

where  $\mathbf{m}$  is the particles mass vector,  $\mathbf{R}$  is the position vector,  $\vec{F}$  is the force exerted on the system and the index  $i$  refers to the  $i$ th particle in the system. This equation of classical mechanics is able to describe fairly rigorously the time evolution of coordinates for macromolecules.

Given a starting configuration and an initial set of velocities, integration of this equation for a short period of time can generate new configurations of atoms position and

velocities. The gradient of the potential energy determines the force  $\vec{F}$ . After a number of integration steps, different local configurations and, possibly, conformational states are sampled.

### 1.3.2.2 Free-energy

A proper description of potential energy is essential for a reliable model, but it is not sufficient to compare with experimental observables. Instead, observables in biophysical models are associated to configurational states, an *ensemble* of micro-states, or configurations, that a system visits in thermodynamic equilibrium.

Free-energy ( $F$ ) is the quantity capable of describing the relative stability between these states or the barriers that separate them. It has a direct relationship to the partition function ( $Z$ ) of a system. The partition function is a function that describes the system statistical properties by taking in account the stability of all configurations including all degrees of freedom [95]. Given two configurational spaces,  $V_A$  and  $V_B$ , it is possible to use such partition functions to estimate free-energy differences ( $F_A - F_B$ ) between states defined by such configurational spaces:

$$\frac{p(\text{state}A)}{p(\text{state}B)} = \frac{\int_{V_A} d\mathbf{r}^N e^{-E(\mathbf{r}^N)/k_B T}}{\int_{V_B} d\mathbf{r}^N e^{-E(\mathbf{r}^N)/k_B T}} = \frac{Z_A}{Z_B} \equiv e^{-(F_A - F_B)/k_B T} \quad (1.12)$$

where  $p(\text{state}X)$  is the probability of finding the system on state X,  $V_X$  is the configurational space volume defining state X,  $\mathbf{r}$  is a degree of freedom,  $E$  is the potential energy,  $k_B$  is the Boltzmann constant,  $T$  is the temperature and  $F_X$  is the free energy of state X. Here, free energy definition is generalized since volumes fluctuations are small in liquid phase for biochemical reactions and differences between Helmholtz and Gibbs free energies are largely irrelevant.

Understanding the relative stability of all relevant states that describe a property is not an easy task. It requires the modeler to sample the potential energy surface (PES), until all degrees of freedom linked to a given property are fully scanned [95]. Usually, biochemical or biophysical processes are investigated through lower dimensional free energy profiles called potential of mean force (PMF).

A single dimensional PMF represents a free-energy profile for a reaction coordinate  $\xi$ . It is defined by the distribution function  $\langle p(\xi) \rangle$  [105]:  $\text{PMF}(\xi) = k_B T \ln[p(\xi)]$ . Thus, differences of local minima and maxima in a PMF can be associated to relative stabilities or activation barriers, respectively. The relative stability of states (minimum regions, basins) of a system distinguish their population, i. e. their equilibrium constant. PMFs are not restricted to only one reaction coordinate and the concept is easily extended to include more coordinates. However, sampling convergence and data interpretation issues limits analysis of PMFs of higher dimensions commonly to two simultaneous reaction coordinates.

In this thesis, PMFs are calculated from distributions obtained from a set of sampling windows with an umbrella potential in different values along  $\xi$  for each window (see section 1.3.2.3). A mathematical reconstruction of the whole data set is done by the weighted histogram analysis method (WHAM) [106, 107], which re-weights the distribution from several windows and discounts the umbrella potential influence. As mentioned, the coordinate  $\xi$  is highly dependent on the process simulated.

### 1.3.2.3 Enhanced Sampling

In the limit of long simulation times, the ergodic theorem [108] states that the temporal average population equals the fully spatial average obtained by integrating the partition function (section 1.3.2.2). However, a complete and straightforward conformational sampling by molecular dynamics might be computationally expensive and inefficient.

Depending on the event timescale, energy barriers might be too high and it is unlikely that they will be overcome in short simulations, hindering the sampling. An alternative for that is to artificially alter the potential energy surface to orientate the system towards a target state. When a reaction coordinate describing the desired event is assumed and a harmonic ( $f(\xi) = k_{umbrella}(\xi - \xi_0)^2$ ) restriction is imposed along this reaction coordinate (fig. 7), this enhanced sampling technique is called umbrella sampling (US).

Umbrella sampling is one of the simplest configurational sampling enhancement

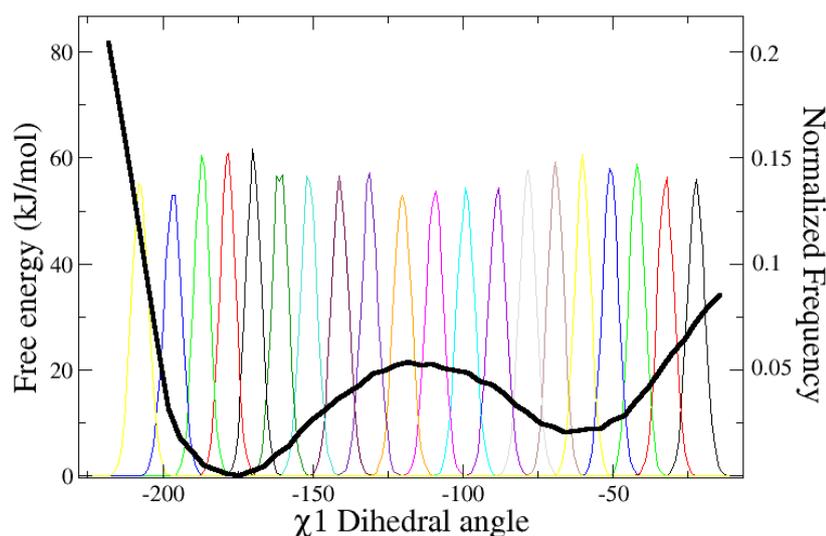


Figure 7 – Umbrella sampling windows and free energy surface for a typical leucine side-chain torsion. Distribution functions extracted from sequential sampling windows are colored in several colors. Sampled free energy surface is represented in a thicker black line. Distributions near low energy regions are spread less and, thus, present taller peaks.

techniques available. Analysis of the resulting distribution of the reaction coordinate in several sampling windows allows the reconstruction of the free-energy profile for that reaction coordinate or degree of freedom. High energy regions, such as barriers separating basins in a surface, return spread distributions. This happens due the fact that a system avoids as much as possible high energy regions even if it has to pay a price for going up the harmonic well of the umbrella potential. This subtle differences on the final distribution shapes are a reflex of the intrinsic free energy surface that the modeler desires to sample.

The definition of a reaction coordinate  $\xi$ , or collective variable (CV), which sampling is enhanced is quite flexible. It is usually composed of simple geometrical degrees of freedom like distances and angles, but it can also be more elaborated. For example, depending on which process the reaction coordinate has to describe, a root mean square deviation (RMSD) from a given reference or the number of contacts between two groups can be used [109].

In chapter II, the distance z-component, normal to membrane plane, is picked as the main CV to study quinone flip-flop and localization. Chapter III applies a combination of two CVs: geometrical distance of Q-head from the residues defining the binding site ;

and a progressive reference RMSD of a predefined path [110] (milestones, fig. 8). They are both chosen to ensure a complete sampling of the quinone translocation through the Q-chamber. Chapter IV also enhances the sampling in two reaction coordinates to describe hydrolysis of an iron-sulfur cluster: combined distances between the attacking and leaving group to the central iron; and a proton distance to a sulfur atom of the thiolate leaving group.

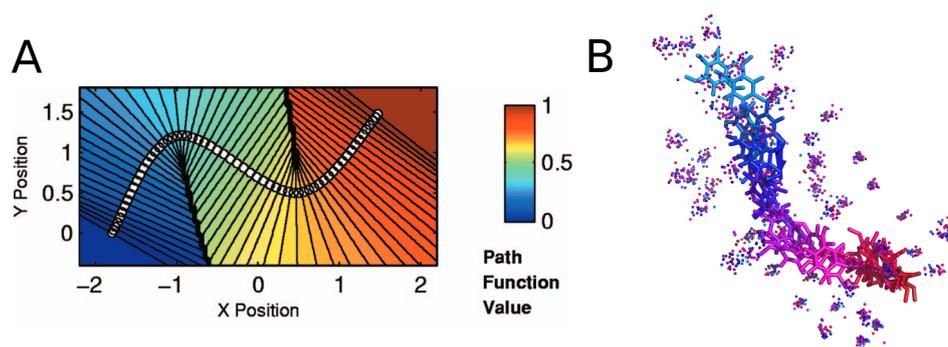


Figure 8 – Milestone concept and path used in chapter III. (A) Milestones in a 2D path are presented in white circles. Path coordinate value along milestones are normalized and colored from blue (0) to red (1). Points in (x,y) are mapped to path values through a distance metric to milestones. This discrete mapping is smoothed with a mathematical interpolation between milestones. Note that path values are ill defined before and after the first and last milestones, respectively. (B) Milestones used in simulations of quinone binding to respiratory complex I of *Thermus thermophilus* are colored from light blue to red accordingly to path values. Each milestone structure included Q-head atoms (sticks) and 55 C $\alpha$  (spheres) of residues in subunits Nqo4, Nqo6 and Nqo8 (fig. 3) in direct contact with quinone inside the chamber. This path collective variable is able to unambiguously describe curved binding paths.

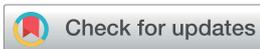


## 2 Effects of lipid composition on membrane distribution and permeability of natural quinones

Murilo Hoiias Teixeira and Guilherme Menegon Arantes

Department of Biochemistry, Instituto de Química, Universidade de São Paulo, SP,  
Brazil

Reprinted with permission from Teixeira, M. H.; Arantes, G. M. Effects of lipid composition on membrane distribution and permeability of natural quinones. *RSC Adv.*, v. 9, p. 16892-16899, 2019. Copyright The Royal Society of Chemistry 2019.



Cite this: *RSC Adv.*, 2019, 9, 16892

# Effects of lipid composition on membrane distribution and permeability of natural quinones†

Murilo Hoias Teixeira and Guilherme Menegon Arantes \*

Natural quinones are amphiphilic molecules that function as mobile charge carriers in biological energy transduction. Their distribution and permeation across membranes are important for binding to enzymatic complexes and for proton translocation. Here, we employ molecular dynamics simulations and free energy calculations with a carefully calibrated classical force-field to probe quinone distribution and permeation in a multi-component bilayer trying to mimic the composition of membranes involved in bioenergetic processes. Ubiquinone, ubiquinol, plastoquinone and menaquinone molecules with short and long isoprenoid tails are simulated. We find that penetration of water molecules bound to the polar quinone head increases considerably in the less ordered and porous bilayer formed by di-linoleoyl (18:2) phospholipids, resulting in a lower free energy barrier for quinone permeation and faster transversal diffusion. In equilibrium, quinone and quinol heads localize preferentially near lipid glycerol groups, but do not perform specific contacts with lipid polar heads. Quinone distribution is not altered significantly by the quinone head, tail and lipid composition in comparison to a single-component bilayer. This study highlights the role of lipid acyl chain unsaturation for permeation and transversal diffusion of polar molecules across biological membranes.

Received 5th March 2019

Accepted 20th May 2019

DOI: 10.1039/c9ra01681c

rsc.li/rsc-advances

## Introduction

Electron transfer chains (ETC) involved in biological energy transduction rely on natural quinones and their redox-coupled quinols (Q molecules‡, Fig. 1) to shuttle electrons between protein complexes and translocate protons across phospholipid membranes, contributing to generation of an electrochemical gradient.<sup>1–4</sup> Clearly, the mechanisms for molecular recognition and binding of Q by respiratory and photosynthetic enzymatic (super)complexes<sup>5,6</sup> depend on the membrane distribution of Q molecules.<sup>7–9</sup>

Diffusion of Q in the membrane has also been suggested to control ETC turnover rates.<sup>10–13</sup> In redox loops translocating protons across the membrane directly through a quinone/quinol pair (or pool), as in the Q-cycle,<sup>14,15</sup> at least two events of Q permeation, or transversal diffusion, occur. In the mitochondrial ETC, quinol production in respiratory complexes I, II

and III take place at the membrane N-side, but the Q<sub>o</sub> site for quinol oxidation in complex III is at the P-side.<sup>16–18</sup> Alternative oxidases cytochrome *bo*<sup>19</sup> and cytochrome *bd*<sup>20</sup> in bacteria also carry out quinol oxidation at the membrane P-side. Thus,

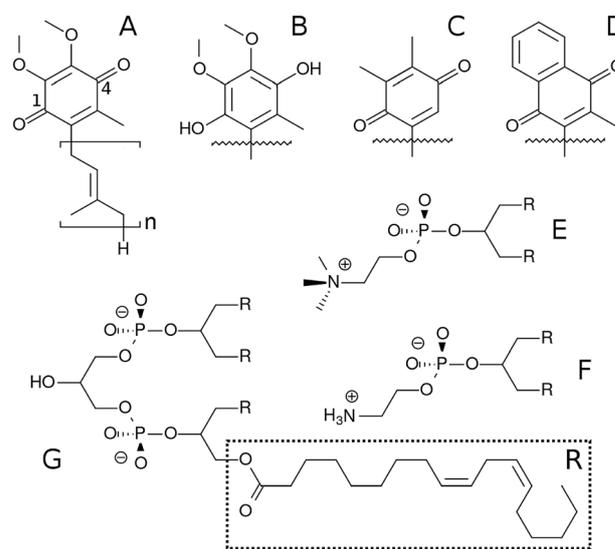


Fig. 1 Chemical structure of Q molecules and lipids studied here. (A) is ubiquinone (UQ<sub>n</sub>), (B) is ubiquinol (UQ<sub>n</sub>H<sub>2</sub>), (C) is plastoquinone (PQ<sub>n</sub>), (D) is menaquinone (MQ<sub>n</sub>). The *n* subscript is the number of bound isoprenoid units, shown only for UQ<sub>n</sub> which also displays positions 1 and 4 of the Q-ring. (E) is 1,2-dilinoleoyl-*sn*-glycero-3-phosphatidylcholine (DLPC), where R is the linoleoyl (18:2) acyl chain. (F) is 1,2-dilinoleoyl-*sn*-glycero-3-phosphatidylethanolamine (DLPE) and (G) is 1'-3'-bis[1,2-dilinoleoyl-*sn*-glycero-3-phospho]-*sn*-glycerol (cardiolipin, LCL).

Department of Biochemistry, Instituto de Química, Universidade de São Paulo, Av. Prof. Lineu Prestes 748, 05508-900, São Paulo, SP, Brazil. E-mail: garantes@iq.usp.br

† Electronic supplementary information (ESI) available: Additional analysis of convergence, lipid mixture and local diffusion, and atom types and charges used in simulations. See DOI: 10.1039/c9ra01681c

‡ Abbreviations: Q is used in general for quinone or quinol analogues with any oxidation state and number of isoprenoid units. Q-head is the aromatic ring and bound groups except for the Q-tail, which is the full isoprenoid chain. UQ<sub>n</sub> is oxidized ubiquinone, PQ<sub>n</sub> is plastoquinone and MQ<sub>n</sub> is menaquinone, where the *n* subscript is the number of bound isoprenoid units. UQ<sub>n</sub>H<sub>2</sub> is reduced ubiquinol. POPC is 16:0,18:1 phosphatidylcholine, DLPC is di-18:2 phosphatidylcholine, DLPE is di-18:2 phosphatidylethanolamine, LCL is tetra-18:2 cardiolipin, ETC is electron transfer chain, MD is molecular dynamics, US is umbrella sampling and COM is center-of-mass.



quinone/quinol recycling depends on Q permeating the membrane back-and-forth and it is important to understand this molecular process in detail.

Most of these proposals focused on Q lateral diffusion over the membrane, a fast and essentially diffusional process.<sup>21</sup> But transversal diffusion is an activated process and hence much slower.<sup>22–25</sup> Unfavorable desolvation and dielectric interactions (with associated energetic barriers) have to be overcome when the polar Q-head penetrates the lipid bilayer.<sup>18,26</sup>

Membrane distribution and lateral diffusion of Q have been experimentally studied (and debated) for decades.<sup>21,25,27–32</sup> Transversal diffusion is harder to measure accurately and has received less attention.<sup>24,25</sup> Molecular dynamics (MD) simulation is a powerful technique to probe both membrane distribution and permeation, but the quality of the results heavily depend on the force-field description of molecular interactions.<sup>33,34</sup> Recently, MD simulations were used to investigate the permeation of ubiquinone on a 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidylcholine (POPC) single-component bilayer and found the Q-head preferentially localizes near lipid glycerol groups.<sup>26</sup>

Lipid composition modulates the properties of membranes and embedded molecules.<sup>35,36</sup> For instance, the fraction of unsaturation (double bond) in lipid acyl chains changes membrane fluidity and permeability so that cells have evolved to homeostatically control this property through the biosynthesis of fatty acids.<sup>36,37</sup> Bioenergetic membranes are enriched with the anionic cardiolipin and its role on the distribution of ubiquinone in a more realistic multi-component lipid bilayer has been explored by simulation.<sup>38</sup> The distribution of natural quinones in membranes with mixed lipid composition was also investigated with more approximate coarse-grained simulations.<sup>39,40</sup>

Here, we employ MD simulations with an all-atom force-field carefully calibrated for Q molecules<sup>26</sup> to investigate the effects of lipid composition and unsaturation of acyl chains on Q distribution and permeation using a multi-component bilayer mimicking membranes involved in bioenergetic processes. Ubiquinone, ubiquinol, menaquinone and plastoquinone (Fig. 1) are studied to probe the role of the Q-head on membrane localization.

## Methods

### Set-up for molecular dynamics simulations

Six model bilayers were studied here, composed by 240 DLPC (Fig. 1), 208 DLPE, 64 cardiolipins (LCL di-anion), 16 Q (UQ<sub>2</sub>, UQ<sub>2</sub>H<sub>2</sub>, UQ<sub>10</sub>, UQ<sub>10</sub>H<sub>2</sub>, PQ<sub>9</sub> or MQ<sub>9</sub>) molecules symmetrically divided between each leaflet, 128 Na<sup>+</sup> cations and 21 728 water molecules. This is the same lipid composition used previously by Róg *et al.*<sup>38</sup> In umbrella sampling simulations, 37 435 water molecules were used to allow for sampling of complete UQ<sub>2</sub> partitioning into the aqueous phase.

The initial configuration for the UQ<sub>10</sub>H<sub>2</sub> bilayer was kindly provided by Prof. Tomasz Róg as used in their publication.<sup>38</sup> Initial geometries for all other bilayers were derived from this configuration by deleting or “mutating” Q-head atoms and adapting the atomic connectivity accordingly.

Force-field parameters for lipids and Na<sup>+</sup> were taken from CHARMM36.<sup>41</sup> Water was described by TIP3P.<sup>42</sup> Our CHARMM<sup>43,44</sup>

compatible force-field previously obtained by careful calibration<sup>26</sup> was used for Q molecules. Additional parameters necessary for MQ and PQ heads were taken directly from the CHARMM specification with partial charges and atom types shown in Table S1 and Fig. S1 (see ESI†). Complete topologies and parameters are available from the authors upon request.

All MD simulations were carried out with GROMACS<sup>45</sup> versions 4.6.7 for equilibrium trajectories and 5.1.3 for free-energy calculations. The NPT ensemble was used and temperature kept at 310 K with the Bussi thermostat<sup>46</sup> and a coupling constant of 0.1 ps with two separate coupling groups (water and everything else). Pressure was kept at 1.0 bar with a weak coupling scheme with a compressibility of 0.5–1.0 10<sup>−5</sup> bar<sup>−1</sup> and a coupling constant of 1 ps. Semi-isotropic coupling in the direction normal to the bilayer was applied. Electrostatics were handled by PME method<sup>47</sup> with a real space cutoff of 1.2 nm, grid spacing of 0.13 nm and quartic interpolation. All covalent hydrogen bonds were constrained using LINCS<sup>48</sup> and van der Waals interactions were truncated from 1.0 to 1.2 nm. No dispersion corrections were applied.<sup>49</sup> The integration time step was set to 2 fs and the nonbonded list was updated every 20 fs.

Unrestrained MD simulations were performed for 300 ns, frames were recorded every 20 ps and the initial 50 ns were discarded to allow for equilibration in the trajectories analyzed here. Mean area was computed as the ratio between the average area of the membrane plane and the number of lipid heavy atoms per leaflet. Atoms instead of molecules were used to allow comparison with multi-component bilayers. Normalized mass densities were calculated to ease comparison between groups with different number of atoms. Contacts between Q atoms and solvent molecules were defined with a 0.3 nm cut-off. Order parameters for methylene units of acyl chains were also calculated.<sup>50</sup> The convergence of calculated properties was checked as shown in Fig. S2.†

### Free energy calculations

Umbrella sampling (US)<sup>51</sup> was used to compute the free energy profile for UQ<sub>2</sub> permeation across the membrane normal. The distance between the UQ<sub>2</sub> head center-of-mass (COM) and the membrane COM along its normal (*z*-axis) was used as the reaction coordinate. The membrane COM was computed from a sum over the lipid atoms within a cylinder centered in the quinone and a 2.0 nm radius. This cylinder scheme helps to avoid artifacts due to membrane undulations, which are common in large membranes.<sup>52</sup>

Each simulation box contained 16 UQ<sub>2</sub> molecules, so 16 different windows were computed simultaneously in each US simulation. To avoid spurious interactions between UQ<sub>2</sub> molecules (for instance, aggregation when in the water phase), an artificial repulsive interaction was included between each Q-head C6 atom pair, using a specific Lennard-Jones pair with zeroed dispersion. Initial configurations for US were obtained at *t* ~ 200 ns from the UQ<sub>2</sub> unrestrained MD trajectory described above.

US windows were chosen equally spaced by 0.250 nm in the ranges *z* = [0.125,1.375] nm and [3.10,4.35] nm, and spaced by 0.125 nm in the range *z* = [1.500,2.875] nm. The umbrella



potential was set with  $k_{\text{umb}} = 1000 \text{ kJ mol}^{-1} \text{ nm}^{-1}$  and a total of 48 windows (obtained from 3 separate US simulations) were used to cover UQ<sub>2</sub> permeation over the complete membrane normal (24 windows for each leaflet). Each US simulation was ran for 300 ns. The reaction coordinate was recorded every 0.2 ps. Free energy profile was symmetrized and calculated with WHAM.<sup>53,54</sup> The initial 30 ns were discarded for equilibration and the remaining 270 ns were used for data accumulation on each window. Statistical uncertainties were estimated as 95% confidence intervals by bootstrap analysis.<sup>55</sup>

## Results & discussion

### Quinone composition and multi-component lipid membranes do not change the Q-head equilibrium distribution

Results from unrestrained molecular dynamics simulations for six membranes with different Q composition are presented in this section. Mass density profiles for lipid groups are shown in Fig. 2. Density distributions in each leaflet are fairly symmetrical (negative or positive membrane normal) which indicates good equilibration and sampling of MD trajectories. Densities of phosphate (P) and acyl glycerol (shown only in Fig. 2A) groups are unaltered in all membranes and suggest that Q composition does not perturb the membrane structure. The multi-component membrane is well mixed and no lipid clusters or domains are found (Fig. S3†).

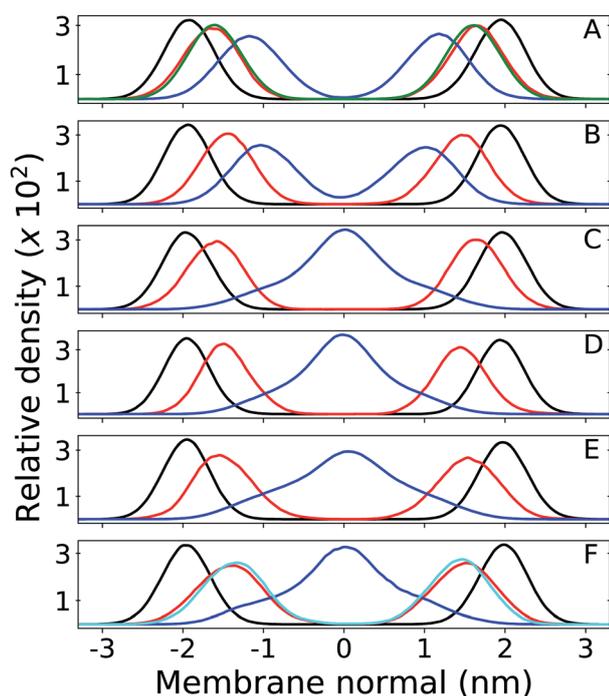


Fig. 2 Relative mass density of multi-component bilayers with different quinone analogue (Q) composition. Phosphate group of DLPC + DLPE is shown in black, Q-head in red and Q-tail in blue. Panel (A) shows the bilayer containing ubiquinone-2 (UQ<sub>2</sub>), where the green profile indicates density of the DLPC + DLPE glycerol group; (B) is ubiquinol-2 (UQ<sub>2</sub>H<sub>2</sub>); (C) is ubiquinone-10 (UQ<sub>10</sub>); (D) is ubiquinol-10 (UQ<sub>10</sub>H<sub>2</sub>); (E) is plastoquinone-9 (PQ<sub>9</sub>) and (F) is menaquinone-9 (MQ<sub>9</sub>), where the cyan profile indicates density of the first Q-ring (6C atoms) only and the full Q-head (10C atoms) is in red. Zero of the membrane normal corresponds to the bilayer center. Data were not symmetrized between the two leaflets.

Density peaks for the Q-head change only 0.3 nm (from 1.4 to 1.7 nm in modulus) between all Q. Density peaks indicate the Q-head equilibrium position across the membrane normal and are in agreement with equilibrium positions previously determined for ubiquinone and ubiquinol in POPC membranes.<sup>26</sup> Q-head density distributions also do not depend on isoprenoid chain length (for instance, compare Fig. 2A and C). For short isoprenoid chains ( $n = 2$ ), the Q-tail density localizes around 1 nm. For longer chains ( $n = 9,10$ ), the Q-tail density peaks at the membrane center and spreads through most of the lipid phase. UQ<sub>n</sub>H<sub>2</sub> ( $n = 2,10$ ) is slightly more buried, with Q-head peak density 0.2 nm lower than ubiquinone and with more penetration of Q-tail density towards the membrane center. This behavior was also observed in POPC membranes and is a consequence of the intramolecular hydrogen bonds (H-bond) often observed between phenolic hydrogen and methoxy oxygen in ubiquinol.<sup>26,56</sup> These H-bonds stabilize the Q-head when it is desolvated and allow for a higher permeation of UQ<sub>n</sub>H<sub>2</sub>.

Fig. 3 shows contacts performed by the Q-head O1 with different lipid groups and solvent. H-bond acceptors display shorter water contacts as the minimum distance shown in Fig. 3E is computed between O1 and any of the water atoms. More interestingly, MQ<sub>9</sub> shows the smallest first hydration peak and only UQ<sub>10</sub> and MQ<sub>9</sub> show a secondary water peak at a distance of  $\sim 0.4$  nm, corresponding to configurations that O1 is not forming an H-bond with water but the other ketonic O4 is. Similar distributions but with smaller differences between Q molecules are observed for contacts with the Q-head O4, except

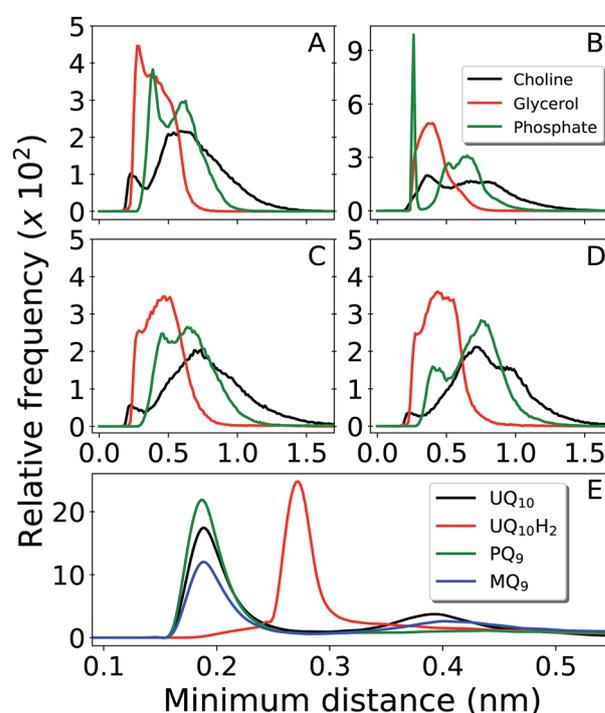


Fig. 3 Distribution of minimum distances around the Q-head oxygen O1 of multi-component bilayers with different Q composition. Contacts with DLPC choline group are shown in black, DLPC + DLPE glycerol in red and DLPC + DLPE phosphate in green. Panel (A) is UQ<sub>10</sub>, (B) is UQ<sub>10</sub>H<sub>2</sub>, (C) is PQ<sub>9</sub> and (D) is MQ<sub>9</sub>. Panel (E) shows water contacts for UQ<sub>10</sub> in black, UQ<sub>10</sub>H<sub>2</sub> in red, PQ<sub>9</sub> in green and MQ<sub>9</sub> in blue. Data obtained from unrestrained MD simulations.



for MQ<sub>9</sub>, which is significantly more hydrated in O4 (data not shown here).

The H-bond donor UQ<sub>10</sub>H<sub>2</sub> frequently makes an H-bond with the anionic phosphate group in DLPC + DLPE (peak at 0.26 nm at Fig. 3B) or in cardiolipins (not shown). H-bond acceptors (UQ<sub>10</sub>, PQ<sub>9</sub> and MQ<sub>9</sub>) do not show frequent contacts with phosphates, neither interact closely with the positive choline groups in DLPC (Fig. 3A, C and D). Almost no H-bonds are observed between any of the Q molecules and either the ethylamine group in DLPE or the bridge glycerol in cardiolipins (not shown). Instead, the Q-head is located near acyl glycerol groups (Fig. 2A). This is very similar to the equilibrium location and contacts performed by Q in a POPC bilayer.<sup>26</sup>

Results for MQ<sub>9</sub> suggest the two rings in the naphthoquinone head have similar average location across the membrane normal (Fig. 2D), but are probably orthogonal to the bilayer plane with O1 less solvated (pointing towards the membrane) and the opposite O4 more often in contact with water (pointing to the water phase). For PQ<sub>9</sub>, the Q-tail density has the lowest peak and spreads more evenly than other Q with long tails. Thus, PQ<sub>9</sub> may have more conformational flexibility which explains the lack of secondary solvation peak (Fig. 3E), without disturbing the average Q-head location.

In comparison to our previous simulations of Q in POPC membranes,<sup>26</sup> both density profiles and contacts performed by the Q-head do not change significantly in this multi-component membrane. The Q-head preferentially resides at the bilayer interface, near the membrane normal occupied by either POPC or DLPC + DLPE glycerol groups (Fig. 2A). No specific contacts have been found between Q and polar lipid heads, except for H-bonds between ubiquinol and phosphate groups which are present in any phospholipid. The equilibrium distribution and location of Q on the membrane models studied so far result from the balance of hydrophobic interactions by the Q-tail and hydration of the Q-head, corresponding to a typical amphiphilic molecule.<sup>57</sup>

### Increased membrane permeability and penetration of water bound to the Q-head

Free energy profiles and analysis obtained from umbrella sampling simulations for UQ<sub>2</sub> permeation across the

membrane are presented in this section. An analogue with only two isoprenoid units was used because it is computationally more efficient to sample and reliable experimental partition coefficients are available for comparison.<sup>58,59</sup>

The calculated free energy profile in Fig. 4A shows three important aspects: the minimum across the membrane normal, which represents the equilibrium position of the Q-head, is located at 1.75 nm as expected from the Q-head density peak in Fig. 2A and close to the minimum observed for UQ<sub>2</sub> permeation in the POPC bilayer (1.65 nm).<sup>26</sup> The insertion or binding free energy for UQ<sub>2</sub> in the multi-component membrane is  $-22 \pm 2$  kJ mol<sup>-1</sup>, in excellent agreement with experiment (Table 1) and with the binding free energy in POPC ( $-21 \pm 1$  kJ mol<sup>-1</sup>).<sup>26</sup> However, the free energy barrier for UQ<sub>2</sub> permeation in the multi-component membrane, given by the free energy difference between the minimum (at 1.75 nm) and the maximum near the center of the bilayer (0 nm), is only  $19 \pm 1$  kJ mol<sup>-1</sup>, much lower than the barrier computed for permeation in POPC ( $47 \pm 2$  kJ mol<sup>-1</sup>).<sup>26</sup>

Fig. 4B shows the permeation barrier converges to within one statistical uncertainty in 200 ns of simulation (170 ns of data acquisition) per US window. In total, 50% more data (300 ns) were collected in each window, supporting the quality of the barrier computed here. Experimental estimates of rates for Q permeation or “flip-flop” by NMR linewidth<sup>25</sup> and redox titration of substrates trapped inside vesicles<sup>24</sup> only suggested rate upper bounds and for lipids with fully saturated acyl chains. To our knowledge, there are not measurements of Q transversal diffusion for bilayers composed of unsaturated acyl chains to compare with the properties determined here. But, we also calculated a local transversal diffusion<sup>60</sup> and a permeation coefficient  $P = 20$  cm s<sup>-1</sup> (Fig. S4†). This is 4 to 5 fold the  $P$  value obtained from simulation of polar aromatic molecules such as phenol and benzoic acid, which have similar (15–20 kJ mol<sup>-1</sup>) permeation free energy barriers, in dioleoyl phosphatidylcholine membranes.<sup>61</sup>

Thus, Q permeation is observed more frequently in the multi-component membrane than in POPC. Although we have not computed profiles for permeation of the other Q studied here, spontaneous flip-flops were observed for UQ<sub>2</sub>H<sub>2</sub>, PQ<sub>9</sub> and MQ<sub>9</sub>, at least once during the 300 ns of unrestrained MD simulations presented in the previous section. Given that UQ<sub>n</sub> did not show spontaneous flip-flops,

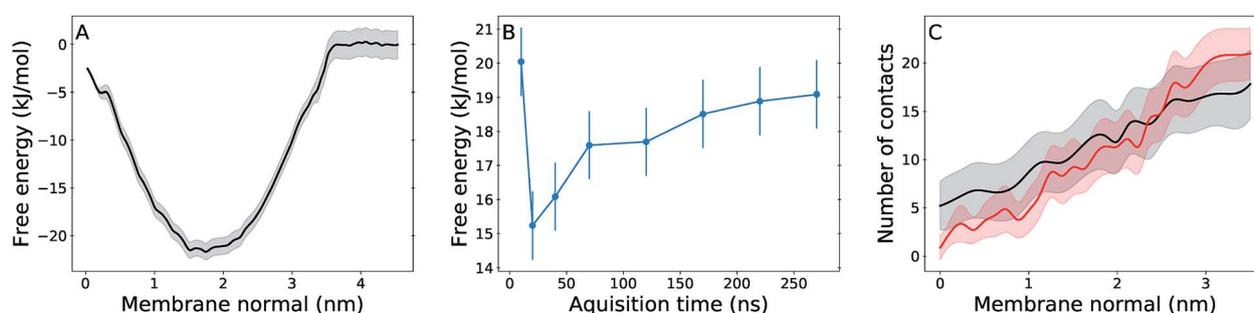


Fig. 4 Free energy and hydration profiles for UQ<sub>2</sub> permeation. Panel (A) shows the free energy profile with statistical uncertainty (gray shadow) for UQ<sub>2</sub> insertion across the membrane normal. (B) shows the convergence of the permeation barrier with increasing acquisition time and fixed 30 ns of initial equilibration discarded from analysis. (C) shows the average number of water contacts with the Q-head for permeation in the multi-component membrane (black) and in the pure POPC membrane (red),<sup>26</sup> with standard deviations shown in shadow. Data obtained from umbrella sampling simulations and symmetrized between the two leaflets.



**Table 1** Experimental and simulated binding free energies (in  $\text{kJ mol}^{-1}$ ) between ubiquinone analogues and multi-component membranes. Experimental free energies were obtained from partition coefficients assuming a temperature of 310 K. Values from Róg *et al.*<sup>38</sup> were estimated from their Fig. 3

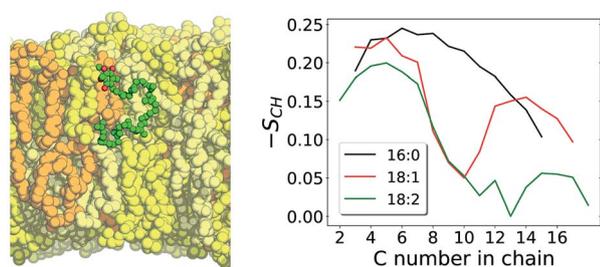
Molecule	Experimental	Róg <sup>38</sup>	Here
UQ <sub>1</sub>	-17 (ref. 59), -20 (ref. 58)	-34 ± 4	
UQ <sub>2</sub>	-21 (ref. 58), -24 (ref. 59)		-22 ± 2

we can speculate that the other Q analogues will show smaller permeation barriers.

Then, what is different in this multi-component membrane that allows permeation more frequently? The mean area per lipid heavy atom is  $0.0126 \text{ nm}^2$  per atom in the multi-component membrane *versus*  $0.0119 \text{ nm}^2$  per atom in POPC.<sup>26</sup> The two unsaturations (positions 9 and 12) in the linoleoyl acyl chain composing all lipids in the multi-component membrane result in a less ordered lipid phase, as shown by lower methylene order parameters computed for linoleoyl chain in comparison to palmitoyl and oleoyl in POPC (Fig. 5). Lower lipid ordering result in a more porous membrane in which molecules can penetrate with a cheaper cavitation energy cost.<sup>22,62</sup>

Accordingly, more water contacts and H-bonds with the Q-head are performed inside the lipid phase for UQ<sub>2</sub> permeation in the multi-component membrane (normal <1 nm in Fig. 4C), leading to less desolvation of the polar Q-head and a lower permeation barrier than in POPC. Water arrest inside the lipid phase during the permeation of polar solutes has been observed in other simulations.<sup>33,63</sup> It should be noted that no water penetration is observed in equilibrium for both POPC or multi-component membranes. But in the rare event of Q permeation, it is easier for water bound to the Q-head to penetrate into a membrane composed by fatty acids with a higher fraction of unsaturations.

A comparison with the simulations performed by Róg *et al.* shows large and important qualitative differences, specially for ubiquinone which was reported to make no contacts with solvent or lipid head groups and localize in the center of the bilayer.<sup>38</sup> It seems unlike that the polar head group in UQ<sub>n</sub>, with four oxygen H-bond acceptors, will be stable in the middle of the lipid phase. Their calculated binding free energy for UQ<sub>1</sub> is in large



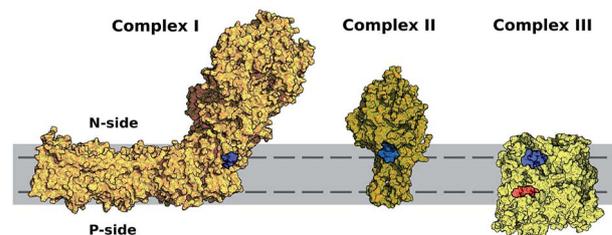
**Fig. 5** Left panel shows a structural snapshot of the multi-component membrane with UQ<sub>10</sub> in its equilibrium position. LCL is shown in orange, DLPC in yellow, DLPE in beige and UQ<sub>10</sub> in green (carbons) and red (oxygens). Right panel shows carbon-hydrogen bond vector order parameters ( $S_{\text{CH}}$ ) calculated with unrestrained MD for acyl chains POPC palmitoyl in black and oleoyl in red,<sup>26</sup> and for DLPC + DLPE linoleoyl acyl chains in green.

disagreement with experimental partition coefficients (Table 1). For ubiquinol, Q-head location and contacts are similar to ours, but strangely depend on the number of isoprenoid units. The lack of symmetry between the two leaflets in the density profiles shown and the irregular dependence of calculated densities and contacts with the number of isoprenoid units question whether appropriate equilibration and sampling were reached in their work.<sup>38</sup> Exactly the same lipid composition and similar simulation procedures were employed here. Thus, most of the qualitative differences found should be due to force-fields. The OPLS derived force-field used by Róg *et al.* has been shown to artificially overbind  $\text{Na}^+$  cations<sup>64</sup> and describe lipid order parameters in disagreement with NMR measurements.<sup>65</sup> On the other hand, the CHARMM36 force-field used here has consistently one of the best performances describing these properties as well as several other membrane structural, thermodynamical and diffusional observables,<sup>41,66,67</sup> including binding free energies and lateral diffusion constants for Q molecules.<sup>26</sup>

### Transversal diffusion of Q may limit turnover of electron transfer chains

Fig. 6 shows the equilibrium position of the Q-head in the multi-component membrane also matches the normal position (in modulus) of entrance sites for Q binding in respiratory complexes, as noted previously for the POPC bilayer.<sup>26</sup> An initial encounter complex formed after membrane diffusion of Q and collision into a respiratory complex will frequently have the appropriate geometry in the membrane normal. Fig. 6 also shows membrane permeation has to occur for quinone/quinol recycling as their sites of reduction/oxidation are placed at opposite membrane sides.

Bilayers with a low fraction of unsaturated chains, such as POPC,<sup>26</sup> may slow down transversal diffusion by up to five orders of magnitude, in comparison to the bilayer with 2 unsaturations per chain studied here. This steep relation is due to the activated nature of the permeation process.<sup>22,23</sup> A low fraction of unsaturation (<20%) may slow Q permeation up to a point it becomes rate-limiting for ETC or even blocks cellular growth.<sup>13</sup> Yet another reason for the high content of cardiolipins, composed mostly by unsaturated acyl chains, in membranes involved in biological energy transduction<sup>71</sup> might be the increment of Q permeability gained by lipid unsaturation.



**Fig. 6** Entrance sites for Q binding in respiratory complexes I–III. Sites for quinone reduction shown in blue and for quinol oxidation in red. Bilayer is represented in solid gray and equilibrium or preferential position of the Q-head across the membrane normal ( $\pm 1.75 \text{ nm}$ ) shown with a dashed line. Structures were taken from PDB ID 4HEA<sup>68</sup> for complex I, 2H88 (ref. 69) for complex II and 2QJP<sup>70</sup> for complex III.



## Conclusions

Molecular dynamics simulations for ubiquinone, ubiquinol, plastoquinone and menaquinone and free energy calculations for permeation of UQ<sub>2</sub> in a multi-component bilayer mimicking the composition of bioenergetic membranes were presented.

Densities of Q-head and Q-tail (of similar length) across the membrane normal change little between each Q studied. The Q-head preferentially localizes near acyl glycerol groups and no specific contacts between the Q-head and polar lipid heads were observed, except for H-bonds formed between ubiquinol and phosphate groups. This is equivalent to the localization and contacts found for Q in a POPC bilayer. The binding free energy and equilibrium position estimated for UQ<sub>2</sub> partition into the multi-component bilayer are also in agreement to what was found in POPC.<sup>26</sup>

We may conclude the equilibrium distribution of Q across model bilayers recently studied by molecular simulation<sup>26,38</sup> do not change significantly with lipid composition, even when containing the anionic cardiolipin. The equilibrium structure and thermodynamics of embedded Q are determined by Q-head hydration and the hydrophobic effect of the isoprenoid tail. This should not be entirely unexpected given that similar partition coefficients have been measured for ubiquinone (with the same number of isoprenoid units) in biological membranes with varied composition.<sup>58,59</sup>

However, permeation of Q across the membrane changes considerably with lipid composition, specially the fraction of acyl chain unsaturation. Lipids with two unsaturations per acyl chain are less ordered, giving a more porous bilayer. Thus, the polar Q-head may carry water molecules during a flip-flop event with a lower free energy barrier, resulting in much faster transversal diffusion across the multi-component membrane studied here.

Membrane permeability modulated by the fraction of acyl chain unsaturation is important for efficient quinone/quinol recycling in opposite membrane sides and for translocation of protons by Q redox loops. Permeability may even limit turnover rates for biological ETC if the fraction of lipid unsaturation is too low. Results in the multi-component bilayer also confirm our previous observation that the equilibrium position of the Q-head across the membrane normal is similar to the position of protein sites for entrance of Q into respiratory complexes.<sup>26</sup> Finally, it should be mentioned that lipid composition may also influence the structure of bioenergetic membranes<sup>72</sup> and of respiratory complexes<sup>73</sup> and their reactivity.<sup>74–76</sup>

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

Funding from FAPESP (project 16/24096-5), CNPq (fellowship number 142112/2015-1) and computational resources from the SDumont cluster in the National Laboratory for Scientific Computing (LNCC/MCTI) are gratefully acknowledged.

## References

- 1 P. Mitchell, Chemiosmotic coupling in oxidative and photosynthetic phosphorylation, *Biochim. Biophys. Acta*, 2011, **1807**, 1507–1538.
- 2 G. W. Canters and E. Vlijgenboom, *Biological Electron Transfer Chains: Genetics, Composition and Mode of Operation*, Springer Netherlands, Dordrecht, 1998.
- 3 R. E. Blankenship, *Molecular Mechanisms of Photosynthesis*, Wiley-Blackwell, USA, 1st edn, 2002.
- 4 D. G. Nicholls and S. J. Ferguson, *Bioenergetics*, Academic Press, Boston, 4th edn, 2013.
- 5 M. Iwai, K. Takizawa, R. Tokutsu, A. Okamoto, Y. Takahashi and J. Minagawa, Isolation of the elusive supercomplex that drives cyclic electron flow in photosynthesis, *Nature*, 2010, **464**, 1210–1213.
- 6 R. Guo, J. Gu, S. Zong, M. Wu and M. Yang, Structure and mechanism of mitochondrial electron transport chain, *Biomed. J.*, 2018, **41**, 9–20.
- 7 F. J. Van Eerden, M. N. Melo, P. W. Frederix, X. Periole and S. J. Marrink, Exchange pathways of plastoquinone and plastoquinol in the photosystem II complex, *Nat. Commun.*, 2017, **8**, 1–8.
- 8 J. Warnau, V. Sharma, A. P. Gamiz-Hernandez, A. D. Luca, O. Haapanen, I. Vattulainen, M. Wikstrom, G. Hummer and V. R. I. Kaila, Redox-coupled quinone dynamics in the respiratory complex I, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, E8413–E8420.
- 9 M. H. Teixeira and G. M. Arantes, Balanced internal hydration discriminates substrate binding to respiratory complex I, *bioRxiv*, 2019, 572404.
- 10 P. R. Rich, A Generalised Model for the Equilibration of Quinones Pools with their Biological Donors and Acceptors in Membrane-Bound Electron Transfer Chains, *FEBS Lett.*, 1981, **130**, 173–178.
- 11 B. Chazotte and C. R. Hackenbrock, Lateral Diffusion as a Rate-limiting Step in Ubiquinone-mediated Mitochondrial Electron Transport, *J. Biol. Chem.*, 1989, **264**, 4978–4985.
- 12 G. Lenaz and M. L. Genova, Mobility and function of Coenzyme Q (ubiquinone) in the mitochondrial respiratory chain, *Biochim. Biophys. Acta*, 2009, **1787**, 563–573.
- 13 I. Budin, T. de Rond, Y. Chen, L. J. G. Chan, C. J. Petzold and J. D. Keasling, Viscous control of cellular respiration by membrane lipid composition, *Science*, 2018, **362**, 1189.
- 14 P. Mitchell, The protonmotive Q cycle: a general formulation, *FEBS Lett.*, 1975, **59**, 137–139.
- 15 M. Wikström, V. Sharma, V. R. I. Kaila, J. P. Hosler and G. Hummer, New Perspectives on Proton Pumping in Cellular Respiration, *Chem. Rev.*, 2015, **115**, 2196–2221.
- 16 K. Fiedorczuk, J. A. Letts, G. Degliesposti, K. Kaszuba, M. Skehel and L. A. Sazanov, Atomic structure of the entire mammalian mitochondrial complex I, *Nature*, 2016, **538**, 406–410.
- 17 E. Maklashina and G. Cecchini, The quinone-binding and catalytic site of complex II, *Biochim. Biophys. Acta*, 2010, **1797**, 1877–1882.



- 18 W. Cramer, S. S. Hasan and E. Yamashita, The Q cycle of cytochrome bc complexes: A structure perspective, *Biochim. Biophys. Acta*, 2011, **1807**, 788–802.
- 19 J. Abramson, S. Riistama, G. Larsson, A. Jasaitis, M. Svensson-Ek, L. Laakkonen, A. Puustinen, S. Iwata and M. Wikström, The structure of the ubiquinol oxidase from *Escherichia coli* and its ubiquinone binding site, *Nat. Struct. Biol.*, 2000, **7**, 910–917.
- 20 S. Safarian, C. Rajendran, H. Müller, J. Preu, J. D. Langer, S. Ovchinnikov, T. Hirose, T. Kusumoto, J. Sakamoto and H. Michel, Structure of a bd oxidase indicates similar mechanisms for membrane-integrated oxygen reductases, *Science*, 2016, **352**, 583–586.
- 21 M. Afri, B. Ehrenberg, Y. Talmon, J. Schmidt, Y. Cohena and A. A. Frimer, Location and Mobility of Ubiquinones of Different Chain Lengths in Artificial Membrane Vesicles, *Chem. Phys. Lipids*, 2004, **131**, 107–121.
- 22 R. J. Pace and S. I. Chan, Molecular motions in lipid bilayers. III. Lateral and transverse diffusion in bilayers, *J. Chem. Phys.*, 1982, **76**, 4241–4247.
- 23 A. Imparato, J. C. Shillcock and R. Lipowsky, Lateral and transverse diffusion in two-component bilayer membranes, *Eur. Phys. J. E*, 2002, **11**, 21–28.
- 24 A. Futami, E. Hurt and G. Hauska, Vectorial redox reactions of physiological quinones I. Requirement of a minimum length of the isoprenoid side chain, *Biochim. Biophys. Acta*, 1979, **547**, 583–596.
- 25 P. B. Kingsley and G. W. Feigenson, 1H-NMR study of the location and motion of ubiquinones in perdeuterated phosphatidylcholine bilayers, *Biochim. Biophys. Acta*, 1981, **635**, 602–618.
- 26 V. V. Galassi and G. M. Arantes, Partition, orientation and mobility of ubiquinones in a lipid bilayer, *Biochim. Biophys. Acta*, 2015, **1847**, 1345.
- 27 H. Katsikas and P. Quinn, The Polyisoprenoid Chain Length Influences the Interaction of Ubiquinones with Phospholipid Bilayers, *Biochim. Biophys. Acta*, 1982, **689**, 363–369.
- 28 B. A. Cornell, M. A. Keniry, A. Post, R. N. Robertson, L. E. Weir and P. W. Westerman, Location and Activity of Ubiquinone 10 and Ubiquinone Analogues in Model and Biological Membranes, *Biochemistry*, 1987, **26**, 7702–7707.
- 29 G. Lenaz, B. Samori, R. Fato, M. Battino, G. P. Castelli and I. Domini, Localization and Preferred Orientations of Ubiquinone Homologs in Model Bilayers, *Biochem. Cell Biol.*, 1992, **70**, 504–514.
- 30 M. Jemiola-Rzeminska, J. Kruk, M. Skowronek and K. Strzalka, Location of Ubiquinone Homologues in Liposome Membranes Studied by Fluorescence Anisotropy of Diphenyl-Hexatriene and Trimethylammonium-Diphenyl-Hexatriene, *Chem. Phys. Lipids*, 1996, **79**, 55–63.
- 31 S. Di Bernardo, R. Fato, R. Casadio, P. Fariselli and G. Lenaz, A High Diffusion Coefficient for Coenzyme Q1H Might be Related to a Folded Structure, *FEBS Lett.*, 1998, **426**, 77–80.
- 32 I. Llorente-Garcia, *et al.*, Single-Molecule in Vivo Imaging of Bacterial Respiratory Complexes Indicates Delocalized Oxidative Phosphorylation, *Biochim. Biophys. Acta*, 2014, **1837**, 811–824.
- 33 C. Neale and R. Pomès, Sampling errors in free energy simulations of small molecules in lipid bilayers, *Biochim. Biophys. Acta*, 2016, **1858**, 2539–2548.
- 34 D. J. Smith, J. B. Klauda and A. J. Sodt, Simulation Best Practices for Lipid Membranes, *Living Journal of Computational Molecular Science*, 2018, **1**, 5966–5996.
- 35 S. Ladha, A. R. Mackie, L. J. Harvey, D. C. Clark, E. J. A. Lea, M. Brullemans and H. Duclouhier, Lateral Diffusion in Planar Lipid Bilayers: A Fluorescence Recovery after Photobleaching Investigation of Its Modulation by Lipid Composition, Cholesterol, or Alamethicin Content and Divalent Cations, *Bioophys. J.*, 1996, **71**, 1364–1373.
- 36 G. van Meer, D. R. Voelker and G. W. Feigenson, Membrane lipids: where they are and how they behave, *Nat. Rev. Mol. Cell Biol.*, 2008, **9**, 112–124.
- 37 S. Ballweg and R. Ernst, Control of membrane fluidity: the OLE pathway in focus, *Biol. Chem.*, 2017, **398**, 215–228.
- 38 P. Kaurola, V. Sharma, A. Vonk, I. Vattulainen and T. Róg, Distribution and dynamics of quinones in the lipid bilayer mimicking the inner membrane of mitochondria, *Biochim. Biophys. Acta*, 2016, **1858**, 2116–2122.
- 39 D. H. de Jong, N. Liguori, T. van den Berg, C. Arnarez, X. Periole and S. J. Marrink, Atomistic and Coarse Grain Topologies for the Cofactors Associated with the Photosystem II Core Complex, *J. Phys. Chem. B*, 2015, **119**, 7791–7803.
- 40 F. J. van Eerden, D. H. de Jong, A. H. de Vries, T. A. Wassenaar and S. J. Marrink, Characterization of thylakoid lipid membranes from cyanobacteria and higher plants by molecular dynamics simulations, *Biochim. Biophys. Acta*, 2015, **1848**, 1319–1330.
- 41 J. B. Klauda, R. M. Venable, J. A. Freites, J. W. O'Connor, D. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. MacKerell and R. W. Pastor, Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types, *J. Phys. Chem. B*, 2010, **114**, 7830–7843.
- 42 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.*, 1983, **79**, 926–935.
- 43 A. D. MacKerell Jr, *et al.*, All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins, *J. Phys. Chem. B*, 1998, **102**, 3586–3616.
- 44 K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov and A. D. Mackerell Jr, CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields, *J. Comput. Chem.*, 2010, **31**, 671–690.
- 45 S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess and E. Lindahl, GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit, *Bioinformatics*, 2013, **29**, 845–854.



- 46 G. Bussi, D. Donadio and M. Parrinello, Canonical sampling through velocity rescaling, *J. Chem. Phys.*, 2007, **126**, 014101.
- 47 T. Darden, D. York and L. Pedersen, Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems, *J. Chem. Phys.*, 1993, **98**, 10089–10092.
- 48 B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, LINCS: A Linear Constraint Solver for Molecular Simulations, *J. Comput. Chem.*, 1997, **18**, 1463–1472.
- 49 C. Anézo, A. H. de Vries, H. Höltje, D. P. Tieleman and S. Marrink, Methodological Issues in Lipid Bilayer Simulations, *J. Phys. Chem. B*, 2003, **107**, 9424–9433.
- 50 T. D. Romo, N. Leioatts and A. Grossfield, Lightweight Object Oriented Structure Analysis: Tools for Building Tools to Analyze Molecular Dynamics Simulations, *J. Comput. Chem.*, 2014, **35**, 2305–2318.
- 51 G. M. Torrie and J. P. Valleau, Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling, *J. Comp. Physiol.*, 1977, **23**, 187–199.
- 52 N. Nitschke, K. Atkovska and J. S. Hub, Accelerating potential of mean force calculations for lipid membrane permeation: System size, reaction coordinate, solute-solute distance, and cutoffs, *J. Chem. Phys.*, 2016, **145**, 125101.
- 53 S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen and P. A. Kollman, The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method, *J. Comput. Chem.*, 1992, **13**, 1011–1021.
- 54 B. Roux, The calculation of the potential of mean force using computer simulations, *Comput. Phys. Commun.*, 1995, **91**, 275–282.
- 55 R. W. Johnson, An introduction to the bootstrap, *Teach. Stat.*, 2001, **23**, 49–54.
- 56 P. R. Rich and R. Harper, Partition coefficients of quinone and hydroquinones and their relation to biochemical reactivity, *FEBS Lett.*, 1990, **269**, 139–144.
- 57 H. A. L. Filipe, M. J. Moreno, T. Róg, I. Vattulainen and L. M. S. Loura, How To Tackle the Issues in Free Energy Simulations of Long Amphiphiles Interacting with Lipid Membranes: Convergence and Local Membrane Deformations, *J. Phys. Chem. B*, 2014, **118**, 3572–3581.
- 58 R. Fato, M. Battino, M. D. Esposti, G. P. Castelli and G. Lenaz, 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*, 1986, **25**, 3378–3390.
- 59 R. Fato, E. Estornell, S. Di Bernardo, F. Pallotti, G. P. Castelli and G. Lenaz, Steady-State Kinetics of the Reduction of Coenzyme Q Analogs by Complex I (NADH:Ubiquinone Oxidoreductase) in Bovine Heart Mitochondria and Submitochondrial Particles, *Biochemistry*, 1996, **35**, 2705–2716.
- 60 G. Hummer, Position-dependent diffusion coefficients and free energies from Bayesian analysis of equilibrium and replica molecular dynamics simulations, *New J. Phys.*, 2005, **7**, 34.
- 61 M. Palaiokostas, W. Ding, G. Shahane and M. Orsi, Effects of lipid composition on membrane permeation, *Soft Matter*, 2018, **14**, 8496–8508.
- 62 J. Tomasi and M. Persico, Molecular Interactions in Solution: An Overview of Methods Based on Continuous Distributions of the Solvent, *Chem. Rev.*, 1994, **94**, 2027–2094.
- 63 C. Neale, W. F. D. Bennett, D. P. Tieleman and R. Pomés, Statistical Convergence of Equilibrium Properties in Simulations of Molecular Solutes Embedded in Lipid Bilayers, *J. Chem. Theory Comput.*, 2011, **7**, 4175–4188.
- 64 A. Catte, M. Girych, M. Javanainen, C. Loison, J. Melcr, M. S. Miettinen, L. Monticelli, J. Määttä, V. S. Oganessian, O. H. S. Ollila, J. Tynkkynen and S. Vilov, Molecular electrometer and binding of cations to phospholipid bilayers, *Phys. Chem. Chem. Phys.*, 2016, **18**, 32560–32569.
- 65 A. Botan, *et al.*, Toward Atomistic Resolution Structure of Phosphatidylcholine Headgroup and Glycerol Backbone at Different Ambient Conditions, *J. Phys. Chem. B*, 2015, **119**, 15045–15088.
- 66 D. Poger, B. Caron and A. E. Mark, Validating lipid force fields against experimental data: Progress, challenges and perspectives, *Biochim. Biophys. Acta*, 2016, **1858**, 1556–1565.
- 67 O. H. S. Ollila and G. Pabst, Atomistic resolution structure and dynamics of lipid bilayers in simulations and experiments, *Biochim. Biophys. Acta*, 2016, **1858**, 2512–2528.
- 68 R. Baradaran, J. M. Berrisford, G. S. Minhas and L. A. Sazanov, Crystal structure of the entire respiratory complex I, *Nature*, 2013, **494**, 443–448.
- 69 L.-S. Huang, J. T. Shen, A. C. Wang and E. A. Berry, Crystallographic studies of the binding of ligands to the dicarboxylate site of Complex II, and the identity of the ligand in the “oxaloacetate-inhibited” state, *Biochim. Biophys. Acta*, 2006, **1757**, 1073–1083.
- 70 L. Esser, M. Elberry, F. Zhou, C.-A. Yu, L. Yu and D. Xia, Inhibitor-complexed structures of the cytochrome bc<sub>1</sub> from the photosynthetic bacterium *Rhodobacter sphaeroides*, *J. Biol. Chem.*, 2008, **283**, 2846–2857.
- 71 J. Comte, B. Maisterrena and D. C. Gautheron, Lipid composition and protein profiles of outer and inner membranes from pig heart mitochondria, *Biochim. Biophys. Acta*, 1976, **419**, 271–284.
- 72 E. K. Eriksson, V. A. Hernández and K. Edwards, Effect of ubiquinone-10 on the stability of biomimetic membranes of relevance for the inner mitochondrial membrane, *Biochim. Biophys. Acta*, 2018, **1860**, 1205–1215.
- 73 S. S. Hasan, E. Yamashita, C. M. Ryan, J. P. Whitelegge and W. A. Cramer, Conservation of Lipid Functions in Cytochrome bc Complexes, *J. Mol. Biol.*, 2011, **414**, 145–162.
- 74 T. H. Haines and N. A. Dencher, Cardiolipin: a proton trap for oxidative phosphorylation, *FEBS Lett.*, 2002, **528**, 35–39.
- 75 P. A. Postila, K. Kaszuba, P. Kuleta, I. Vattulainen, M. Sarewicz, A. Osyczka and T. Róg, Atomistic determinants of co-enzyme Q reduction at the Q<sub>r</sub>-site of the cytochrome bc<sub>1</sub> complex, *Sci. Rep.*, 2016, **6**, 33607.
- 76 A. H. de Vries, A. E. Mark and S. J. Marrink, The Binary Mixing Behavior of Phospholipids in a Bilayer: A Molecular Dynamics Study, *J. Phys. Chem. B*, 2004, **108**, 2454–2463.



## Supporting Information

### Effects of lipid composition on membrane distribution and permeability of natural quinones

Murilo Hoias Teixeira and Guilherme Menegon Arantes

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

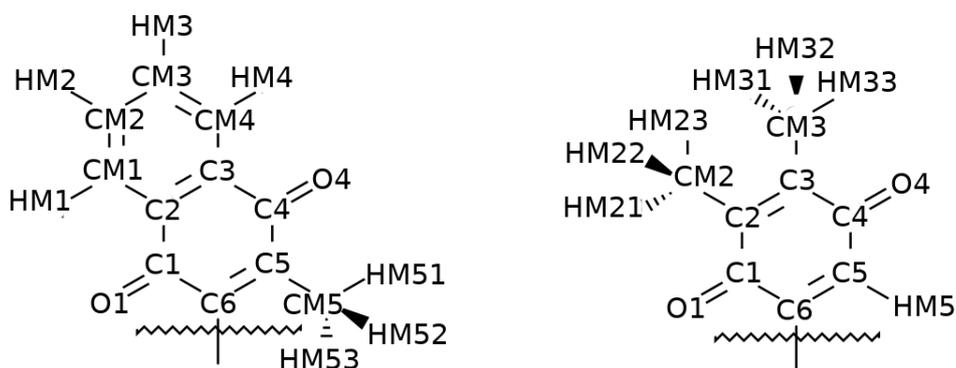


Figure S1: Atom names for PQ and MQ quinone heads.

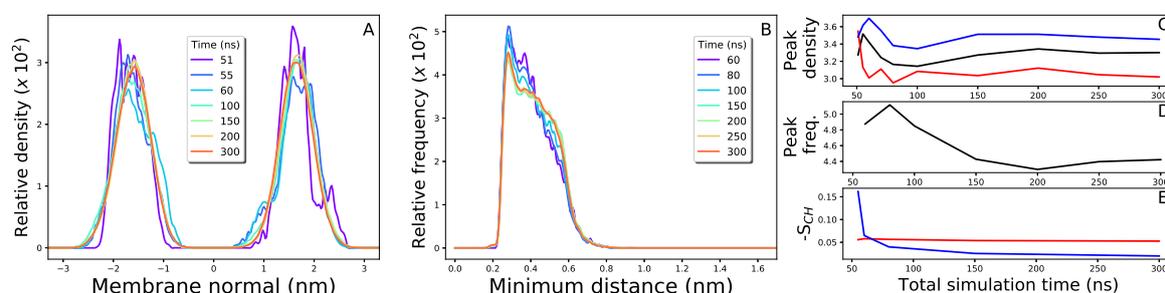


Figure S2: Convergence of calculated properties with total simulation time. Panel A shows the relative mass density of the UQ<sub>10</sub> head along the membrane normal as in Fig. 2C, calculated for various simulation times as colored in the legend. Panel B shows the distribution of minimum distances between the UQ<sub>10</sub> head oxygen O1 and DLPC+DLPE glycerol as in Fig. 3A for various simulation times. Panel C shows the peak of relative density ( $\times 10^2$ ) for phosphate group of DLPC+DLPE in black, Q-head in red and Q-tail in blue. Panel D shows the peak of minimum distance ( $\times 10^2$ ) between the UQ<sub>10</sub> head oxygen O1 and DLPC+DLPE glycerol. Panel E shows the C<sub>10</sub>-H bond vector order parameter for linoleoyl acyl chains (18:2,  $S_{CH}$  for C<sub>10</sub>) average in red as in Fig. 5 and standard deviation in blue. The first 50 ns were discarded to allow for equilibration in the trajectories analyzed. Properties converge to within 1-2% of the longest-time values in 150 ns of simulation or less. Similar convergence behaviour was observed for other molecules and properties not shown here.

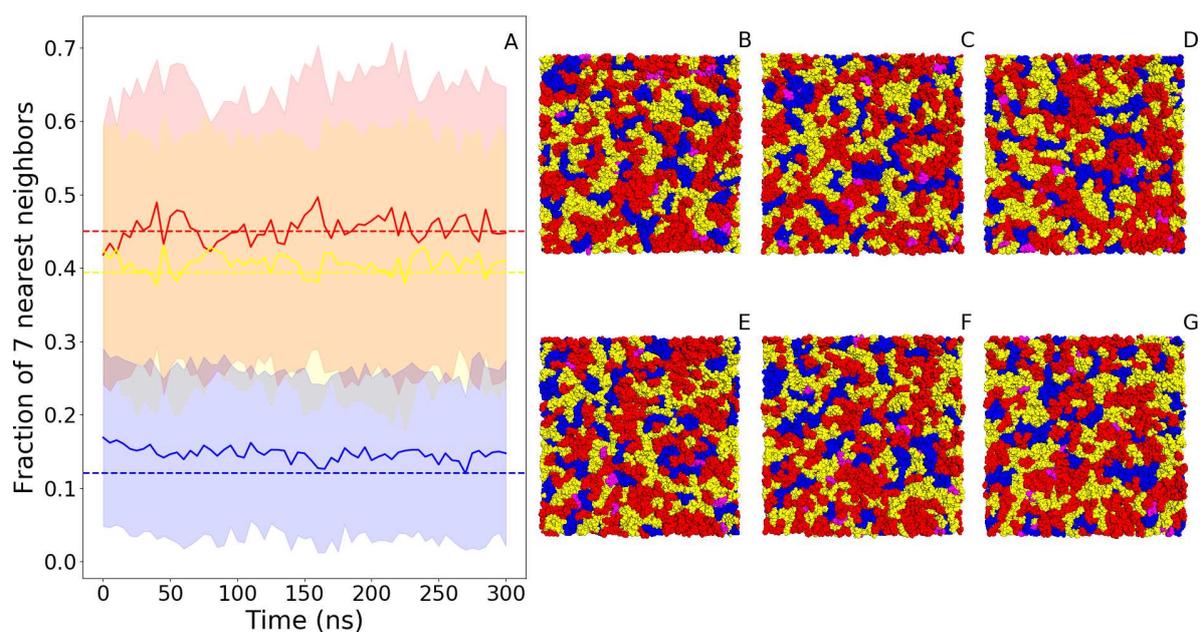


Figure S3: Multi-component membrane is well mixed. Panel A shows the average fraction of DLPC (yellow), DLPE (red) and LCL (blue) neighbors to DNPC (solid lines, with standard deviation in shadow) calculated as previously described.<sup>76</sup> Dashed lines show the expected fraction for an ideal mixture. Panels B-G show snapshots from 50 to 300 ns separated by 50 ns for the upper leaflet with lipids colored as in panel A plus magenta for UQ<sub>10</sub>. No lipid clusters or domains are observed.

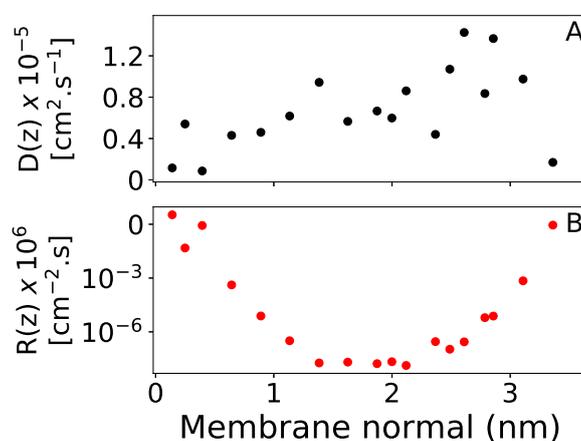


Figure S4: Coefficients of local transversal diffusion [panel A,  $D(z)$ ] and local resistance [panel B,  $R(z)$ ] for UQ<sub>2</sub> permeation along the membrane normal ( $z$ ). As previously described,<sup>60,61</sup> coefficients were calculated by  $D(z) = \frac{\text{var}(z)}{\tau}$  and  $R(z) = \exp[\frac{\Delta G(z)}{RT}]D(z)^{-1}$  with the variance ( $\text{var}$ ) and autocorrelation characteristic time ( $\tau$ ) of the normal coordinate  $z$  obtained from each US simulation window and symmetrized for the two leaflets. The diffusion profile is rather noisy because of the  $\tau$  values, calculated by integration of the autocorrelation function. Nevertheless,  $\tau$  values have small contributions to the resistance profile, which is determined mainly by the exponential free energy contribution,  $\Delta G(z)$ . A permeation coefficient may be calculated as  $P^{-1} = \int R(z)dz$ , integrated in the membrane range  $z=[-3.0,3.0]$  nm to give  $P=20$  cm s<sup>-1</sup>. To our knowledge, there is no experimental value for Q permeation to compare with.

Table S1: Atomic types and partial charges for PQ and MQ quinone heads. See figure S1 for atom names.

<b>PQ head</b>		
Atom name	Atom Type	Partial Charge
C5	CA	-0.0575
C6	CA	-0.0575
C4	CA	0.57
O4	O	-0.57
C3	CA	-0.115
HM5	HP	0.115
C2	CA	-0.115
C1	CA	0.57
O1	O	-0.57
CM3	CT3	-0.155
HM31	HA3	0.09
HM32	HA3	0.09
HM33	HA3	0.09
CM2	CT3	-0.155
HM21	HA3	0.09
HM22	HA3	0.09
HM23	HA3	0.09
<b>MQ head</b>		
Atom name	Atom Type	Partial Charge
C5	CA	0.09
C6	CA	0.09745
C4	CA	0.57
O4	O	-0.57
C3	CA	0.00
CM4	CA	-0.115
C2	CA	0.00
CM1	CA	-0.115
C1	CA	0.57
O1	O	-0.57
CM5	CQ31	-0.45745
HM51	HA3	0.09
HM52	HA3	0.09
HM53	HA3	0.09
CM3	CA	-0.115
HM3	HP	0.115
HM4	HP	0.115
CM2	CA	-0.115
HM2	HP	0.115
HM1	HP	0.115

### 3 Balanced internal hydration discriminates substrate binding to respiratory complex I

Murilo Hoiias Teixeira and Guilherme Menegon Arantes

Department of Biochemistry, Instituto de Química, Universidade de São Paulo, SP,  
Brazil

Reprinted with permission from Teixeira, M. H.; Arantes, G. M. Balanced internal hydration discriminates substrate binding to respiratory complex I. *Biochim. Biophys.*

*Acta, Bioenerg.*, v. 1860, p. 541-548, 2019. Copyright 2019 Elsevier B.V.



# Balanced internal hydration discriminates substrate binding to respiratory complex I

Murilo Hoias Teixeira, Guilherme Menegon Arantes\*

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



## ARTICLE INFO

**Keywords:**  
Ubiquinone  
Molecular recognition  
Molecular dynamics simulations  
Free energy profile

## ABSTRACT

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

## 1. Introduction

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

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

subunits and 980 kDa mass) involved in complex assembly, stability and specialized metabolic roles.

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

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

\* Corresponding author.

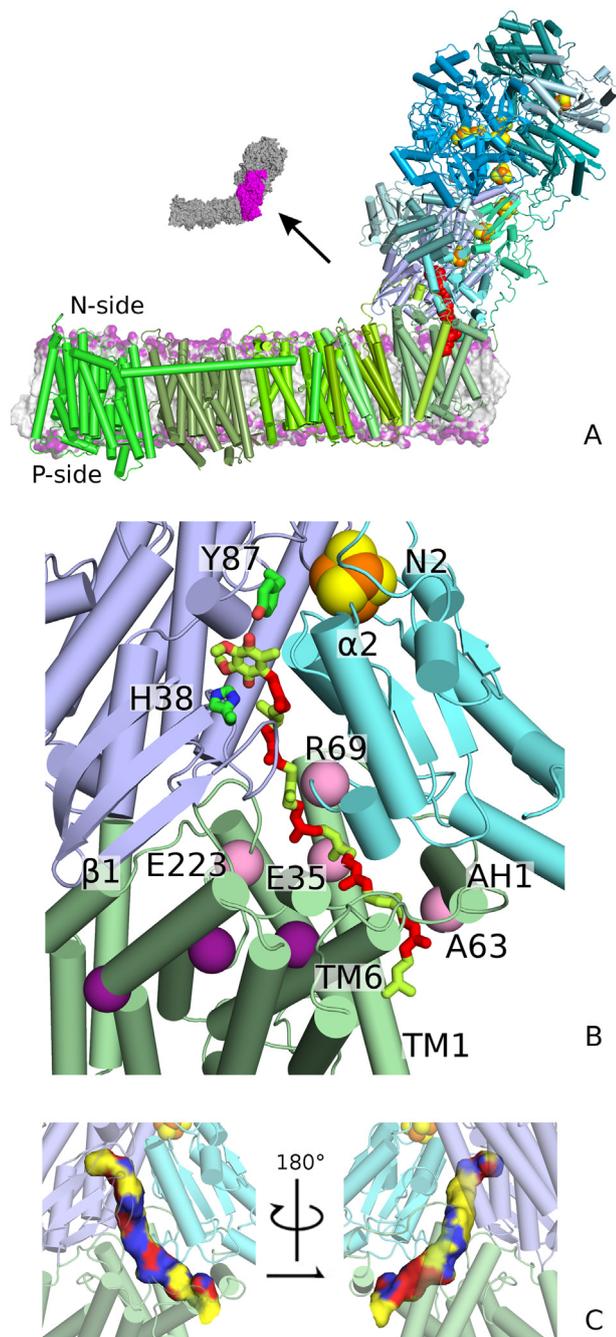
E-mail address: [garantes@iq.usp.br](mailto:garantes@iq.usp.br) (G. Menegon Arantes).

<https://doi.org/10.1016/j.bbabio.2019.05.004>

Received 6 November 2018; Received in revised form 16 May 2019; Accepted 28 May 2019

Available online 04 June 2019

0005-2728/ © 2019 Elsevier B.V. All rights reserved.



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

required for full enzymatic activity. X-ray data show the inhibitor piericidin A and the substrate analogue decylubiquinone, both containing short hydrophobic tails, bind in the Q-chamber at this position [8,9]. The chamber continues through an amphipathic region, roughly composed by a charged surface facing the membrane arm and by a large

hydrophobic patch on the opposite side (Fig. 1B and C). It ends exposed to the lipid membrane in an exit formed by Nqo8 helices TM1, TM6 and amphipathic helix AH1. The Q-chamber geometry and the free volume available to accommodate Q inside are similar among all determined active structures [8,9,11,10,13].

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

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

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

## 2. Materials & methods

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

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

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

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

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

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

Modeling of the Q binding process requires a balanced description of Q interactions with the hydrated protein interior and the lipid membrane. This was accomplished by using a calibrated force-field for Q, previously shown to give very good agreement between simulated and experimental water-to-membrane partition coefficients and free energies [25,71]. Interactions of protein, lipids and ions were described with the all-atom CHARMM36 force-field [31,32]. Water was represented by standard TIP3P [33]. FeS centers were described using the Chang & Kim [34] parameters with corrections proposed by McCullagh & Voth [35]. These corrections are important to keep the cuboidal structure in  $\text{Fe}_4\text{S}_4$  clusters.

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

## 2.2. Reaction coordinates and free-energy calculations

Two reaction coordinates were used to describe the Q binding process: the distance between the centers of mass (dCOM) of the six carbon atoms in the Q ring and ten  $\text{C}_\alpha$  of residues in subunit Nqo4 surrounding the Q-head in the reactive site (Fig. S2); and a pathway collective variable (Path CV) [40] using a distance metric [16,17] as implemented in PLUMED 2.3.1 [41] between the Q-head and 55  $\text{C}_\alpha$  of residues in subunits Nqo4, Nqo6 and Nqo8 (Fig. S2) in direct contact with Q inside the chamber. The Path CV was evaluated with respect to 30 milestone configurations representing progressive binding of Q along the chamber (see ESI Methods). These two kinds of reaction coordinates have already been used successfully to describe ligand binding along protein tunnels or cavities [16,17,18,19]. Results are presented using the shifted coordinate  $\text{sdCOM} = 4 \text{ nm} - \text{dCOM}$ , so that

entrance of Q from the membrane into the Q-chamber proceeds from low to high values and is seen from left to right in the figures shown. This representation is more intuitive than using dCOM or Path CV, which run from high to low values during entrance of Q. All the conclusions drawn here are equivalent when projections over either dCOM (and its sdCOM representation) or the Path CV are used to describe the results (Fig. S3 and ESI).

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

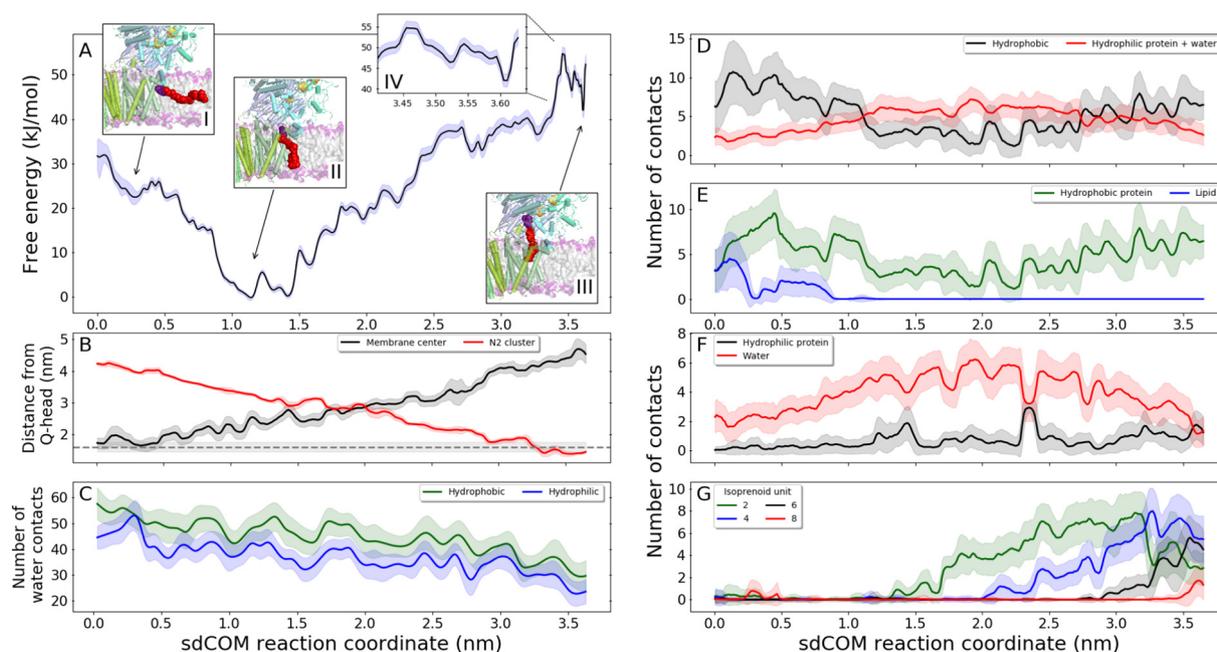
## 3. Results

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

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

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

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



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

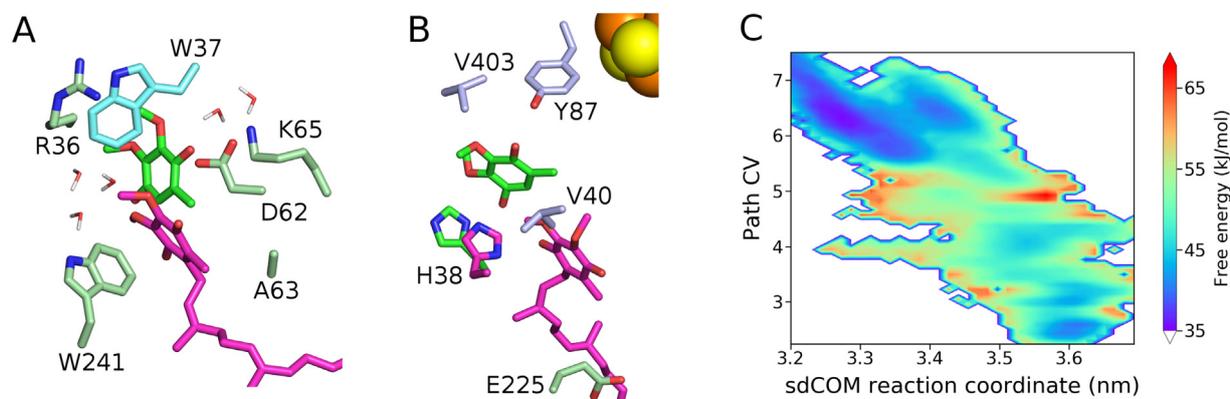
to the Q-chamber dehydrate due to entrance of Q and partial expulsion of water molecules from the chamber to the aqueous phase (Fig. 2C). Q-head interactions with hydrophobic groups decrease, hydrophilic contacts are formed and hydration increases from an average of 2 water contacts in sdCOM = 0.3 nm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

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

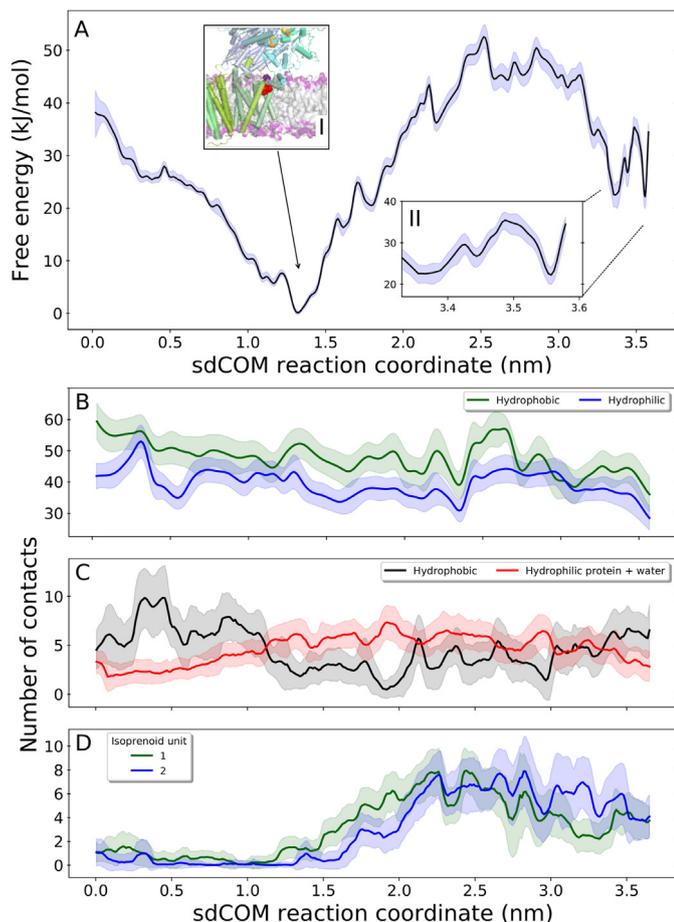
In the minimum found at  $3.2 < \text{sdCOM} < 3.4$  nm, a Q carbonyl oxygen can form a hydrogen bond with the protonated Nqo4 His38

side-chain (pre-RS position, Fig. 3B). Q is still far from Nqo4 Tyr87 and the other closest acidic residue, Nqo8 Glu225, is more than  $8 \text{ \AA}$  away. In order to form hydrogen bonds with both His38 and Tyr87, the barrier at sdCOM = 3.45 nm has to be transposed.

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



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



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

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

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

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

Although natural complex I substrates have 6 to 10 isoprenoid units, experiments are often carried out with more soluble short-tail analogues [45,46,47,1,27]. The effect of Q-tail length on binding was investigated by simulating  $Q_2$  as shown in Fig. 4. The free energy profile is similar to  $Q_{10}$  binding up to Q entrance in the BS ( $\text{sdCOM} < 1.5 \text{ nm}$ , Fig. 4A). The only difference is the relative smaller stability of the hydrated BS pose for  $Q_2$  ( $\text{sdCOM} = 1.1 \text{ nm}$ , similar to a binding site

identified for  $Q_1$  in another simulation study [24]). Hydration of the  $Q_2$ -tail and the chamber interior, and contacts performed by the  $Q_2$ -head (Fig. 4 B–D) are also equivalent to those performed by  $Q_{10}$  up to this point of the binding process.

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

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

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

## 4. Discussion

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

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

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

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

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

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

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

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

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

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

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

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

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

The experimental turnover rate of complex I catalysis is similar for reduction of Q with 10 to 4 isoprenoid units (average  $k_{cat} = 380 \pm 39 \text{ s}^{-1}$ ), but decreases significantly for Q with shorter tail ( $k_{cat} = 138 \pm 7 \text{ s}^{-1}$  for  $Q_2$  reduction) [50]. This rate drop-off may be interpreted by assuming the intrinsic barrier for redox conversion of oxidized Q to quinol ( $QH_2$ ) bound at the RS is independent of Q-tail length and by noting that free energies calculated here show similar barrier heights for  $Q_2$  and  $Q_{10}$  but higher stability for  $Q_2$  (favored by 20 kJ/mol when bound at or near the RS). Thus, the rate-limiting step

in the slower  $Q_2$  turnover should change in relation to Q with long tails and take place after formation of bound quinol, probably product release. Due to high stability of Q bound to the BS and possible fast Q entrance inside complex I (high  $k_{on}$ ), a second short-tail substrate molecule may easily enter the chamber after the first one, and block dissociation once the first is reduced. This would not occur for Q with a long tail which occupies the BS and is also in line with the auto-inhibition of complex I activity observed in high concentrations of Q with short tails [45,46,58].

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

## 5. Conclusions

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

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

## Transparency document

The [Transparency document](#) associated with this article can be found, in online version.

## Acknowledgments

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

## Supplementary data

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

## References

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

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

## Electronic Supplementary Information

# Balanced internal hydration discriminates substrate binding to respiratory complex I

Murilo Hoias Teixeira and Guilherme Menegon Arantes

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

## Supplementary Methods

The structure of complex I from *T. thermophilus*[8] did not contain a bound NADH, so this substrate was not included in the computational model. The NADH binding site is located more than 50Å from the Q-chamber. Given the total simulation time per sampling window, it is unlikely that any structural perturbation near the Q-chamber could have been caused by lack of NADH in the model. A disulfide bond was included between residues Nqo2 C144 and C172. All missing atoms in chain terminal regions were not built in the model nor considered for analysis. The only missing terminus that might be located close to the Q-chamber is Nqo6 residues 1-15 (sequence MALKDLFERDVQELE). This N-terminus has amphipathic character with many charged residues and the sequence is not conserved among species. Thus, it should not penetrate in the membrane nor perturb Q attachment to the nearby chamber exit formed by Nqo8 helices TM1, TM6 and AH1. It should be mentioned that other Q-chamber exits exposed to the lipid membrane have been proposed, including TM1 and TM7, or TM5 and TM6, both together with AH1, all from subunit Nqo8 (or the equivalent ND1 in bovine)[8, 10, 13]. These other proposals were not considered here as they show narrower exits and should require significant protein movement for passage of the Q-head.

Although Nqo4 loop  $\beta 1$ – $\beta 2$  and Nqo8 loop TM5–TM6 have been observed in different conformations[9, 11, 10], the position of these loops in the *T. thermophilus* structure should be appropriate for describing Q binding as it was co-crystallized with Q analogues[8].

All protein hydrogen atoms were built with the GROMACS utility[35] during model construction. The NPT ensemble was used in all simulations. Temperature was maintained with the Bussi thermostat[63] and a coupling constant of 0.1 ps with two separate coupling groups (water plus ions and everything else). Pressure was kept with the Berendsen barostat[64] during the initial equilibration and with the Parrinello-Rahman barostat for the umbrella sampling simulations with a coupling constant of 1 ps and a compressibility of  $1.10^{-5}$  bar<sup>-1</sup>. Semi-isotropic coupling in the direction normal to the bilayer was applied. A real space cutoff of 1.2 nm was used for PME electrostatics and van der Waals interactions. No dispersion corrections were applied. All covalent bonds were constrained using LINCS[65].

For both Q<sub>10</sub> and Q<sub>2</sub> simulations, US windows were separated by 0.1 nm in the dCOM range 0.43-3.43 nm and 3.60-4.10 nm that covers binding through the entire Q-chamber and the membrane exit. Residues used for the definition of both reaction coordinates are shown in Fig.S2. Additional windows were placed with reference dCOM and Path CV as shown in Table S1. Coordinate distributions were checked for sufficient overlap between adjacent 2D-windows, as required by the WHAM procedure. Two Supplementary Files are included: coordinates of the 30 milestone configurations used for calculation of Path CV; and the initial Q<sub>10</sub>-bound structure used

for all simulations. Results can be fully reproduced using these files and the information described here. In all analysis shown, a contact was considered when the distance between any pair of atoms in each group was smaller than 3.5Å. Table S2 shows the atoms included on each group. Isoprenoid units were composed of 5 carbons and the 8 bound hydrogens, except for the terminal isoprenoid with an extra hydrogen.

The choice of reaction coordinate is crucial for reliable calculation of free energy profiles. A simple distance metric between a reference point in the protein and another in the ligand would suffice to describe transit along the Q-chamber if it was homogeneously thin and tight, like a perfect tunnel. However, the Q-chamber BS and RS are relative bulky cavities where different poses and interactions may correspond to the same reference distance. In these sites, a simple distance metric becomes ill-defined and sampling with only one control coordinate such as dCOM would be flawed. In order to help resolving Q transit inside the full chamber, the Path CV coordinate based on milestone configurations along the binding pathway was used here[17]. Two-dimensional free energy surfaces are shown in Fig. S3A and B. They are rather diagonal, indicating that presentation of results using projections on either dCOM or Path CV coordinates would lead to equivalent conclusions. For instance, the projected profiles in Fig. S3C and D are qualitatively similar displaying the same number of local minima, maxima and relative energies, except for the end of the Path CV profile (Path CV>26) which is more compressed due to the smaller number of milestone configurations used in this region. Two-dimensional profiles are used here to resolve the profile minima in the RS region, as shown in Figs. 3C and Fig. S3E.

Given the size and complexity of the simulated system, it is unlike that all degrees of freedom were fully sampled in our simulations. However, the most important ones – hydration of the chamber interior and Q-tail – appear to have been well sampled. Fig. S8 shows the relaxation of internal water molecules bound near the Q-chamber. Internal water diffuses slower (first peak in panels S8A-C) than bulk water (second peak) as expected, but fully exchanges from the protein interior to the bulk in less than 10-15 ns. Thus, hydration of different groups in the Q-chamber and the Q-tail equilibrated in less 50 ns in all US windows. Figs. S1D/E show that stabilization within statistical uncertainty of the free energy calculated for Q bound in the RS state ( $\Delta G_{RS}$ ) is observed with an equilibration time of 40-50 ns and acquisition time of 80-90 ns per US window. This point of the profile has the slowest convergence over all US simulations presented here. Error bars (blue shadows) in free energy profiles indicate the sampling statistical uncertainty only. Imperfections in model composition and structure or in the interaction force-field may introduce systematic errors, which unfortunately are difficult to estimate. Nevertheless, the favorable comparison between three different profiles obtained from different starting structures for Q<sub>10</sub> (Figs. 2A and S1C) and Q<sub>2</sub> (Fig. 4A) suggest that our results have at least semi-quantitative precision.

## Supplementary Results & Discussion

### *Structure of Nqo6 loop $\alpha 2-\beta 2$*

As described in the Methods section of the main text, the protein model used here was based on the structure of the respiratory complex I from *T. thermophilus* (PDB ID: 4HEA)[8], with Nqo6 loop  $\alpha 2-\beta 2$  rebuilt using the loop in the ovine enzyme as a template (Fig. S1A). The resulting loop structure after relaxation is very close to other mammalian enzymes, including the recent mouse structure (PDB ID: 6G2J)[13] determined with a 3.3Å resolution (root mean-squared deviation, RMSD, of 0.14 nm, Table S3). The original (4HEA) Nqo6 loop  $\alpha 2-\beta 2$  structure is quite different (RMSD of 0.56 nm). Hence the RMSD of 0.18 nm found for Q-chamber in comparison to the 4HEA structure in Table S3 is caused almost completely by deviations on this Nqo6 loop.

In preliminary simulations shown in Fig. S1, the missing Nqo6 residues 65-69 were built with MODELLER *de novo*, *i.e.* without a structural template. A rather different loop structure was obtained as shown in Fig. S1B, with divergent positions of highly conserved Arg69 and Arg62. This difference had a profound impact on the calculated free energy (Fig. S1C). The profile in  $\text{sdCOM} < 1.5$  nm is similar to Fig. 2A in the main text, as Q is still distant from Nqo6 loop  $\alpha 2$ - $\beta 2$ . However, the bumps found in between  $1.5 < \text{sdCOM} < 2.5$  nm (Fig. S1C) are due to artificial interactions of Arg62 with the Q-head. These are not observed in the main text simulations as Arg62 is distant from the Q-chamber when the Nqo6 loop structure is based on the ovine enzyme (Fig. S1B). Differences in the profiles are even more pronounced when Q is bound at the RS. In fact, no (meta)stable free energy minimum is found in the RS, as Arg69 artificially binds the Q-head instead of the protonated Nqo4 His38 (Fig. S1B). Thus, caution should be used when building loop segments *de novo*, particularly if the built residues are in contact or near the protein region to be studied in detail.

#### *Modelling limitations*

Based on the complex I structure co-crystallized with a substrate analogue[8], our initial model should be appropriate to study Q binding in the internal chamber formed by subunits Nqo4, Nqo6 and Nqo8. This Q-chamber is rather stable and shows similar conformations among different species (Table S3). But the flexible Nqo6 loop  $\alpha 2$ - $\beta 2$ , originally with a missing segment[8], had its structure rebuilt here using the ovine loop[11] as a template. Such structural refining turned out to be important to describe Q stability in the RS (Fig. S1).

Given an appropriate structural model and a reasonable description of Q interactions[25], accurate simulation of Q binding also requires sampling the configurational space. The narrow and conformationally restrict Q-chamber suggests that configurational freedom of groups inside will be small and that description of ligand binding with a simple distance coordinate may be accurate. However, the BS and the RS are bulky cavities and a more elaborate control coordinate such as the Path CV[17] had to be employed (see SI Methods). Formation of the initial ligand-protein encounter complex with Q still in the membrane or crawling over the protein surface in order to find the exposed chamber exit were not considered here. Although these will contribute to the total binding free energy, simulating processes with Q outside the chamber requires sampling a much larger configurational space or using elaborate spacial restrictions[66].

#### *Protein contacts and flexibility along Q-binding*

Q transit may be described starting from the chamber exit when the Q-head is in contact with apolar residues in helices Nqo8 TM1, TM6 and AH1 (Fig. S5B). Once inside the chamber and towards the BS, the Q-head is close to the Nqo8 loop TM5-TM6 (in the membrane arm side) and TM1 (in the opposite side), and with Nqo6  $\beta 1$ -strand and loop  $\alpha 1$ - $\beta 1$ . After the BS, when the chamber has a turn upwards, the Q-head interacts mostly with Nqo6 helix  $\alpha 2$ , loop  $\alpha 2$ - $\beta 1$ ,  $\beta 1$ -strand and turn  $\alpha 4$ - $\alpha 5$ . Once in the RS, the Q-head forms contacts with Nqo6  $\alpha 2$  (and the nearby loop  $\alpha 2$ - $\beta 1$  that binds the N2 cluster) and Nqo4 loop  $\beta 1$ - $\beta 2$ , end of  $\alpha 4$  and C-terminus helices and loop  $\alpha 2$ - $\alpha 3$  (Fig. 1B).

Besides the BS, pre-RS and RS positions already described in the main text, the Q-head may directly contact charged residues at  $\text{sdCOM}=1.8$  nm and hydrogen bond with Nqo8 Arg45 and Nqo6 Arg69 (Fig. S5A). The Q-tail forms several hydrophobic contacts with apolar residues particularly when isoprenoid units approach the chamber exit, the Nqo6 helix  $\alpha 2$  and the Nqo4 loop  $\beta 1$ - $\beta 2$  (Fig. 1).

The Q-chamber backbone and all described secondary structures remain stable during the simulations. For Q<sub>10</sub> binding, RMSD of C <sub>$\alpha$</sub>  position for Q-chamber residues vary between 0.1-0.3 nm along the binding

coordinate. Most of this can be traced to thermal fluctuations of Nqo4 loop  $\beta 1-\beta 2$ , Nqo6 loop  $\alpha 2-\beta 2$  and Nqo8 loop TM5–TM6 (Table S3). Fig. S9 shows representative RMSD trajectories calculated for entire chains Nqo4, Nqo6 and Nqo8, and for only Nqo8 residues E163, E213 and R216 possibly involved in activation of the E-channel, for US windows with Q bound in the RS and in the BS. Structural fluctuations during the trajectories are small ( $<0.1$  nm), as expected for stable structures. The initial RMSD to the X-ray structure express structural relaxation in water during model equilibration and, for the individual residues, also includes the protein response to Q transit inside the chamber (see below section *Possible role of Q binding in activation of the E-channel*).

#### *Point mutations and Q-chamber stability*

Point mutations on residues in the Q-chamber perturb assembly of the complex and its enzymatic activity (Table S5)[8]. Several of these mutations are known to be pathological in humans[6, 7]. Interestingly, mutations on the less flexible residues Nqo8 E35, R36, R45, D62, performing stable ionic contacts (Fig. S7) in the opposite side of the membrane arm, near Nqo8 helices TM1 and AH1, are more deleterious and reduce the content of assembled complex I, unless a conservative mutation is observed. These four residues also show the highest conservations among the Q-chamber (Table S4). There is more plasticity on the network of ionic contacts on the side of the E-channel which fluctuate often during our simulations (Fig. S7). Mutations (Nqo8 E163, E213, R216, E223, E225, E248 and R294) are less deleterious and activity is maintained even when a charge is reversed (such as E163K, Table S5). Thus, ionic contacts on the side of Nqo8 helices TM1 and AH1 have an essential structural role and keep the Q-chamber architecture. Contacts on the E-channel side near Nqo8 helices TM5 and TM6, and Nqo6  $\beta 1$  strand may fluctuate without disrupting the enzyme structure. These “controlled” fluctuations may be explored by evolution in order to modulate enzyme activity. Indeed, it has been shown recently that localized unfolding of Nqo8 TM5 and TM6 and Nqo6  $\beta 1$  is involved in the active-to-deactive transition observed in mammalian complex I[67].

#### *Possible role of Q binding in activation of the E-channel*

The network of ionic contacts near the chamber is perturbed by passage of Q, specially contacts located on the membrane arm side (Fig. S7). The Q-head forms hydrogen bonds with Nqo8 Arg36 in the BS (sdCOM=1.4 nm, Fig. 2F), leading to rupture of the Arg36–Glu66(Nqo6) salt bridge, which otherwise remains stable along Q binding. Contact Arg36–Asp62, although formed transiently before the BS, is only stable in sdCOM $>1.4$  nm, after Q passage towards the RS (Fig. S7C). Nqo6 Arg69 forms transient salt bridges with both Nqo8 Glu223 and Glu225. After passage of Q and hydrogen bonding with the Q-head (sdCOM=2.3 nm, Fig. 2F), only Arg69 bridge with Glu225 remains stable and the contact with Glu223 may no longer be formed (sdCOM $>2.3$  nm, Fig. S7B). When not coordinated with Arg69, the flexible Glu223 side-chain approaches Nqo8 Arg216 (in sdCOM=0.3, 1.2, 2.2 and 3.1 nm, Fig. S7D) disturbing its contact with Glu213, which in turn may approach Glu163 (Fig. S7F), the following residue in the E-channel (Fig. 1B). Note that Glu223 never approaches Glu213 itself (Fig. S7E). Additionally, when Q is in the RS (sdCOM $> 3.5$  nm), the otherwise stable bridge Arg216–Glu248 in Nqo8 is slightly perturbed (Fig. S7D), bridge Arg216–Glu213 has the closest contact over the binding process and Glu213 is the farthest from Glu163 (Fig. S7F).

This highly hydrated network of ionic contacts has been proposed to be involved in a putative Nqo8 proton pumping channel and in the coupling between the Q redox process and proton pumping in the membrane arm, via the E-channel[8, 21, 23, 24]. The above observations allow us to speculate a possible role for Q transit in

the coupling mechanism. When Q is unbound from complex I, the proton pumps should be deactivated as no electron transport will take part. This is in line with the large Glu213–Glu163 distances seen for  $\text{sdCOM} < 2.0$  nm. As Q loads the chamber, the ionic network is perturbed as described above, Glu213 approaches Glu163, and the E-channel may propagate a “ready-to-go” signal to the proton channels. As soon as Q is reduced and binds the  $\text{H}^+$  from Nqo4 His38, Asp139 (Asp160 in bovine numbering) may twist and contact Arg294 which in turn complete activation of the E-channel to trigger proton transfer.

Besides the rebuilt Nqo6 segment D55-P72 (see Materials & Methods in the main text and discussion above in section Structure of Nqo6 loop  $\alpha 2$ – $\beta 2$ ), Nqo8 Arg216 and Glu163 also occupy positions in our initial protein model and during the US simulations that are fairly different from the *T. thermophilus* X-ray structure[8]. Arg216 is located in a flexible loop, near rebuilt Arg69, partially exposed to the Q-chamber and distant 18 Å from the Glu213 side chain in the X-ray structure. During initial model equilibration and over all simulated US windows (Fig. S9C-D), Arg216 moved away from the Q-chamber towards Glu213 (6 Å side chain distance in our equilibrated model), in line with chamber occupation by a Q molecule with long tail, electrostatic repulsion from Arg69, intrinsic loop flexibility and the smaller side chain distance (12 Å) observed in the ovine structure[11]. The increased Glu163–Glu213 side chain distance observed in our initial model as well as along Q transit (specially when  $\text{sdCOM} < 2.0$  nm, Fig. S7F) in comparison to the X-ray structure (6 Å) is brought by a local unwinding of the TM helix holding Glu163 (*e.g.*, Fig. S9C). However, this unwinding is reversible and depends on the Q position along the binding process ( $2.0 < \text{sdCOM} < 3.5$  nm), leading to smaller Glu163–Glu213 distances (Fig. S7F) and restoration of the Glu163 X-ray structure (Fig. S9D). Although this fluctuation of Glu163 position may be caused by Q transit inside the chamber and be related to the coupling mechanism, as proposed in the previous paragraph, it may also be an artifact of the anionic protonation state chosen to model Glu163. Anionic Glu213 suffers only side chain rotations and has a backbone conformation similar to the *T. thermophilus* X-ray structure.

In summary, the network of charged residues close to the BS may function as an ionic interface separating the hydrophobic exterior (apolar residues and lipid) from the more hydrated Q-chamber interior, much like the zwitterionic head in a phospholipid is at the membrane interface (Fig. 2B). Accurate molecular simulations of the correlated motions found on this network of residues in complex I may require more realistic model compositions such as representing the several possible protonation states of residues in the Q-chamber and E-channel and the electrochemical gradient across the membrane.

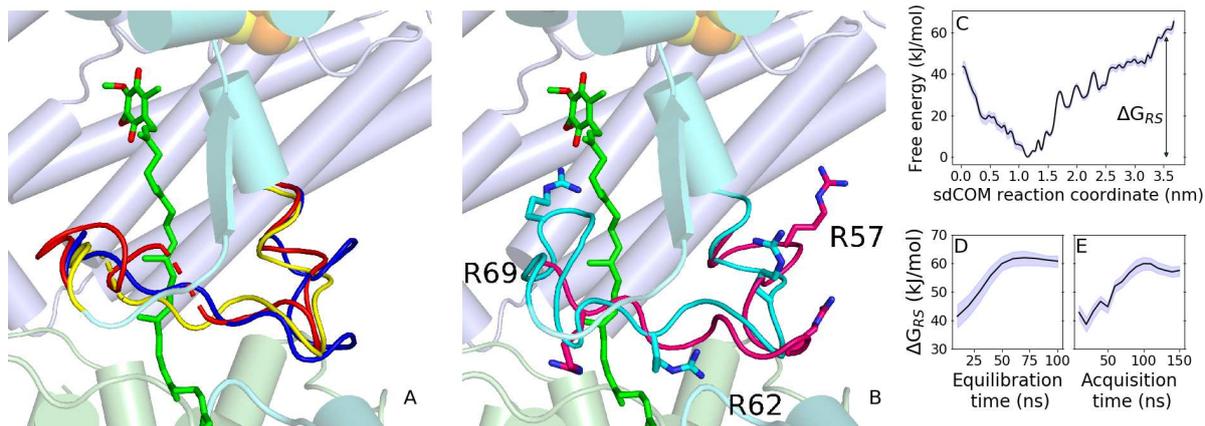


Figure S1: Structure of the Nqo6 loop  $\alpha 2-\beta 2$  and its impact on the free energy profile for Q binding. (A) Close view in cartoon of subunits Nqo4 (light blue), Nqo6 (cyan) and Nqo8 (pale green) with Q<sub>10</sub> modeled in green sticks and loop structures from ovine in blue (PDB ID 5LC5), *T. thermophilus* in yellow (PDB ID 4HEA) and *Y.lipolytica* in red (PDB ID 4WZ7). Note the missing segment 65-69 from the *T. thermophilus* in dashed yellow backbone. (B) Same view as (A), but with the loop of the relaxed initial protein model used in the main text simulations (Fig. 2) in magenta (compare with the blue loop in panel A), and the loop from *T. thermophilus* with the missing residues 65-69 built with MODELLER *de novo* in cyan. Arg69, Arg62 and Arg57 are also shown. (C) Free energy profile in black with statistical uncertainty in blue shadow obtained with the Nqo6 segment 65-69 built *de novo* as shown in panel B. Free energy convergence of Q bound in the RS ( $\Delta G_{RS}$ ) with increasing (D) equilibration and (E) acquisition simulation time per US window.

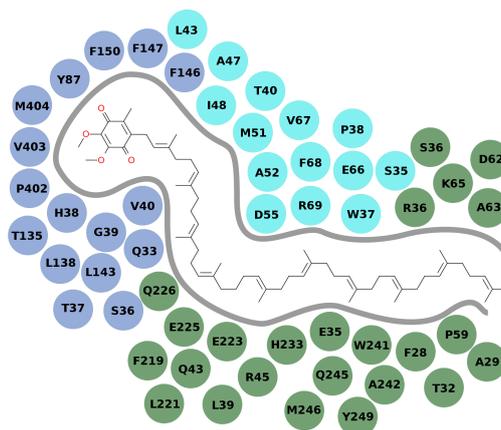


Figure S2: Residues along the Q-chamber which C $\alpha$  was used for definition of dCOM and Path CV reaction coordinates. Bound Q<sub>10</sub> is also shown. For dCOM, Nqo4 (light blue) residues Q33, H38, G39, V40, T135, L143, T144, P402, V403 and M404 were used. For Path CV, Nqo4 residues Q33, S36, T37, H38, G39, V40, Y87, T135, L138, L143, F146, F147, F150, P402, V403, M404, Nqo6 (cyan) residues S35, W37, P38, T40, L43, A47, I48, M51, A52, D55, E66, V67, F68, R69, A70, and Nqo8 (pale green) residues F28, A29, T32, E35, R36, L39, Q43, R45, P59, D62, A63, K65, S66, F219, L221, E223, E225, Q226, H233, W241, A242, Q245, M246, Y249 were used.

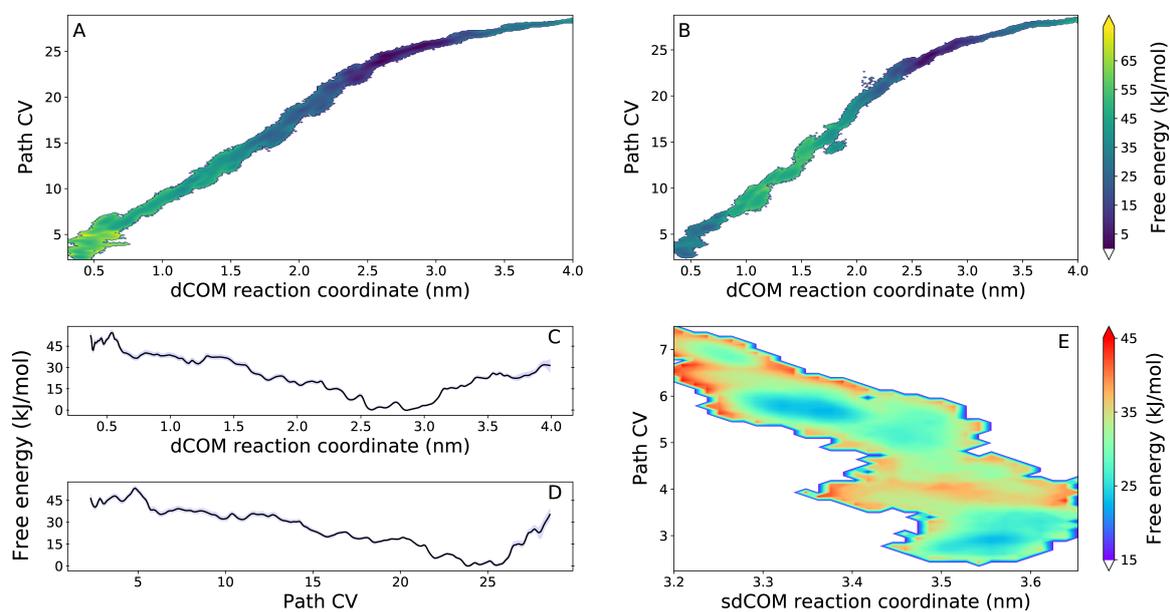


Figure S3: Free energy profiles for Q binding along the chamber in complex I. Complete two-dimensional free energy profiles calculated for (A) Q<sub>10</sub> and (B) Q<sub>2</sub> binding. Projection of the Q<sub>10</sub> 2D minimum free energy on (C) dCOM and (D) Path CV coordinates. Note the dCOM coordinate (instead of the shifted sdCOM) is used in the x-axis of panels A-D. Here, binding from the membrane into the chamber proceeds from high to low values for dCOM and Path CV reaction coordinates. (E) Two-dimensional free energy profile for Q<sub>2</sub> binding in the RS region.

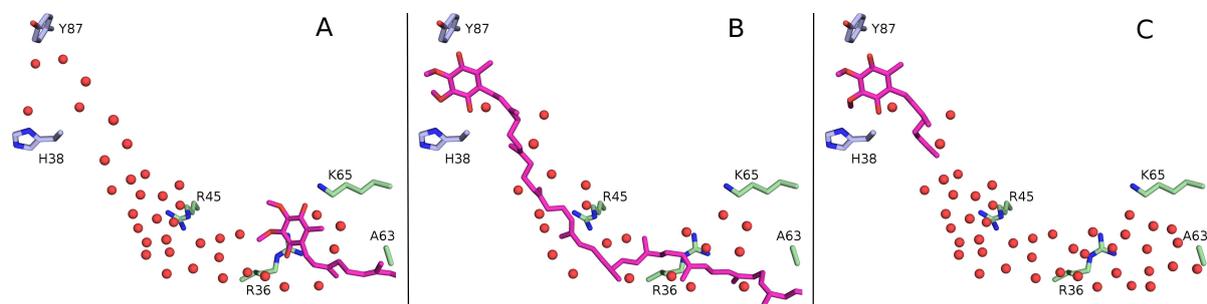


Figure S4: Chamber internal hydration with water oxygens shown in red spheres. Q in magenta occupies the (A) BS, (B) Q<sub>10</sub> in RS, and (C) Q<sub>2</sub> in RS positions.

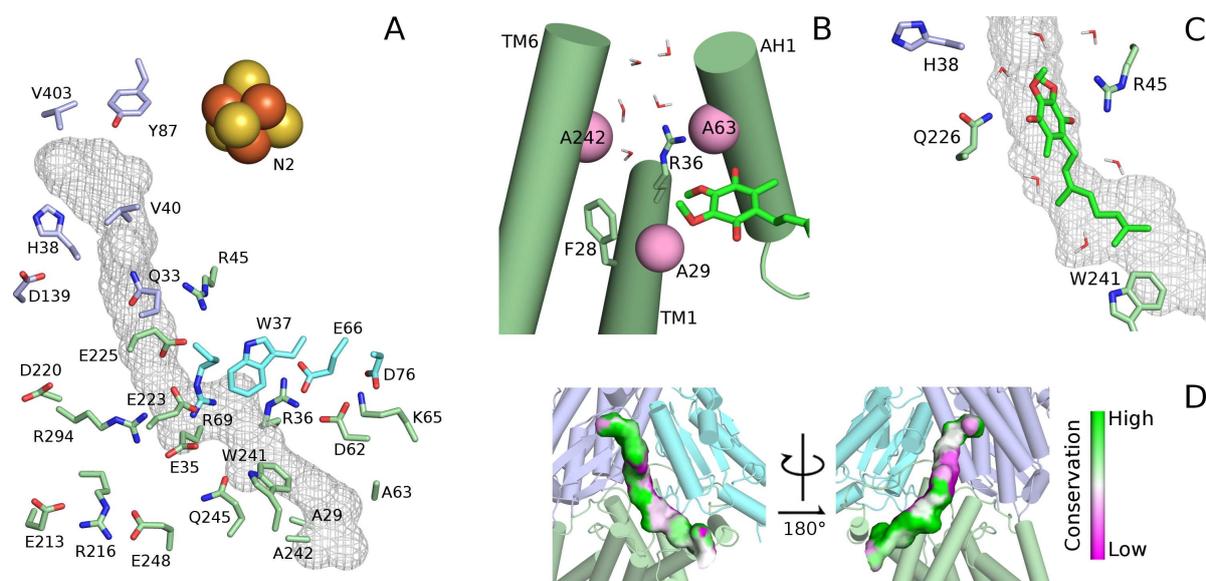


Figure S5: Protein environment of Q in the binding chamber. Key residues are shown for subunits Nqo4 (in light blue), Nqo6 (cyan) and Nqo8 (pale green). (A) Chamber overview with residue position and molecular surface mesh of Q bound in the RS. (B) CE position, with Q outside the chamber exit. Helices TM1, TM6 and AH1 are shown in cartoon with Ala residues as pink spheres. Water molecules are shown to indicate chamber hydration. This is the only panel displayed in a different viewing perspective. (C) Q<sub>2</sub> in green on the high free energy region for its binding (sdCOM=2.5 nm). Surface mesh as in panel A. (D) Molecular surface of Q bound in the RS colored by conservation of the nearest protein residue in *T. thermophilus* structure. The conservation score was calculated as described in Table S4.

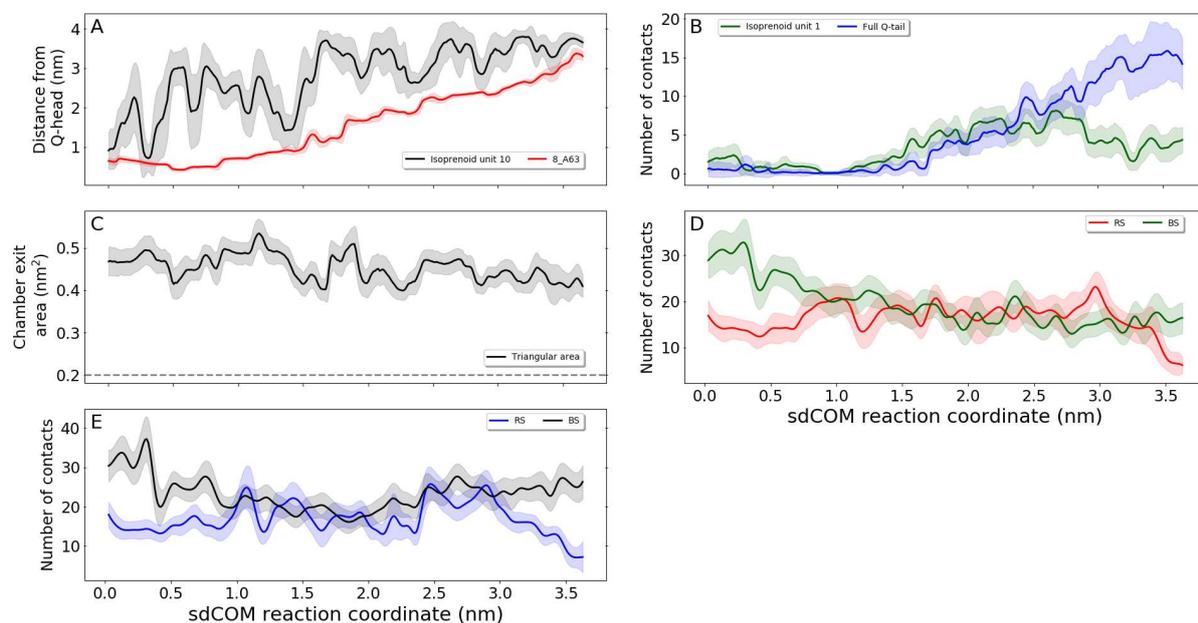


Figure S6: Additional simulation analysis. (A) Distance of the Q-head COM to the C<sub>α</sub> from Nqo8 Ala63 in the chamber exit in red and to the COM of isoprenoid unit 10 in black. (B) Hydration of the full Q<sub>10</sub>-tail in blue and of isoprenoid unit 1 in green. (C) Area of the triangle determined by the C<sub>α</sub> position of Nqo8 Ala29 (helix TM1), Ala63 (helix AH1) and Ala242 (helix TM6) in the Q-chamber exit to the lipid membrane (Fig. S5B). For the original crystal structure (PDB 4HEA)[8], this area is 0.37 nm<sup>2</sup>. The dashed gray line shows the estimated transversal area of the Q-head (0.2 nm<sup>2</sup>=0.7 nm × 0.3 nm, respectively the distance between carbonyl oxygens and the van der Waals thickness of the Q ring). Total hydration of BS and RS residues in the Q-chamber (D) for Q<sub>10</sub> and (E) for Q<sub>2</sub> binding.

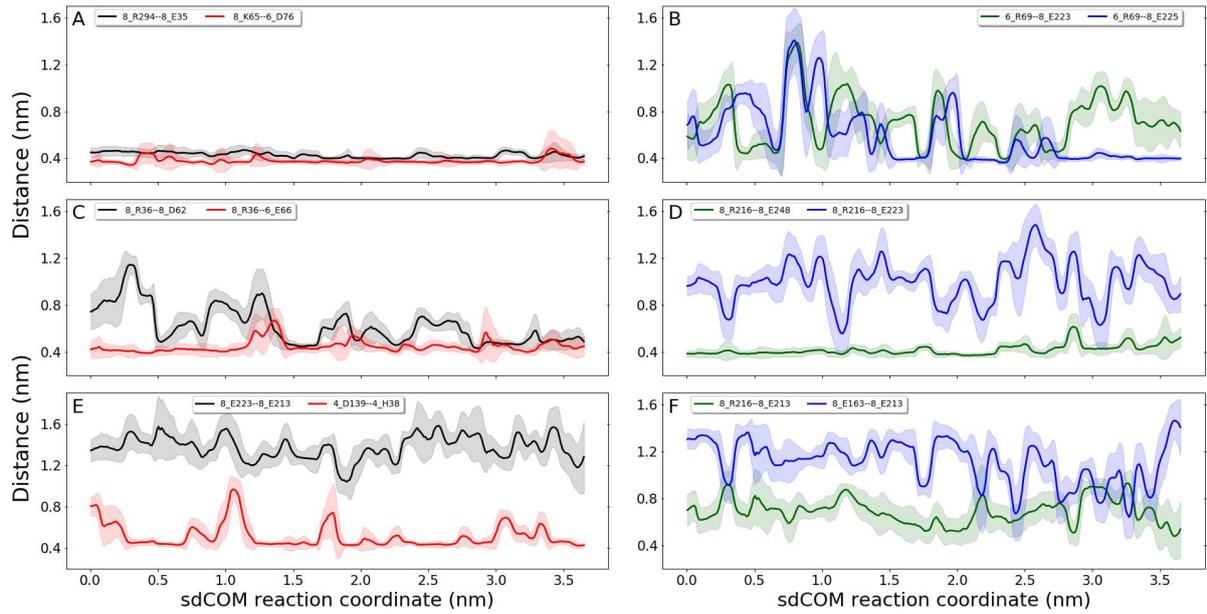


Figure S7: Network of ionic contacts near the Q-chamber. Pair distances along Q binding between side chains of (A) Nqo8\_R294–8\_E35 (black) and 8\_K65–6\_D76 (red), (B) 6\_R69–8\_E223 (green) and 6\_R69–8\_E225 (blue), (C) 8\_R36–8\_D62 (black) and 8\_R36–6\_E66 (red), (D) 8\_R216–8\_E248 (green) and 8\_R216–8\_E223 (blue), (E) 8\_E223–8\_E213 (black) and 4\_D139–4\_H38 (red), and (F) 8\_R216–8\_E213 (green) and 8\_E163–8\_E213 (blue). Distances were computed between  $C_{\zeta}$  for Arg,  $N_{\zeta}$  for Lys,  $C_{\delta}$  for Glu,  $C_{\gamma}$  for Asp and COM of imidazole ring for His.

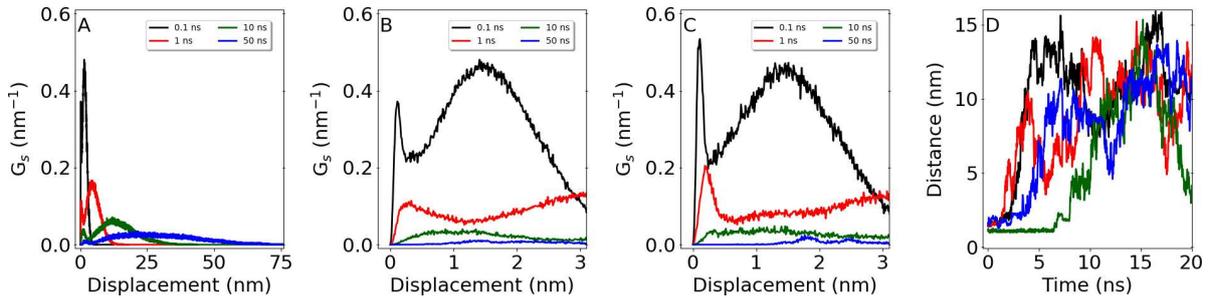


Figure S8: Relaxation of internal water molecules. (A, B and C) Self-part of the van Hove correlation function,  $G_s(r, \Delta t)$  [69, 70], for a group of  $\sim 30$  water molecules initially bound within 0.5 nm of Q and the protein Q-chamber, calculated for  $\Delta t = 0.1, 1, 10, 50$  ns as colored in the legend. This function gives the probability for a water molecule to show position displacements ( $r$ ) in  $\Delta t$  time intervals. Panels A and B show the same data obtained from a US window with  $dCOM = 0.43$  nm (RS) and panel C from  $dCOM = 2.93$  nm (BS), discarding the initial 50 ns for equilibration. (D) Representative trajectories of water distance to the Q-head for 4 water molecules initially bound inside the Q-chamber.

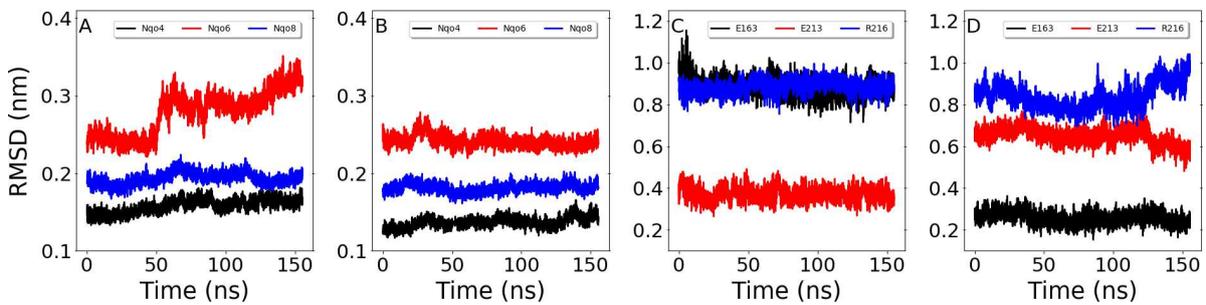


Figure S9: Representative trajectories of root mean-squared deviation of  $C_{\alpha}$  position for chains Nqo4, Nqo6 and Nqo8 (panels A and B) and of heavy-atom position for Nqo8 residues E163, E213 and R216 (panels C and D) as colored in the legend, in comparison to the *T. thermophilus* X-ray structure. Panels A and C obtained from US window with  $dCOM = 0.43$  nm (RS) and panels B and D for  $dCOM = 2.93$  nm (BS), discarding the initial 50 ns for equilibration.

Table S1: Reference dCOM (in nm) and Path CV of additional windows used in US simulations<sup>a</sup>.

Q <sub>10</sub>		Q <sub>2</sub>	
dCOM	Path CV	dCOM	Path CV
0.42	3.31	0.48	
0.43	2.80	0.48	3.90
0.43	3.10	0.58	5.60
0.43	3.20	0.73	6.40
0.43	3.50	1.11	10.0
0.46	3.72	1.85	16.7
0.48		2.33	21.8
0.48	3.90	2.77	
0.48	4.30	3.52	27.6
0.48	4.60		
0.48	4.90		
0.54	5.10		
0.54	5.20		
1.43	12.0		
1.93	16.6		
2.48			
2.77			
3.20	26.5		

<sup>a</sup> In addition to windows in dCOM range 0.43-3.43 nm and 3.60-4.10 nm separated by 0.1nm. No restrain was applied to Path CV when a reference value for this coordinate is not listed.

Table S2: List of atoms included on each group used for simulation and analysis.

Group	Atoms
Water	Oxygen (OW)
Q-head	Heavy atoms in Q except for the isoprenoid tail
Q-head oxygens	Four oxygen atoms in Q
Lipid tail	All atoms in both palmitoyl and oleoyl POPC hydrocarbon chains (except for carbonyl group)
Q-chamber C <sub>α</sub>	C <sub>α</sub> of residues Nqo4 33, 37, 38, 39, 40, 62, 87, 88, 131, 132, 135, 138, 143, 146, 147, 150, 398, 401, 402, 403, 404, residues Nqo6 37, 40, 41, 42, 43, 44, 45, 47, 48, 51, 52, 67, 68, 69, 70, 81 and residues Nqo8 28,29, 32, 35, 36, 39, 45, 59, 60, 62, 63, 65, 216, 219, 225, 226, 241, 242, 245, 246, 249, 294.
Nqo4 loop β1-β2	Nqo4 residues 32-41
Nqo6 loop α2-β2	Nqo6 residues 55-70
Nqo8 loop TM5-TM6	Nqo8 residues 220-238
BS residues	Side-chains of residues Nqo8 28, 29, 32, 35, 36, 59, 60, 62, 63, 65, 226, 241, 242, 245, 246, 249, 294
RS residues	Side-chains of residues Nqo4 38, 39, 87, 88, 131, 132, 135, 150, 398, 401-404 and Nqo6 44, 45, 48
Apolar (hydrophobic) protein	Side chains of G, P, A, V, I, L, M, F and W (except for N <sub>ε1</sub> and H <sub>ε1</sub> )
Charged and polar (hydrophilic) protein	Backbone N and hetero atoms in side chains of R, K, D, E, H, S, T, N, Q, C, Y

Table S3: Root mean-squared deviation (in nm) of the  $C_{\alpha}$  position for different structural models and groups of residues in the Q-chamber. The initial protein model (see Methods) was used as a reference structure.

Group	Structural model <sup>a</sup>					
	Q <sub>10</sub> simulation	4HEA	4WZ7	5LC5	5LNK	6G2J
Q-chamber	0.10-0.20	0.18	0.24	0.17	0.21	0.16
Nqo4 loop $\beta 1$ - $\beta 2$	0.12-0.26	0.02	0.25	0.21	0.38	0.20
Nqo6 loop $\alpha 2$ - $\beta 2$	0.13-0.21	0.56	0.47	0.18	0.15	0.14
Nqo8 loop TM5-TM6	0.14-0.29	0.03	0.55	0.24	0.47	0.24

<sup>a</sup> The average plus one-standard deviation RMSD range observed over the full reaction coordinate of the Q<sub>10</sub> US simulation is shown. The given structural codes correspond to PDB IDs 4HEA for *T. thermophilus*[8], 4WZ7 for *Y. lipolytica*[9], 5LC5 for bovine[10], 5LNK for ovine[11] and 6G2J for mouse[13].

Table S4: Conservation among species of Q-chamber residues in subunits Nqo4, Nqo6, Nqo8. Score was calculated after a multiple sequence alignment of 300 species with ConSurf 2016[68] (using a maximal identity between sequences of 90% and a minimal identity for homologs of 30% as elsewhere[8], with otherwise default settings). More negative scores indicate higher conservation.

Chain	<i>T. thermophilus</i>	<i>E. coli</i>	<i>Y. lipolytica</i>	Bovine	Ovine	Mouse	Conservation score
Nqo4	Q 33	N 223	Q 90	Q 54	Q 54	Q 54	-0.868
Nqo4	H 38	H 228	H 95	H 59	H 59	H 59	-1.178
Nqo4	V 40	A 230	V 97	V 61	V 61	V 61	-1.077
Nqo4	Y 87	Y 277	Y 144	Y 108	Y 108	Y 108	-1.208
Nqo4	D 139	D 329	D 196	D 160	D 160	D 160	-0.859
Nqo4	V 403	V 594	V 460	V 424	V 424	V 424	-1.080
Nqo6	W 37	W 55	W 77	W 46	W 46	W 56	-0.780
Nqo6	R 62	R 80	R 102	R 71	R 71	R 81	-0.978
Nqo6	E 66 <sup>a</sup>	E 84	-	-	-	-	+0.182
Nqo6	R 69	R 87	R 108	R 77	R 77	R 87	-0.522
Nqo6	D 76	D 94	D 115	D 84	D 84	D 94	-1.146
Nqo8	F 28	G 29	V 19	V 17	V 17	M 17	-0.786
Nqo8	A 29	A 30	A 20	A 18	A 18	A 18	-0.862
Nqo8	E 35	E 36	E 26	E 24	E 24	E 24	-1.223
Nqo8	R 36	R 37	R 27	R 25	R 25	R 25	-1.228
Nqo8	R 45	R 46	R 36	R 34	R 34	R 34	-1.281
Nqo8	D 62	D 63	D 53	D 51	D 51	D 51	-1.261
Nqo8	A 63	M 64	A 54	A 52	A 52	A 52	-0.249
Nqo8	K 65	K 66	K 56	K 54	K 54	K 54	-1.225
Nqo8	E 213	V 206	E 196	E 192	E 192	E 192	-1.175
Nqo8	R 216	R 209	R 199	R 195	R 195	R 195	-1.197
Nqo8	D 220	D 213	D 203	D 199	D 199	D 199	-1.250
Nqo8	E 223	E 216	E 206	E 202	E 202	E 202	-1.184
Nqo8	E 225	E 218	E 208	E 204	E 204	E 204	-1.159
Nqo8	Q 226	Q 219	S 209	S 205	S 205	S 205	-1.001
Nqo8	W 241	F 234	F 224	F 220	F 220	F 220	-0.866
Nqo8	A 242	G 235	V 225	A 221	A 221	A 221	-0.655
Nqo8	Q 245	F 238	F 228	F 224	F 224	F 224	-0.798
Nqo8	E 248	E 241	E 231	E 227	E 227	E 227	-1.162
Nqo8	E 294	R 286	R 297R	R 274	R 274	R 274	-1.219

<sup>a</sup> Although Nqo6 E66 is not conserved (a gap may be found in the alignment), the side chain of PSST D80 in the mammalian enzymes occupies the same spacial position as Nqo6 E66 in *T. Thermophilus*.

Table S5: Point mutations in residues near the Q-chamber with the assembly of the mutated protein complex (or content of complex I), enzymatic activity of the mutant in relation to wild-type and associated human disease. Data taken from [8], except when explicitly referenced.

Subunit	Residue in <i>T. Thermophilus</i>	Mutated to	Assembly	Activity	Associated disease
Nqo4	Y87	F/W[60]	Yes	10%	
Nqo6	W37	A[47]	Yes	40%	
		I	Reduced	40%	
		E	Reduced	< 5%	
Nqo8	E35	K	Reduced	6%	LHON and MELAS <sup>a</sup>
		A/Q	Reduced	18-27%	
		D	Yes	> 50%	
		D	Yes	55%	
	R36	A	No	0%	MELAS
		K	Reduced	10%	
	R45	A	No	0%	Hypertrophic cardiomyopathy
		K	Reduced	20%	
	D62	A/N	No	0%	
		E		25-90%	
	A63	T	Yes	90%	LHON
	K65	A		30%	
	E163	K	Yes	80%	LHON
		A	Yes	24%	
	E213	D/Q/V		> 30%	
	R216	A/K/H/F		> 30%	
		Q			Progressive encephalomyopathy and Leigh syndrome
	D220	A	Reduced	10%	
		E		40%	
	E223	A		> 60%	
	E225	A		35%	
	E248	A/Q	Yes	> 60%	
	R294	A		60%	

<sup>a</sup> LHON is Leber's Hereditary Optical Neuropathy and MELAS is Mitochondrial Encephalomyopathy, Lactic Acidosis and Stroke-like episodes.



# 4 Modeling the Hydrolysis of Iron-Sulfur Clusters

Murilo H. Teixeira<sup>a</sup>, Felipe Curtolo<sup>a</sup>, Sofia R. G. Camilo<sup>a</sup>, Martin J. Field<sup>b,c</sup>, Peng Zheng<sup>d</sup>, Hongbin Li<sup>e</sup> and Guilherme M. Arantes<sup>a</sup>

<sup>a</sup>Department of Biochemistry, Instituto de Química, Universidade de São Paulo, SP, Brazil

<sup>b</sup>CNRS, CEA, IRIG, Laboratoire de Chimie et Biologie des Métaux, Université Grenoble Alpes, Grenoble, France

<sup>c</sup>Institut Laue-Langevin, Grenoble, France

<sup>d</sup>State Key Laboratory of Coordination Chemistry, School of Chemistry and Chemical Engineering, Nanjing University, Jiangsu, China

<sup>e</sup>Department of Chemistry, University of British Columbia, British Columbia, Canada

Reprinted with permission from Teixeira, M. H.; Arantes, G. M. Modeling the Hydrolysis of Iron-Sulfur Clusters. *J. Chem. Inf. Model.*, v. 60, p. 653-660, 2020.

Copyright 2020 American Chemical Society.

# Modeling the Hydrolysis of Iron–Sulfur Clusters

Murilo H. Teixeira,<sup>†</sup> Felipe Curtolo,<sup>†</sup> Sofia R. G. Camilo,<sup>†</sup> Martin J. Field,<sup>‡,§,||</sup> Peng Zheng,<sup>||</sup> Hongbin Li,<sup>⊥,||</sup> and Guilherme M. Arantes<sup>\*,†,||</sup>

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

<sup>‡</sup>CNRS, CEA, IRIG, Laboratoire de Chimie et Biologie des Métaux, Université Grenoble Alpes, 17 rue des Martyrs, 38000 Grenoble, France

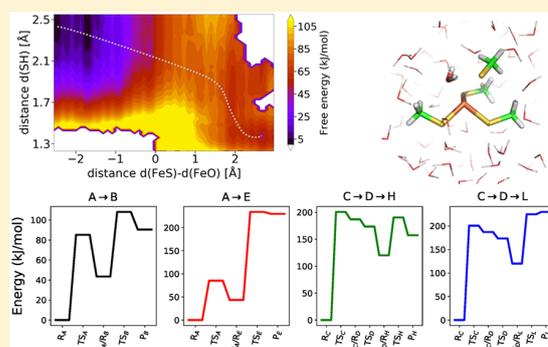
<sup>§</sup>Institut Laue-Langevin, BP 156, 41 Avenue des Martyrs, 38042 Grenoble, Cedex 9, France

<sup>||</sup>State Key Laboratory of Coordination Chemistry, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing, Jiangsu 210023, China

<sup>⊥</sup>Department of Chemistry, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada

## Supporting Information

**ABSTRACT:** Iron–sulfur (FeS) clusters are essential metal cofactors involved in a wide variety of biological functions. Their catalytic efficiency, biosynthesis, and regulation depend on FeS stability in aqueous solution. Here, molecular modeling is used to investigate the hydrolysis of an oxidized (ferric) mononuclear FeS cluster by bare dissociation and water substitution mechanisms in neutral and acidic solution. First, approximate electronic structure descriptions of FeS reactions by density functional theory are validated against high-level wave function CCSD(T) calculations. Solvation contributions are included by an all-atom model with hybrid quantum chemical/molecular mechanical (QM/MM) potentials and enhanced sampling molecular dynamics simulations. The free energy profile obtained for FeS cluster hydrolysis indicates that the hybrid functional M06 together with an implicit solvent correction capture the most important aspects of FeS cluster reactivity in aqueous solution. Then, 20 reaction channels leading to two consecutive Fe–S bond ruptures were explored with this calibrated model. For all protonation states, nucleophilic substitution with concerted bond breaking and forming to iron is the preferred mechanism, both kinetic and thermodynamically. In neutral solution, proton transfer from water to the sulfur leaving group is also concerted. Dissociative reactions show higher barriers and will not be relevant for FeS reactivity when exposed to solvent. These hydrolysis mechanisms may help to explain the stability and catalytic mechanisms of FeS clusters of multiple sizes and proteins.



## 1. INTRODUCTION

Iron and sulfur are abundant elements on the Earth's surface. They were recruited in the form of iron–sulfur (FeS) clusters early during the evolution of life resulting in ancient proteins with the ability to catalyze electron transfer reactions.<sup>1</sup> Essential biological processes such as cellular respiration and photosynthesis rely on enzymes equipped with FeS clusters. In fact, the largest class of metalloproteins comprise proteins which carry FeS clusters as cofactors.<sup>2</sup> From structural and electronic points of view, FeS clusters sit between transition metal atoms and solid surfaces.<sup>3</sup> Usually, clusters are formed by 1 to 8 iron nuclei (mostly in ferric form but may be reduced to ferrous) bridged by inorganic sulfide anions and connected to the protein scaffold by thiolate groups in cysteine side chains.

The stability and biosynthesis of FeS clusters will naturally depend on their reactivity in aqueous solution.<sup>4</sup> Catalytic mechanisms and regulation of enzymes enclosing FeS clusters

are also modulated by the connectivity of Fe–S bonds. In nitrogenase, sulfide hemilability has been observed as an intermediate step for H<sub>2</sub> exchange in the reduction cycle of the FeMo cofactor.<sup>5,6</sup> In aconitase, the FeS cluster is exposed in the active site to directly coordinate the citrate substrate and may break down in response to cellular levels of iron.<sup>7</sup>

The reactivity of protein-bound FeS clusters with molecular oxygen has been studied experimentally in detail since an oxidative environment may convert exposed FeS clusters to unstable species that quickly decompose.<sup>6–8</sup> But their stability upon hydrolysis and substitution reactions has received less attention, partly because FeS clusters are often buried in

**Special Issue:** Molecular Simulation in Latin America: Coming of Age

**Received:** September 30, 2019

**Published:** December 2, 2019

proteins and inaccessible to water. Synthetic FeS analogues have been studied extensively and provided details of ligand exchange and protonation chemistry.<sup>8,9</sup> Another elegant approach based on single-molecule force spectroscopy has been proposed recently to partially unfold and expose FeS clusters, allowing us to study their reactivity in biologically relevant conditions.<sup>10–13</sup>

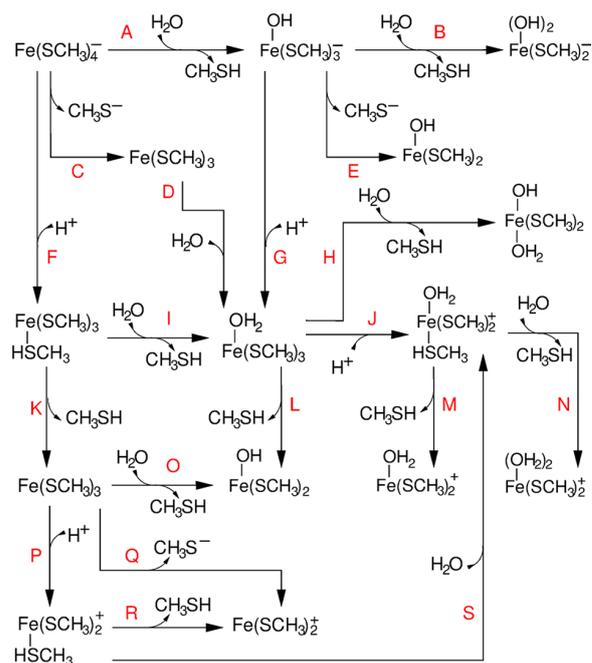
Quantum chemical calculations including environmental effects may also be used to investigate the reactivity of transition metal complexes in solution.<sup>14</sup> But modeling FeS clusters is particularly challenging. Besides the multiple Fe–S bonds, several protonation, oxidation, and spin states may be populated,<sup>3</sup> contributing to an explosion in the number of possible reactive channels. Their electronic structures show many low-lying and near-degenerate states that may cross, leading to multiple-state reactivity.<sup>15,16</sup> In polynuclear FeS clusters, strong electron correlation and long-range spin coupling effects complicate enormously the theoretical description<sup>17,18</sup> such that electronic structure calculations on FeS clusters have been proposed as an example where upcoming quantum computers could exceed the power of classical supercomputers.<sup>19</sup>

Here, we look into the stability of FeS clusters starting with a rather pedestrian model, a mononuclear iron center bound to four thiolate ligands, [1Fe–4S]. This is the simplest FeS cluster and reflects the cofactor found in rubredoxin, an electron-transfer protein in Gram-negative bacteria.<sup>20</sup> Applying several electronic structure methods, we investigate exhaustively the reactions of the model compound  $\text{Fe}(\text{SCH}_3)_4$  with water or bare dissociation in various protonation states, and mechanisms up to the second Fe–S bond rupture, as described in Figure 1.

The remainder of the text is organized as follows. Given the intricate electronic structure of FeS clusters, we initially benchmark the performance of approximate density functional theory (DFT) against high-level quantum chemical (QC) calculations. Free energy contributions and solvation effects captured by explicit all-atom hybrid QM/MM simulations are compared to models using implicit solvation to find an efficient yet reliable DFT description of FeS stability in aqueous solution. Readers only interested in FeS reactivity can skip all methodological details and model calibration and jump to Section 3.3 where FeS hydrolysis reactions are discussed.

## 2. METHODS

**2.1. Quantum Chemical Calculations.** All molecular geometries for stationary points (reactants, products, and transition states, TS) were optimized with the OLYP<sup>22,23</sup> functional, the 6-31+G(2df,p)<sup>24</sup> basis set, and the PCM implicit water solvation model<sup>25</sup> with the Gaussian 09 program (rev. A1).<sup>26</sup> Previous calculations showed this functional provides a reasonable description of Fe–S bond dissociation.<sup>27</sup> Two-dimensional scans were used to search for TS in substitution (AnDn) reactions, restraining a linear combination of forming (Fe–O) and breaking (Fe–S) bonds and the proton coordinate when this atom was transferred to the sulfur leaving group (see more details of reaction coordinates below). These scans were done with the pDynamo library<sup>28</sup> version 1.9 interfaced with ORCA program version 3.0.1,<sup>29</sup> the OLYP functional, the def2-SVP basis set,<sup>30</sup> and the COSMO implicit water model.<sup>31</sup> Only electrostatic contributions were included in the implicit solvation models. Iron compounds were always in the oxidized and sextet spin states.



**Figure 1.** Scheme of the iron–sulfur reactions studied here. Substitution reactions (AnDn in the current nomenclature<sup>21</sup>) are denoted with concomitant water input and leaving group output arrows. Dissociation reactions (Dn) show only the output leaving group arrow. Besides the protonation steps, reaction S is the only addition reaction (An) shown. Molecules  $\text{Fe}(\text{SCH}_3)_3$  (formed by reactions C and K) and  $\text{Fe}(\text{OH})(\text{SCH}_3)_2$  (formed by reactions E, L, and O) are repeated to avoid an even clumsier scheme.

Zero-point energies and thermal and entropic contributions were evaluated for stationary points using vibrational frequencies and the harmonic oscillator and rigid rotor approximations.<sup>32</sup> No empirical scaling was applied to frequencies.<sup>33</sup>

Single-point energies for isolated optimized geometries were obtained with ORCA and the following functionals: OLYP, OPBE,<sup>22,34</sup> B3LYP,<sup>23,35</sup> TPSS,<sup>36</sup> M06,<sup>37</sup> M06L,<sup>38</sup> B97, and  $\omega$ B97X.<sup>39,40</sup> Spin-polarized orbitals and standard integration grids were adopted. The def2-TZVP basis set<sup>30</sup> and resolution of identity with the TZV/J<sup>41</sup> auxiliary basis were used. A second-order SCF optimization had to be activated to obtain convergence in several cases. All calculations were carried out with  $C_1$  point-group symmetry. Dispersion interactions were added to some DFT functionals with Grimme's D3 correction<sup>42</sup> using the Becke and Johnson damping.<sup>43,44</sup> Relativistic effects were evaluated at the M06/def2-TZVP level with the Douglas–Kroll–Hess (DKH) approximation.

A semiempirical potential specifically parametrized to model Fe–S bond dissociation<sup>27</sup> based on the PM6 general parametrization<sup>45</sup> with *d*-orbitals was also tested. Complete active space configuration interaction calculations<sup>46</sup> with seven electrons in seven orbitals were performed with orbitals determined from restricted-open shell (ROHF) calculations with fractional occupation.<sup>47</sup> Semiempirical calculations were done with the pDynamo library.<sup>27</sup>

Wave function CCSD(T)<sup>46</sup> calculations with the domain-based local pair natural orbital (DLPNO) approximation,<sup>48</sup> ROHF orbitals, and auxiliary basis def2-TZVP/C were

obtained with ORCA version 4.1.1<sup>49</sup> and employed as reference energies for calibrations of the DFT methods.

**2.2. Hybrid QM/MM Free Energy Simulations in Explicit Solvent.** Reaction A (Figure 1) was simulated in explicit water with an all-atom model containing  $\text{Fe}(\text{SCH}_3)_4^-$ , 1689 water molecules, and a  $\text{Na}^+$  ion to neutralize the system in a cubic box of 3.72 nm side. The system was initially relaxed during 2 ns of molecular dynamics simulation with a classical molecular mechanical (MM) potential and periodic boundary conditions using previous parameters for the FeS center, CHARMM36<sup>50</sup> for  $\text{Na}^+$ , and TIP3P for water.<sup>51</sup> Then, all water molecules further than 16 Å from the Fe center were frozen, and the remainder of the system relaxed again for 5 ps of molecular dynamics with a QM/MM potential,<sup>52</sup> where only  $\text{Fe}(\text{SCH}_3)_4^-$  and a harmonically restrained nearby water molecule were treated in the QC region at the OLYP/def2-SVP level of theory and all other atoms treated in the MM region. A standard electrostatic QM/MM embedding without cut offs or long-range electrostatic corrections was used as implemented in the pDynamo library 1.9 interfaced with ORCA 3.0.1.<sup>28</sup>

This system and QM/MM potential were used for umbrella sampling (US)<sup>53</sup> with molecular dynamics simulations of a two-dimension free energy profile for reaction A. Two reaction coordinates were employed: The difference between the distances of breaking (Fe–S) and forming (Fe–O) iron bonds,  $d(\text{FeS})-d(\text{FeO})$ , described the iron reaction, and the distance between a proton initially bound to water and the sulfur of the leaving group,  $d(\text{SH})$ , described the coupled proton transfer.<sup>54,55</sup> Reaction coordinates were explored between ranges  $-2.5 \leq d(\text{FeS})-d(\text{FeO}) \leq 2.9$  Å and  $1.3 \leq d(\text{SH}) \leq 2.5$  Å, separated by steps of 0.2 and 0.3 Å, restrained with harmonic potentials with force constant  $k_{\text{umb}} = 1000$  and  $500$  kJ/mol Å<sup>-2</sup>, respectively, for the iron and the proton coordinates. A total of 146 US windows were sampled during 6 ps each, resulting in an aggregate simulation time with the QM(DFT)/MM potential of 0.9 ns. A Langevin dynamics integrator was used with a time step of 1 fs, friction coefficient  $\gamma = 25$  ps<sup>-1</sup>, and temperature of 300 K.<sup>56</sup> The two-dimensional free energy surface was pieced together using the weighted histogram analysis method (WHAM)<sup>57</sup> with the initial 1 ps of each US window discarded for equilibration. Statistical uncertainties were estimated as 95% confidence intervals by bootstrap analysis.<sup>58</sup>

### 3. RESULTS AND DISCUSSION

**3.1. Performance of Approximate Electronic Structure Methods.** Structures of transition-metal complexes determined with DFT methods are generally in good agreement with experiments.<sup>59</sup> However, iron–sulfur clusters show strong correlation effects that may not be described correctly by standard functionals. Previously, it was shown that left–right correlation important for the dissociation of Fe–S bonds is appropriately included in the OLYP functional,<sup>27</sup> so this method was employed here for geometry optimization and sampling. The quality of obtained structures may be assessed by the calculated Fe–S bond length for  $\text{Fe}(\text{SCH}_3)_4^-$  (2.32 Å) which agrees reasonably with the Fe–S lengths (2.27–2.30 Å) observed in the high-resolution crystal structure of the [1Fe-4S]-containing rubredoxin protein (PDB ID 8RXN).<sup>20</sup>

Table 1 shows the performance of several DFT functionals for calculating relative reaction energies and barriers in comparison to the gold-standard electronic structure method

**Table 1. Performance of DFT and Semiempirical Methods in Comparison to the CCSD(T) Reference<sup>a</sup>**

Functional	MUE	MAE
B3LYP	19.5	60.8
B3LYP-D3	9.2	28.9
B97-D3	13.9	31.8
M06	6.8	16.9
M06L	8.2	29.4
M06L/SVP	13.7	45.3
OLYP	27.4	81.5
OLYP/SVP	19.7	63.2
OPBE	24.9	69.0
PM6R	176.6	481.4
TPSS	13.7	52.7
$\omega$ B97X-D3	4.3	26.9

<sup>a</sup>Mean unsigned error (MUE) and maximum absolute error (MAE) are shown for relative energies (in kJ/mol) of 34 TS and product species in FeS cluster reactions shown in Figure . All DFT calculations were done with the def2-TZVP basis set, except M06L/SVP and OLYP/SVP were done with def2-SVP.

CCSD(T)<sup>46</sup> for 34 stationary points (products and TS) of all reactions in Figure 1. No solvent contribution was included in this comparison. The lowest mean error is observed for the  $\omega$ B97X-D3 range-separated hybrid functional with dispersion corrections, followed by the hybrid M06 functional, which shows the lowest maximum error. The MAE observed for all functionals tested correspond to the TS of reactions C and E where left–right (multiconfigurational) correlation is significant.<sup>27</sup> The performance of B3LYP-D3 and M06L is also good. Comparison with B3LYP shows that addition of dispersion corrections is important and justifies the good performance of M06 and M06L which account for dispersion in the original parametrizations.<sup>37</sup> Fortuitous error cancelation when using OLYP with smaller split-valence basis sets suggests that employing this level of theory should give reasonable results for geometry optimizations and sampling. Using a split-valence basis with M06L does not lead to such error cancelation and degrades the performance of this functional.

Possible exceptions where dispersion effects may be essential for a correct structural description are TS for addition (An) and dissociation (Dn) steps involving neutral attacking or leaving groups, such as TS for reactions D, K, L, M, R, and S (Figure 1). These species were reoptimized with the M06/6-31+G(2df,p) level of theory. Energy differences between reoptimized structures and the original OLYP geometries are smaller than 5 kJ/mol for  $\text{TS}_R$  and  $\text{TS}_S$ , around 10 kJ/mol for  $\text{TS}_D$ ,  $\text{TS}_K$ , and  $\text{TS}_M$  and 30 kJ/mol for  $\text{TS}_L$ . The considerable difference for the latter species is due to an incomplete optimization done with OLYP and the floppy nature of this TS. These corrections were implemented in the results presented in Section 3.3 but do not change any of the qualitative conclusions. Thus, even for FeS species where dispersion plays a major role, geometries obtained with the more approximate functional OLYP prove reasonable. Dispersion effects on geometries will be significantly smaller for reactant and product species, as well as for TS of the substitution (AnDn) reactions where stronger contributions such as electrostatic effects from  $\text{H}^+$  transfer dominate.

The PM6R semiempirical method specifically calibrated for Fe–S dissociation reactions and 3 orders of magnitude faster than DFT methods was also tested.<sup>27</sup> Unfortunately, the

method has unacceptable performance (Table 1) for all substitution reactions that involve oxygen atoms because this element was not included in the PM6R reparametrization. Relativistic corrections within the DHK framework were also tested, but their maximum contribution for relative energies relevant for FeS reactivity is only 3 kJ/mol, with an average of 1 kJ/mol. Thus, relativistic corrections were ignored in the remainder of this study.

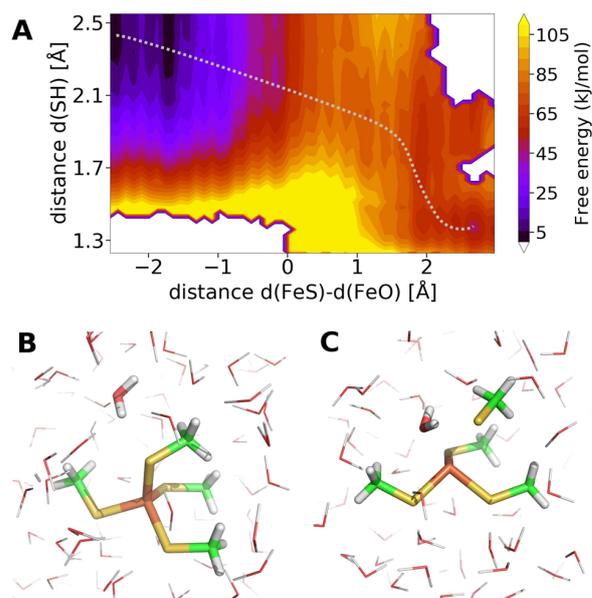
Previously, the bare dissociation of Fe–S bonds was suggested to proceed through crossings between different spin states (quartets).<sup>16</sup> However, for the substitution reactions with water and the leaving group protonations (CH<sub>3</sub>SH) studied here, the energy gap between sextet and quartet states is more than 50 kJ/mol.<sup>12</sup>

An analysis of computational timings (Table S1) is necessary to find the best relation between accuracy and computational cost. The efficiency of the DLPNO–CCSD(T) method is impressive and comparable to much more approximate hybrid and range-separated functionals. However, it is still impossible to run geometry optimizations with the DLPNO approximation since analytic gradients are not implemented, and sampling millions of geometries necessary for free energy simulations would be prohibitively expensive. The hybrid functionals  $\omega$ B97X–D3 and M06 give the best performance and may be applied for geometry optimizations but are still too expensive for sampling. The efficiency of M06L and other generalized gradient approximated (GGA) functionals is 1 order of magnitude better than the best performance hybrid functionals but still too demanding for sampling. Acceptable efficiency is obtained with a split-valence basis set, so the OLYP/def2–SVP level was used as a reasonable compromise between accuracy and cost-effective DFT treatment for the hybrid potential free-energy simulations.

**3.2. Free Energies and Solvation Effects.** The effect of water solvation on FeS reactivity was examined here for reaction A using two solvent models. This reaction was chosen as representative because it involves charged species and is important for FeS stability in aqueous solution (Section 3.3). We applied an all-atom explicit solvent model with a QM/MM hybrid potential description,<sup>52,54,55</sup> where individual water molecules may interact through electrostatic and van der Waals forces directly with the reactive molecules and polarize their electronic structure<sup>52</sup> (see Figure 2B for a model snapshot). This was compared with a more approximate implicit solvent model<sup>25,60</sup> where the electron cloud from reactive molecules are polarized by the dielectric response of a continuum that does not describe water structure or specific contacts (such as hydrogen bonds).

Reaction A actually involves two processes: iron transfer, i.e., breaking one Fe–S bond and forming the Fe–O bond with water, and proton transfer from water to the leaving group sulfur atom. These were described by distance reaction coordinates  $d(\text{FeS})-d(\text{FeO})$  and  $d(\text{SH})$ , respectively.

The two-dimensional free energy surface in Figure 2 shows that reaction A proceeds through a typical nucleophilic substitution mechanism ( $S_N2$  or  $A_nD_n$  in the updated nomenclature) with a late TS, where the Fe–S bond is largely broken and the Fe–O bond is partially formed. The transferred proton is still fully bound to water and coordinated with the leaving group sulfur (at 1.98 Å distance, Table 2). A snapshot of the TS structure is shown in Figure 2C. In the reactant state, the reactive water hydrogen bonds to the leaving group sulfur (Figure 2B), and the surface topology is rather



**Figure 2.** Hydrolysis of  $\text{Fe}(\text{SCH}_3)_4^-$  (reaction A in Figure 1) in explicit solvent. (A) Two-dimensional free energy profile for the reaction in aqueous solution obtained with a hybrid QM/MM potential. Gray dashed line shows the approximate minimum free energy pathway. Snapshots of the (B) reactant and (C) transition states show the reactive atoms in sticks with Fe in orange, S in yellow, O in red, and C in green.

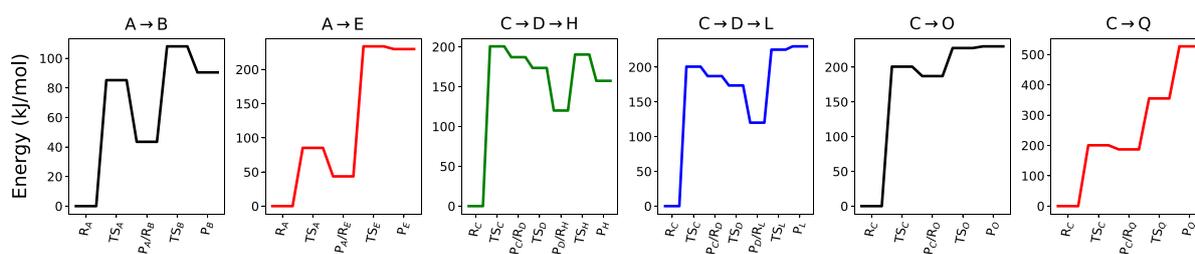
**Table 2. Free Energies ( $\Delta G$ , kJ/mol) and Reaction Coordinates (Å) for Stationary Species of Reaction A in Explicit (QM/MM, as shown in Figure 2) and Implicit (M06+COSMO) Solvents**

	Reactant	TS	Product
QM/MM			
$d(\text{SH})^a$	2.40	1.98	1.39
$d(\text{FeS})-d(\text{FeO})^a$	-2.43	1.28	2.75
$\Delta G$	0.0	$82 \pm 5$	$51 \pm 5$
M06+COSMO			
$d(\text{SH})^b$	2.47	1.79	1.35
$d(\text{FeS})-d(\text{FeO})^b$	-1.60	0.89	3.68
$\Delta G$	0.0	85	43

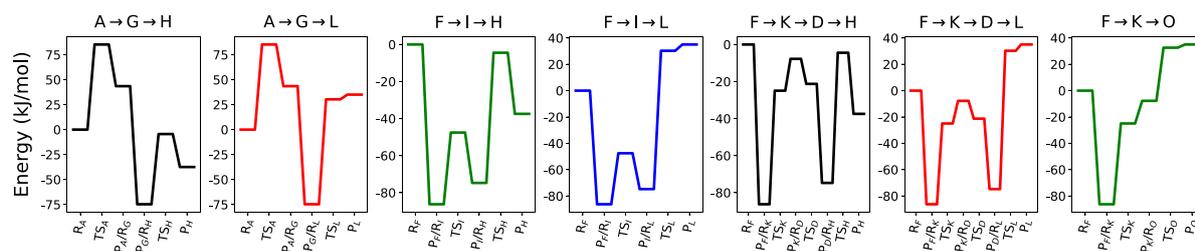
<sup>a</sup>Mean values from state basins. Standard deviations are 0.04–0.05 Å for all states, except for  $d(\text{FeS})-d(\text{FeO})$  in the reactant state where the deviation is 0.16 Å. <sup>b</sup>Geometries of ion–molecule complexes were used for reactant and product states.

shallow because of water flexibility. But the surface rises up steeply for proton transfer if the Fe–S bond is not significantly broken, as expected from the basicity of the bound thiolate group in comparison to water (or its conjugate base, hydroxide). The TS region is rather flat [ $0.5 < d(\text{FeS})-d(\text{FeO}) < 1.5$  Å and  $1.7 < d(\text{SH}) < 2.1$  Å] with the minimum free energy pathway passing through a barrier of  $82 \pm 5$  kJ/mol (Table 2). The product basin is deep with a reaction free energy of  $51 \pm 5$  kJ/mol. Thus, reaction A is not spontaneous with slow activation kinetics (in relation to the thermal energy).

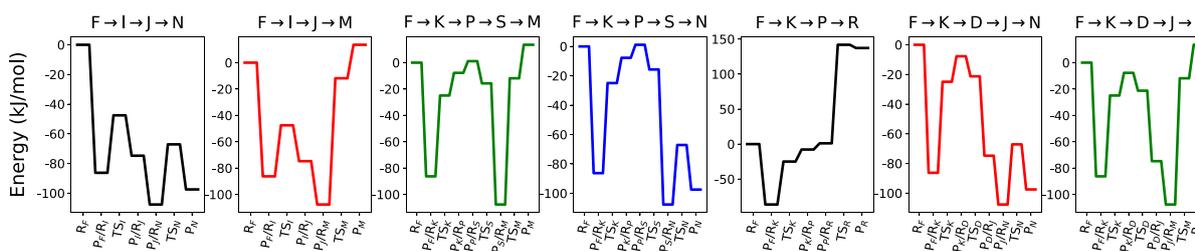
Reaction A in implicit solvent also follows an  $A_nD_n$  mechanism with a late TS (Table 2). Proton reaction coordinates at stationary species are equivalent to those found in the explicit model. Differences in the iron reaction



**Figure 3.** Reaction profiles for hydrolysis of oxidized mononuclear FeS cluster model in neutral aqueous solution. Reaction combinations are shown on top of each profile using the letter code given in Figure 1. Reaction species are shown in the abscissa using R for reactant, TS for transition state, and P for product, with the corresponding reaction subscripted. Energies were obtained with M06/def2-TZVP and the COSMO implicit water solvent, and geometries were obtained with OLYP/6-31+G(2df,p) and the PCM implicit water solvent.



**Figure 4.** Reaction profiles for the hydrolysis of oxidized mononuclear FeS cluster model with one protonation, corresponding to mildly acidic solution. See the legend in Figure 3 for more details.



**Figure 5.** Reaction profiles for the hydrolysis of oxidized mononuclear FeS cluster model with two protonations, corresponding to highly acidic solution. See the legend in Figure 3 for more details.

coordinate are due to water flexibility and the shallow free energy surface observed above (Figure 2A). Thus, similar reaction mechanisms are obtained in both solvent models when using the same electronic structure method (OLYP with split-valence basis). Geometrical differences are due to the underlying topology of the free energy surface and lack of solvent hydrogen bonding in the implicit model.

Free energies in implicit solvent usually include electronic, thermal, and entropic effects from the reactive molecules and solvent contributions.<sup>60</sup> The M06/def2-TZVP level of theory gives high quality electronic energies for FeS reactions (Table 1), so it was used here with the COSMO implicit solvent model<sup>31</sup> that accounts for the electrostatic component of solvation free energies. Zero-point energies and thermal (enthalpic and entropic) contributions from reactive centers may be calculated by harmonic oscillator and rigid rotor approximations. However, these contributions are relatively small (Table S2) and may be unreliable for entropic effects when low-frequency vibrational modes are present.<sup>61</sup> This is indeed the case for TS<sub>A</sub> resulting in spurious entropies (Table S2). Thus, thermal and entropic effects were not included, and relative free energies in our implicit water model contained only electronic and solvent contributions.

Table 2 shows that reaction and barrier free energies are within 8 kJ/mol (or 3 kJ/mol considering the statistical uncertainty of free energy simulations) between the two solvent models. This good agreement is partially fortuitous, due to error cancelation and inclusion of different contributions. Nevertheless, we may conclude from these two sections of model calibration that geometries determined by the OLYP functional with a split-valence basis set and energetics obtained at the M06/def2-TZVP level together with implicit solvation provide a reasonably reliable and efficient procedure. This level of theory was used to explore the reactivity of FeS clusters in aqueous solution in the next section.

**3.3. Hydrolysis and Protonation Reactions.** Consecutive cleavage of two Fe–S bonds should be enough to disrupt mononuclear and polynuclear FeS clusters. Here, all feasible combinations of the 19 reactions shown in Figure 1 that start with the ferric mononuclear FeS cluster model and lead to two Fe–S bond ruptures were considered. Reaction combinations were divided in groups without, with one, and with two protonations of reactive species. These groups may correspond to reaction sequences in neutral, mildly acidic, and highly acidic aqueous solutions, respectively.

Given the high barrier (see below) and the unlikely event of thiolate dissociation at low pH, no combination involving a

thiolate anion in acidic media (e.g.,  $C \rightarrow D \rightarrow J \rightarrow N$  or  $F \rightarrow K \rightarrow Q$ ) was examined. TS for protonation reactions were not considered as barriers for proton transfer and were assumed to be smaller than other reactions. This is certainly true for proton addition to anions but may not hold for reactions J and P, where protonation and hydrolysis may be concerted. After applying the above criteria, there are 20 possible reaction combinations for which free energy profiles of reactions are shown in Figures 3, 4, and 5.

In neutral aqueous solution (Figure 3), nucleophilic substitution reactions (A and B) have high barriers (85 and 65 kJ/mol, respectively). These may be thermally activated and lead to products more stable than dissociative reactions (C, E, and Q), which also have much higher barriers (200, 190, and 170 kJ/mol). Thus, the  $A \rightarrow B$  reaction sequence will account for the rupture of up to two Fe–S bonds in FeS clusters, and no thiolate dissociation will be observed in neutral solution.<sup>12</sup> This also suggests that spin-crossings previously identified in thiolate dissociation (Dn) reactions of stretched rubredoxin<sup>16</sup> should not contribute to the reactivity of FeS exposed to water. In both A and B reactions, a late TS is found with the proton transferred from the attacking water to the leaving group simultaneously with the reaction at the iron center. Details of this concerted reaction were given above for reaction A in explicit solvent.

Reaction profiles are significantly stabilized when one reactive species is protonated, such as in mildly acidic solution or when only part of a FeS cluster is exposed to acidic solution (Figure 4). If protonation takes place after the first Fe–S bond rupture, the barrier for the second rupture via dissociative reaction L is 105 kJ/mol, lower than thiolate dissociation but still higher than the barrier for the substitution (AnDn) reaction H, 70 kJ/mol.  $P_H$  is also 75 kJ/mol more stable than  $P_L$ ; thus, the sequence  $A \rightarrow G \rightarrow H$  is kinetic and thermodynamically preferred. Similar to  $TS_A$  discussed above,  $TS_H$  is also late with the Fe–S bond largely broken and the Fe–O bond almost formed [ $d(\text{FeS})-d(\text{FeO}) = 1.2 \text{ \AA}$ ]. The transferred proton is still bound to water and coordinates the sulfur in the leaving group [ $d(\text{SH}) = 2.0 \text{ \AA}$ ].

In acidic solution, it is most likely that protonation will take place before bond rupture, as soon as the FeS cluster is exposed to solvent. Protonation stabilizes the mononuclear FeS cluster by 85 kJ/mol but also facilitates the Fe–S bond rupture. Reactions I (AnDn) and K (Dn) may compete as their barriers are 40 and 60 kJ/mol, respectively. But  $P_I$  is much more stable than  $P_K$ . Protonation of the leaving group advances Fe–S rupture in  $TS_I$  resulting in an early TS [ $d(\text{SH}) = 1.4 \text{ \AA}$  and  $d(\text{FeS})-d(\text{FeO}) = 0.2 \text{ \AA}$ ]. The second Fe–S rupture follows the discussion of the previous paragraph, so the sequence  $F \rightarrow I \rightarrow H$  should be observed in mildly acidic solution. Other reaction combinations will give unstable products.

Bond cleavage in FeS clusters fully exposed to acidic solution or in highly acidic media will proceed with both thiol leaving groups protonated (Figure 5). The first Fe–S rupture should occur via reaction I, as in the previous paragraph. The second Fe–S rupture will proceed via the substitution (AnDn) reaction N with a barrier of 40 kJ/mol, whereas the dissociative reaction M has a high barrier, 100 kJ/mol.  $TS_N$  is again a late TS [ $d(\text{SH}) = 1.4 \text{ \AA}$  and  $d(\text{FeS})-d(\text{FeO}) = 1.0 \text{ \AA}$ ] but with dissociative character, as Fe–S and Fe–O bonds are almost not formed. Note that the first leaving group protonation will halve reaction barriers of substitution reactions (e.g.,  $TS_A \times$

$TS_I$ ) or even cut to one-third those of dissociative reactions ( $TS_C \times TS_K$ ). But a second protonation has no effect on the barrier of substitution reactions and actually increases the barrier of dissociative reactions (M and R). We conclude the combination  $F \rightarrow I \rightarrow J \rightarrow N$  is preferred in highly acidic solution. Sequences  $F \rightarrow K \rightarrow P \rightarrow S \rightarrow N$  and  $F \rightarrow K \rightarrow D \rightarrow J \rightarrow N$  will lead to the same products, but they have to climb higher barriers and hence will be slower.

## 4. CONCLUSIONS

Iron–sulfur reactivity is ubiquitous and fundamental in biological processes but inherently difficult to study both using experiments or theory. Here, we presented a modeling study of the hydrolysis reactions of a mononuclear oxidized FeS cluster [1Fe-4S], as found in the rubredoxin protein. Several molecular electronic structure methods were compared with the CCSD(T) gold-standard level of theory in order to find suitable descriptions of FeS reactivity, following our previous studies.<sup>12,15,16,17,27</sup> Solvation and free energy contributions were evaluated with an all-atom hybrid QM/MM potential giving an accomplished view of water substitution reactions on iron and allowing the calibration of more approximate implicit solvent models. We found that geometries determined by the OLYP functional with a split-valence basis set and energetics obtained with a hybrid functional such as M06 or  $\omega$ B97X-D3 and a triple- $\zeta$  basis set, both together with implicit solvation, provide a reasonably reliable and efficient procedure to model FeS reactivity in solution.

Hydrolysis channels of the FeS center with up to two consecutive Fe–S bond ruptures in several protonation states were investigated exhaustively with this calibrated model. Substitution with a concerted TS (AnDn mechanism) is the preferred mechanism, kinetic and thermodynamically, for all protonation states. Purely dissociative (Dn) mechanisms should not play a role for FeS reactivity when clusters are exposed to solvent. For rate-limiting steps, the TS character is determined by a FeS cluster charge and leaving group protonation. For charged clusters (reactions A and N) or for a neutral center with a deprotonated leaving group (reaction H), the TS shows a largely broken Fe–S bond. Fe–O bond forming is advanced when the leaving group is deprotonated (A and H) or late, otherwise (I and N).

These results clearly indicate that functional FeS clusters should be buried in proteins not only to protect from oxidative degradation but also from hydrolysis reactions, especially under acid catalysis. Because of the lower barriers found here after protonation, it may also be suggested that enzymes employ general-acid catalysis to synthesize FeS clusters<sup>4</sup> and to change iron coordination, such as sulfide hemilability in nitrogenase.<sup>5,6</sup> The theoretical methods and reaction mechanisms depicted here will be general and helpful to understand the FeS reactivity in aqueous solution for multiple cluster sizes and proteins partners.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jcim.9b00881>.

Two tables with computational timings and with thermal and entropic contributions (PDF)

## AUTHOR INFORMATION

## Corresponding Author

\*E-mail: garantes@iq.usp.br.

## ORCID

Martin J. Field: 0000-0001-8674-7997

Peng Zheng: 0000-0003-4792-6364

Hongbin Li: 0000-0001-7813-1332

Guilherme M. Arantes: 0000-0001-5356-7703

## Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

Funding from FAPESP (Grants No. 2016/24096-5 and No. 2018/08311-9) is gratefully acknowledged. P.Z. is thankful for the funding from the National Natural Science Foundation of China (Grant No. 21771103) and Natural Science Foundation of Jiangsu Province (Grant No. BK20160639).

## REFERENCES

- (1) Beinert, H. Iron-Sulfur Proteins: Ancient Structures, Still Full of Surprises. *JBC, J. Biol. Inorg. Chem.* **2000**, *5*, 2–15.
- (2) Beinert, H.; Holm, R. H.; Munck, E. Iron-Sulfur Clusters: Nature's Modular, Multipurpose Structures. *Science* **1997**, *277*, 653–659.
- (3) Noodleman, L.; Peng, C.; Case, D.; Mouesca, J.-M. Orbital Interactions, Electron Delocalization and Spin Coupling in Iron-Sulfur Clusters. *Coord. Chem. Rev.* **1995**, *144*, 199–244.
- (4) Layer, G.; Ollagnier-de Choudens, S.; Sanakis, Y.; Fontecave, M. Iron-Sulfur Cluster Biosynthesis. *J. Biol. Chem.* **2006**, *281*, 16256–16263.
- (5) Raugei, S.; Seefeldt, L. C.; Hoffman, B. M. Critical Computational Analysis Illuminates the Reductive Elimination Mechanism that Activates Nitrogenase for N<sub>2</sub> Reduction. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, E10521–E10530.
- (6) Sippel, D.; Rohde, M.; Netzer, J.; Trncik, C.; Gies, J.; Grunau, K.; Djurdjevic, I.; Decamps, L.; Andrade, S. L. A.; Einsle, O. A Bound Reaction Intermediate Sheds Light on the Mechanism of Nitrogenase. *Science* **2018**, *359*, 1484–1489.
- (7) Beinert, H.; Kennedy, M. C.; Stout, C. D. Aconitase as Iron-Sulfur Protein, Enzyme, and Iron-Regulatory Protein. *Chem. Rev.* **1996**, *96*, 2335–2373.
- (8) Rao, P. V.; Holm, R. H. Synthetic Analogues of the Active Sites of Iron-Sulfur Proteins. *Chem. Rev.* **2004**, *104*, S27–S60.
- (9) Henderson, R. A. Proton Transfer to Synthetic FeS-Based Clusters. *Coord. Chem. Rev.* **2005**, *249*, 1841–1856.
- (10) Zheng, P.; Li, H. Highly Covalent Ferric Thiolate Bonds Exhibit Surprisingly Low Mechanical Stability. *J. Am. Chem. Soc.* **2011**, *133*, 6791–6798.
- (11) Zheng, P.; Chou, C.-C.; Guo, Y.; Wang, Y.; Li, H. Single Molecule Force Spectroscopy Reveals the Molecular Mechanical Anisotropy of the FeS<sub>4</sub> Metal Center in Rubredoxin. *J. Am. Chem. Soc.* **2013**, *135*, 17783–17792.
- (12) Zheng, P.; Arantes, G. M.; Field, M. J.; Li, H. Force Induced Chemical Reactions on the Metal Center in a Single Metalloprotein Molecule. *Nat. Commun.* **2015**, *6*, 7569.
- (13) Nunes-Alves, A.; Arantes, G. M. Mechanical Unfolding of Macromolecules Coupled to Bond Dissociation. *J. Chem. Theory Comput.* **2018**, *14*, 282–290.
- (14) Cramer, C. J.; Truhlar, D. G. Density Functional Theory for Transition Metals and Transition Metal Chemistry. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10757–10816.
- (15) Schroder, D.; Shaik, S.; Schwarz, H. Two-State Reactivity as a New Concept in Organometallic Chemistry. *Acc. Chem. Res.* **2000**, *33*, 139–145.
- (16) Arantes, G. M.; Bhattacharjee, A.; Field, M. J. Homolytic Cleavage of Fe–S Bonds in Rubredoxin Under Mechanical Stress. *Angew. Chem., Int. Ed.* **2013**, *52*, 8144–8146.
- (17) Arantes, G. M.; Taylor, P. R. Approximate Multiconfigurational Treatment of Spin-Coupled Metal Complexes. *J. Chem. Theory Comput.* **2010**, *6*, 1981–1989.
- (18) Sharma, S.; Sivalingam, K.; Neese, F.; Chan, G. K.-L. Low-Energy Spectrum of Iron-Sulfur Clusters Directly from Many-Particle Quantum Mechanics. *Nat. Chem.* **2014**, *6*, 927–933.
- (19) Reiher, M.; Wiebe, N.; Svore, K. M.; Wecker, D.; Troyer, M. Elucidating Reaction Mechanisms on Quantum Computers. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 7555–7560.
- (20) Dauter, Z.; Sieker, L. C.; Wilson, K. S. Refinement of Rubredoxin from *Desulfovibrio vulgaris* at 1.0 Å With and Without Restraints. *Acta Crystallogr., Sect. B: Struct. Sci.* **1992**, *48*, 42–59.
- (21) Guthrie, R. D.; Jencks, W. P. IUPAC Recommendations for the Representation of Reaction Mechanisms. *Acc. Chem. Res.* **1989**, *22*, 343–349.
- (22) Handy, N. C.; Cohen, A. J. Left-Right Correlation Energy. *Mol. Phys.* **2001**, *99*, 403–412.
- (23) Lee, C.; Yang, W.; Parr, R. Development of the Colle-Salvetti Correlation Energy Formula into a Functional of the Electron Density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785–789.
- (24) Ditchfield, R.; Hehre, W.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1971**, *54*, 724–728.
- (25) Tomasi, J.; Persico, M. Molecular Interactions in Solution: An Overview of Methods Based on Continuous Distributions of the Solvent. *Chem. Rev.* **1994**, *94*, 2027–2094.
- (26) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, P. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision A.1; Gaussian, Inc.: Wallingford, CT, 2009.
- (27) Arantes, G. M.; Field, M. J. Ferric-Thiolate Bond Dissociation Studied with Electronic Structure Calculations. *J. Phys. Chem. A* **2015**, *119*, 10084–10090.
- (28) Field, M. J. The pDynamo Program for Molecular Simulations using Hybrid Quantum Chemical and Molecular Mechanical Potentials. *J. Chem. Theory Comput.* **2008**, *4*, 1151–1161.
- (29) Neese, F. The ORCA program system. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 73–78.
- (30) Weigend, F.; Ahlrichs, R. 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.* **2005**, *7*, 3297–3305.
- (31) Klamt, A. Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. *J. Phys. Chem.* **1995**, *99*, 2224–2235.
- (32) McQuarrie, D. A. *Statistical Mechanics*, 1st ed.; Harper and Row: New York, 1976.
- (33) Hehre, W. J.; Radom, L.; Schleyer, P. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*, 1st ed.; Wiley: New York, 1986.
- (34) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

- (35) Becke, A. D. Density Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648.
- (36) Tao, J. M.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta-Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91*, 146401.
- (37) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
- (38) Zhao, Y.; Truhlar, D. G. A New Local Density Functional for Main Group Thermochemistry, Transition Metal Bonding, Thermochemical Kinetics, and Noncovalent Interactions. *J. Chem. Phys.* **2006**, *125*, 194101.
- (39) Chai, J.-D.; Head-Gordon, M. Systematic Optimization of Long-Range Corrected Hybrid Density Functionals. *J. Chem. Phys.* **2008**, *128*, No. 084106.
- (40) Lin, Y. S.; Li, G. D.; Mao, S. P.; Chai, J. D. Long-Range Corrected Hybrid Density Functionals with Improved Dispersion Corrections. *J. Chem. Theory Comput.* **2013**, *9*, 263–722.
- (41) Weigend, F. Accurate Coulomb-Fitting Basis Sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.
- (42) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate *ab initio* Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
- (43) Becke, A. D.; Johnson, E. R. A Density Functional Model of the Dispersion Interaction. *J. Chem. Phys.* **2005**, *123*, 154101.
- (44) Johnson, E. R.; Becke, A. D. A Post Hartree-Fock Model of Intermolecular Interactions. *J. Chem. Phys.* **2005**, *123*, No. 024101.
- (45) Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods V: Modification of NDDO Approximations and Application to 70 Elements. *J. Mol. Model.* **2007**, *13*, 1173–1213.
- (46) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*, 1st ed.; Wiley: New York, 2000.
- (47) Granucci, G.; Toniolo, A. Molecular Gradients for Semiempirical CI Wavefunctions with Floating Occupation Molecular Orbitals. *Chem. Phys. Lett.* **2000**, *325*, 79–85.
- (48) Riplinger, C.; Pinski, P.; Becker, U.; Valeev, E. F.; Neese, F. Sparse maps-A Systematic Infrastructure for Reduced Scaling Electronic Structure Methods. II. Linear Scaling Domain Based Pair Natural Orbital Coupled Cluster Theory. *J. Chem. Phys.* **2016**, *144*, No. 024109.
- (49) Neese, F. Software Update: The ORCA Program System, Version 4.0. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*, No. e1327.
- (50) Huang, J.; MacKerell, A. D., Jr CHARMM36 All-Atom Additive Protein Force Field: Validation Based on Comparison to NMR Data. *J. Comput. Chem.* **2013**, *34*, 2135–2145.
- (51) Jørgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (52) Field, M. J.; Bash, P. A.; Karplus, M. A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations. *J. Comput. Chem.* **1990**, *11*, 700–733.
- (53) Torrie, G. M.; Valleau, J. P. Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling. *J. Comput. Phys.* **1977**, *23*, 187–199.
- (54) Arantes, G. M. The Catalytic Acid in the Dephosphorylation of the Cdk2- pTpy/CycA Protein Complex by Cdc25B Phosphatase. *J. Phys. Chem. B* **2008**, *112*, 15244–15247.
- (55) Arantes, G. M.; Ribeiro, M. C. C. A Microscopic View of Substitution Reactions Solvated by Ionic Liquids. *J. Chem. Phys.* **2008**, *128*, 114503.
- (56) Allen, M.; Tildesley, D. *Computer Simulation of Liquids*, 1st ed.; Oxford University Press: New York, 1987.
- (57) Roux, B. The Calculation of the Potential of Mean Force Using Computer Simulations. *Comput. Phys. Commun.* **1995**, *91*, 275–282.
- (58) Johnson, R. W. An Introduction to the Bootstrap. *Teaching Statistics* **2001**, *23*, 49–54.
- (59) Buhl, M.; Kabrede, H. Geometries of Transition-Metal Complexes from Density-Functional Theory. *J. Chem. Theory Comput.* **2006**, *2*, 1282–1290.
- (60) Chien, T. C.; Dias, L.; Arantes, G.; Santos, L.; Triboni, E.; Bastos, E. L.; Politi, M. J. 1-(2-Quinolyl)-2-naphthol: A New Intra-Intermolecular Photoacidphotobase Molecule. *J. Photochem. Photobiol., A* **2008**, *194*, 37–48.
- (61) Ayala, P.; Schlegel, H. B. Identification and Treatment of Internal Rotation in Normal Mode Vibrational Analysis. *J. Chem. Phys.* **1998**, *108*, 2314–2325.

# Supporting Information

## Modeling the Hydrolysis of Iron-Sulfur Clusters

Murilo H. Teixeira, Felipe Curtolo, Sofia R. G. Camilo, Martin J. Field, Peng Zheng, Hongbin Li and Guilherme M. Arantes\*

*Department of Biochemistry, Instituto de Química, Universidade de São Paulo, Av. Prof. Lineu Prestes 748, 05508-900, São Paulo, SP, Brazil; Université Grenoble Alpes, CNRS, CEA, IRIG, Laboratoire de Chimie et Biologie des Métaux, 17 rue des Martyrs, 38000 Grenoble, France; Institut Laue-Langevin, BP 156, 41 Avenue des Martyrs, 38042 Grenoble Cedex 9, France; State Key Laboratory of Coordination Chemistry, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing, Jiangsu, 210023, China; Department of Chemistry, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada*

Email: [garantes@iq.usp.br](mailto:garantes@iq.usp.br)

Table S1: Serial computational timings (in minutes) for a complete energy calculation of the isolated  $\text{Fe}(\text{SCH}_3)_4^-$  molecule with the def2-TZVP basis (except M06L/SVP and OLYP/SVP) and the ORCA program running on an Intel E5-2670 CPU.

Method	Time
B3LYP-D3	24
CCSD(T)	165
M06	24
M06L	2
M06L/SVP	0.7
OLYP	2
OLYP/SVP	0.5
$\omega$ B97X-D3	52

Table S2: Total electronic and nuclear energy in implicit solvent (M06+COSMO) and in gas phase (M06) for the same geometries, zero-point energy (ZPE), enthalpic (H) and entropic (S) contributions at 300 K in implicit solvent for stationary species along the water substitution of  $\text{Fe}(\text{SCH}_3)_4^-$  (reaction A in Fig. 1). Absolute values in *Hartree* are shown for the reactant state. For TS and product state, relative values to the reactant are shown in kJ/mol.  $S_{>900K}$  was obtained removing contributions from vibrational temperatures  $<900$  K.

	Reactant	TS	Product
M06+COSMO	-3092.681575	85.2	43.4
M06 (gas)	-3092.591188	51.7	44.0
ZPE	0.172002	3.1	-3.8
H	0.193168	-6.3	-7.3
S	0.092555	-67.1	5.0
$S_{>900K}$	0.058681	-52.4	19.4



## 5 Conclusion

Quinones and iron-sulfur clusters are fundamental cofactors of the respiratory chain. They provide chemical mechanisms for electron transport inside proteins and proton translocation across membranes. Details on these molecular processes and structural features associated to them were modelled and presented in this thesis.

Distribution of quinones with different head and tail compositions (Ch. 2, fig. 2) in a multicomponent membrane are similar. A constant equilibrium position among the series tested indicates that Q-head molecular interactions with lipid head groups and interfacial water are dominant for quinone localization. These results agree with previous simulations of ubiquinone in pure POPC membranes. This position in membranes is optimal for quinone binding in respiratory complexes, which all present their Q-binding sites entrances in equivalent regions (Ch. 2, fig. 6).

When comparing pure POPC and multicomponent membranes, not all properties are similar. Permeation across membranes is considerably different, with higher transversal diffusion rates for the mixed membrane tested. The composition mimicking the inner mitochondrial membrane presents higher levels of lipid chain unsaturation and make this type of membrane more fluid and less compact. This extra internal space allows water dragging with the Q-head during quinone flip-flop, lowering energy barriers, increasing flip-flop rates. An easier transversal diffusion is essential to keep high turnover rates in ETC, since its complexes have quinone binding sites in both membrane leaflets.

Specifically for respiratory complex I, quinone migrates from the membrane interface up to close to the N2 cluster where catalysis occurs. This diffusion happens in a long, internal and narrow chamber (Ch. 3, fig. 1). The simulations presented on chapter 3 reveals a second binding position near the chamber entrance. This intermediate binding pose may be involved in the coupling between quinone reduction and proton translocation in the membrane subunit of complex I. Additionally, an auto-inhibitory effect of UQ<sub>2</sub> molecules on complex I is likely to be associated to a simultaneous binding of two

substrates molecules (Ch. 3, sec. 4.2).

High hydration levels of the Q-chamber were also identified, in contrast with previous averaged and limited-resolution experimental structural pictures provided by X-ray and cryo-EM techniques. The Q-chamber is large enough to accommodate few water molecules upon quinone binding and a balancing of internal protein dehydration and substrate hydration regulate the whole binding energetics (Ch. 3, sec. 4.1 and fig. S4). Our main simulation results for complex I - secondary binding site and hydration of the Q-chamber - have been confirmed by experiments published after our results, confirming the predictive power of our simulations. Binding of amphipatic molecules on amphipatic sites may be likely regulated on other proteins in nature.

On chapter 4, Fe-S cluster hydrolysis mechanisms were investigated. The lowest energy and, thus, more likely mechanism to happen is a substitution with a concerted transition state (Ch. 4, fig. 2). This mechanism is preferred in all protonation states tested - i.e. pH scenarios included. Modeling in low pH presented the most favorable kinect and thermodynamically reaction (Ch. 4, fig. 5). These results suggest that enzymes must employ acid catalysis for Fe-S clusters formation, modification and degradation reactions and such clusters are likely to be buried in proteins to avoid hydrolysis to keep their functionality.

Partition and diffusion of amphipatic molecules in membranes, complex binding processes through long tunnels and metal clusters hydrolysis mechanisms may be found in several examples in nature. All methods applied and results obtained in this thesis are not limited to understand only ETC related phenomena, but may also be helpful for designing simulations to study these biochemical and biophysical processes in other similar systems.

## 6 References

- 1 NATVIG, L. Why Computational Science and Engineering Should be of Interest to Computer Scientists. *Norsk Informatikkonferanse*, v. 15, p. 1–11, 1994.
- 2 MOORE, G. E. Understanding Moore’s law: four decades of innovation. *Choice Reviews Online*, v. 44, p. 44–6278–44–6278, 2007.
- 3 ARUTE, F. et al. Quantum supremacy using a programmable superconducting processor. *Nature*, Springer US, v. 574, p. 505–510, 2019.
- 4 SHAO, Y. et al. Advances in methods and algorithms in a modern quantum chemistry program package. *Phys. Chem. Chem. Phys.*, v. 8, p. 3172–3191, 2006.
- 5 HELGAKER, T.; JØRGENSEN, P.; OLSEN, J. *Molecular Electronic-Structure Theory*. 1<sup>st</sup>. ed. New York: Wiley, 2000.
- 6 GHAFARIZADEH, A. et al. PhysiCell: An open source physics-based cell simulator for 3-D multicellular systems. *PLoS Comput. Biol.*, v. 14, p. 1–20, 2018.
- 7 GUISONI, N.; MAZZITELLO, K. I.; DIAMBRA, L. Modeling active cell movement with the Potts model. *Front. Phys.*, v. 6, 2018.
- 8 NUNES-ALVES, A.; ZUCKERMAN, D. M.; ARANTES, G. M. Escape of a Small Molecule from Inside T4 Lysozyme by Multiple Pathways. *Biophys. J.*, v. 114, p. 1058–1066, 2018.
- 9 NUNES-ALVES, A.; ARANTES, G. M. Ligand-receptor affinities computed by an adapted linear interaction model for continuum electrostatics and by protein conformational averaging. *J. Chem. Inf. and Model.*, v. 54, p. 2309–2319, 2014.
- 10 GALASSI, V. V.; ARANTES, G. M. Partition, orientation and mobility of ubiquinones in a lipid bilayer. *Biochim. Biophys. Acta - Bioenerg.*, v. 1847, p. 1560–1573, 2015.
- 11 ARANTES, G. M.; LOOS, M. Specific parametrisation of a hybrid potential to simulate reactions in phosphatases. *Phys. Chem. Chem. Phys.*, v. 8, p. 347–353, 2006.
- 12 BORDES, I.; CASTILLO, R.; MOLINER, V. Theoretical Study of the Phosphoryl Transfer Reaction from ATP to Dha Catalyzed by DhaK from Escherichia coli. *J. Phys Chem. B*, v. 121, p. 8878–8892, 2017.
- 13 DEFELIPE, L. A. et al. Protein Topology Determines Cysteine Oxidation Fate: The Case of Sulfenyl Amide Formation among Protein Families. *PLoS Comput. Biol.*, v. 11, p. 1–25, 2015.
- 14 KOGA, N. et al. Principles for designing ideal protein structures. *Nature*, v. 491, p. 222–7, 2012.
- 15 CUSTÓDIO, F. L.; BARBOSA, H. J.; DARDENNE, L. E. A multiple minima genetic algorithm for protein structure prediction. *Appl. Soft Comput.*, v. 15, p. 88–99, 2014.

- 16 THIEL, W.; HUMMER, G. Methods for computational chemistry. *Nature*, p. 96–97, 2013.
- 17 BELEVICH, N.; Von Ballmoos, C.; VERKHOVSKAYA, M. Activation of Proton Translocation by Respiratory Complex I. *Biochemistry*, v. 56, p. 5691–5697, 2017.
- 18 BUDIN, I. et al. Viscous control of cellular respiration by membrane lipid composition. *Science*, v. 1189, p. 1186–1189, 2018.
- 19 AGIP, A.-N. A. et al. Cryo-EM structures of complex I from mouse heart mitochondria in two biochemically defined states. *Nature Struct. & Mol. Biol.*, v. 25, p. 548–556, 2018.
- 20 HUANG, L. S. et al. Crystallographic studies of the binding of ligands to the dicarboxylate site of Complex II, and the identity of the ligand in the "oxaloacetate-inhibited" state. *Biochim. Biophys. Acta - Bioenerg.*, v. 1757, p. 1073–1083, 2006.
- 21 MAKLASHINA, E. et al. Fumarate reductase and succinate oxidase activity of *Escherichia coli* complex II homologs are perturbed differently by mutation of the flavin binding domain. *J. Biol. Chem.*, v. 281, p. 11357–11365, 2006.
- 22 MULLER, F. L.; LIU, Y.; Van Remmen, H. Complex III releases superoxide to both sides of the inner mitochondrial membrane. *J. Biol. Chem.*, v. 279, p. 49064–49073, 2004.
- 23 OSYCZKA, A. et al. Reversible redox energy coupling in electron transfer chains. *Nature*, v. 427, p. 607–612, 2004.
- 24 GAMIZ-HERNANDEZ, A. P. et al. Terminal Electron-Proton Transfer Dynamics in the Quinone Reduction of Respiratory Complex I. *J. Am. Chem. Soc.*, v. 139, p. 16282–16288, 2017.
- 25 BARRAGAN, A. M. et al. Identification of Ubiquinol Binding Motifs at the Q<sub>o</sub> Site of the Cytochrome bc<sub>1</sub> Complex. *J. Phys. Chem. B*, v. 112, p. 11571–11576, 2015.
- 26 KLINGEN, A. R. et al. Redox-linked protonation state changes in cytochrome bc<sub>1</sub> identified by Poisson-Boltzmann electrostatics calculations. *Biochim. Biophys. Acta - Bioenerg.*, v. 1767, p. 204–221, 2007.
- 27 WALLER, T. C. et al. Compartment and hub definitions tune metabolic networks for metabolomic interpretations. *GigaScience*, v. 9, p. 1–18, 2020.
- 28 VOET, D.; VOET, J. *Biochemistry*. 4<sup>th</sup>. ed. New York: Wiley, 2011.
- 29 BARADARAN, R. et al. Crystal structure of the entire respiratory complex I. *Nature*, v. 494, p. 443–448, 2013.
- 30 RUSTIN, P.; RÖTIG, A. *Inborn errors of complex II - Unusual human mitochondrial diseases*. 2002. 117–120 p.
- 31 BÉNIT, P.; LEBON, S.; RUSTIN, P. Respiratory-chain diseases related to complex III deficiency. *Biochim. Biophys. Acta - Mol. Cell Res.*, v. 1793, p. 181–185, 2009.
- 32 SAZANOV, L. A. A giant molecular proton pump : structure and mechanism of respiratory complex I. *Nature*, v. 16, p. 375–388, 2015.
- 33 HIRST, J. Mitochondrial Complex I. *Annu. Rev. Biochem.*, v. 82, p. 551–575, 2013.

- 
- 34 PAREY, K. et al. High-resolution cryo-EM structures of respiratory complex I: Mechanism, assembly, and disease. *Sci. Adv.*, v. 5, p. 1–10, 2019.
- 35 AGIP, A. A. et al. Mammalian Respiratory Complex I Through the Lens of Cryo-EM. *Annu. Rev. Biophys.*, v. 48, p. 165–184, 2019.
- 36 WAGNER, T.; ERMLER, U.; SHIMA, S. The methanogenic CO<sub>2</sub> reducing-and-fixing enzyme is bifunctional and contains 46 [4Fe-4S] clusters. *Science*, v. 354, p. 114–117, 2016.
- 37 PETERS, J. W. X-ray Crystal Structure of the Fe-Only Hydrogenase (CpI) from *Clostridium pasteurianum* to 1.8 Angstrom Resolution. *Science*, v. 282, p. 1853–1858, 1998.
- 38 VOLBEDA, A. et al. Crystal structure of the nickel-iron hydrogenase from *Desulfovibrio gigas*. *Nature*, v. 373, p. 580–587, 1995.
- 39 MATHIESEN, C.; HÄGERHÄLL, C. 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.*, v. 1556, p. 121–132, 2002.
- 40 SHARMA, V. et al. Redox-induced activation of the proton pump in the respiratory complex I. *Proc. Natl. Acad. Sci. U. S. A.*, v. 112, p. 11571–11576, 2015.
- 41 Di Luca, A.; GAMIZ-HERNANDEZ, A. P.; KAILA, V. R. I. Symmetry-related proton transfer pathways in respiratory complex I. *Proc. Natl. Acad. Sci. U. S. A.*, v. 114, p. E6314–E6321, 2017.
- 42 WIKSTRÖM, M. et al. New Perspectives on Proton Pumping in Cellular Respiration. *Chem. Rev.*, v. 115, p. 2196–2221, 2015.
- 43 WARNAU, J. et al. Redox-coupled quinone dynamics in the respiratory complex I. *Proc. Natl. Acad. Sci. U. S. A.*, v. 115, p. E8413–E8420, 2018.
- 44 HAAPANEN, O.; DJURABEKOVA, A.; SHARMA, V. Role of second quinone binding site in proton pumping by respiratory complex I. *Front. Chem.*, v. 7, p. 1–14, 2019.
- 45 IVERSON, T. M. Catalytic mechanisms of complex II enzymes: A structural perspective. *Biochim. Biophys. Acta - Bioenerg.*, v. 1827, p. 648–657, 2013.
- 46 CROFTS, A. R. et al. The q-cycle reviewed: How well does a monomeric mechanism of the bc<sub>1</sub> complex account for the function of a dimeric complex? *Biochim. Biophys. Acta*, v. 1777, p. 1001–1019, 2008.
- 47 CROFTS, A. R. et al. The Q-Cycle Mechanism of the bc<sub>1</sub> Complex: A Biologist's Perspective on Atomistic Studies. *J. Phys. Chem. B*, v. 121, p. 3701–3717, 2017.
- 48 CROFTS, A. R. et al. The mechanism of ubihydroquinone oxidation at the Q<sub>o</sub>-site of the cytochrome bc<sub>1</sub> complex. *Biochim. Biophys. Acta - Bioenerg.*, v. 1827, p. 1362–1377, 2013.
- 49 ESSER, L. et al. Inhibitor-complexed structures of the cytochrome bc<sub>1</sub> from the photosynthetic bacterium *Rhodobacter sphaeroides*. *J. Biol. Chem.*, v. 283, p. 2846–2857, 2008.

- 50 YOSHIKAWA, S.; MURAMOTO, K.; SHINZAWA-ITOH, K. Proton-Pumping Mechanism of Cytochrome c Oxidase . *Annu. Rev. Biophys.*, v. 40, p. 205–223, 2011.
- 51 JUNGE, W.; NELSON, N. ATP synthase. *Annu. Rev. Biochem.*, v. 84, p. 631–657, 2015.
- 52 RICH, P. R.; HARPER, R. Partition coefficients of quinones and hydroquinones and their relation to biochemical reactivity. *FEBS Lett.*, v. 269, p. 139–144, 1990.
- 53 LENAŽ, G. Quinone specificity of Complex I. *Biochim. Biophys. Acta - Bioenerg.*, v. 1364, p. 207–221, 1998.
- 54 LIU, B. L. et al. The relationship between the structure of plastoquinone derivatives and their biological activity in Photosystem II of spinach chloroplasts. *Photosynth. Res.*, v. 30, n. 2-3, p. 95–106, 1991.
- 55 FEDOR, J. G. et al. Correlating kinetic and structural data on ubiquinone binding and reduction by respiratory complex I. *Proc. Natl. Acad. Sci. U. S. A.*, v. 114, p. 12737–12742, 2017.
- 56 MARCHAL, D. et al. Electrochemical measurement of lateral diffusion coefficients of ubiquinones and plastoquinones of various isoprenoid chain lengths incorporated in model bilayers. *Biophys. J.*, v. 74, p. 1937–1948, 1998.
- 57 FATO, R. et al. 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*, v. 25, p. 3378–3390, 1986.
- 58 BARILE, C. J. et al. The Flip-Flop Diffusion Mechanism across Lipids in a Hybrid Bilayer Membrane. *Biophys. J.*, v. 110, p. 2451–2462, 2016.
- 59 ZICKERMANN, V. et al. Analysis of the pathogenic human mitochondrial mutation ND1/3460, and mutations of strictly conserved residues in its vicinity, using the bacterium *Paracoccus denitrificans*. *Biochemistry*, v. 37, p. 11792–11796, 1998.
- 60 KURKI, S. et al. Mutagenesis of three conserved Glu residues in a bacterial homologue of the ND1 subunit of complex I affects ubiquinone reduction kinetics but not inhibition by dicyclohexylcarbodiimide. *Biochemistry*, v. 39, p. 13496–13502, 2000.
- 61 FATO, R. et al. Steady-state kinetics of the reduction of coenzyme Q analogs by complex I (NADH:ubiquinone oxidoreductase) in bovine heart mitochondria and submitochondrial particles. *Biochemistry*, v. 35, p. 2705–2716, 1996.
- 62 KAUROLA, P. et al. Distribution and dynamics of quinones in the lipid bilayer mimicking the inner membrane of mitochondria. *J. Biol. Chem.*, v. 1858, p. 2116–2122, 2016.
- 63 KATSIKAS, H.; QUINN, P. J. The polyisoprenoid chain length influences the interaction of ubiquinones with phospholipid bilayers. *Biochim. Biophys. Acta - Biomembranes*, v. 689, p. 363–369, 1982.
- 64 KATSIKAS, H.; QUINN, P. J. Fluorescence Probe Studies of the Distribution of Ubiquinone Homologues in Bilayers of Dipalmitoylglycerophosphocholine. *Eur. J. Biochem.*, v. 131, p. 607–612, 1983.

- 
- 65 METZ, G. et al. NMR Studies of Ubiquinone Location in Oriented Model Membranes: Evidence for a Single Motionally-Averaged Population. *J. Am. Chem. Soc.*, v. 117, p. 564–565, 1995.
- 66 JEMIOLA-RZEMINSKA, M. et al. Location of ubiquinone homologues in liposome membranes studied by fluorescence anisotropy of diphenyl-hexatriene and trimethylammonium-diphenyl-hexatriene. *Chem. Phys. Lipids*, v. 79, p. 55–63, 1996.
- 67 HAUSS, T. et al. Localization of coenzyme Q10 in the center of a deuterated lipid membrane by neutron diffraction. *Biochim. Biophys. Acta - Bioenerg.*, v. 1710, p. 57–62, 2005.
- 68 KINGSLEY, P. B.; FEIGENSON, G. W. <sup>1</sup>H-NMR study of the location and motion of ubiquinones in perdeuterated phosphatidylcholine bilayers. *Biochim. Biophys. Acta - Bioenerg.*, v. 635, p. 602–618, 1981.
- 69 STIDHAM, M. A.; MCINTOSH, T. J.; SIEDOW, J. N. The location of ubiquinone-10 in phospholipid bilayers was analyzed using a variety of physical techniques . Specifically , we examined the hypothesis that ubiquinone localizes at the geometric center of phospholipid bilayers . Light microscopy of dipalmitoyl. *Biochim. Biophys. Acta*, v. 767, p. 423–431, 1984.
- 70 CORNELL, B. A. et al. Location and Activity of Ubiquinone 10 and Ubiquinone Analogues in Model and Biological Membranes. *Biochemistry*, v. 26, p. 7702–7707, 1987.
- 71 LENAŽ, G. et al. Localization and preferred orientations of ubiquinone homologs in model bilayers. *Biochem. Cell Biol.*, v. 70, p. 504–514, 1992.
- 72 SCHEURING, S.; STURGIS, J. N. Dynamics and diffusion in photosynthetic membranes from *Rhodospirillum rubrum*. *Biophys. J.*, v. 91, p. 3707–3717, 2006.
- 73 CHAZOTTE, B.; HACKENBROCK, C. R. Lateral diffusion as a rate-limiting step in ubiquinone-mediated mitochondrial electron transport. *J. Biol. Chem.*, v. 264, p. 4978–4985, 1989.
- 74 MATHAI, J. C. et al. Rate-limiting step in electron transport. Osmotically sensitive diffusion of quinones through voids in the bilayer. *J. Biol. Chem.*, v. 268, p. 15442–15454, 1993.
- 75 JOHNSON, D. C. et al. Structure, Function, and Formation of Biological Iron-Sulfur Clusters. *Annu. Rev. Biochem.*, v. 74, p. 247–281, 2005.
- 76 BRZÓSKA, K.; MĘCZYŃSKA, S.; KRUSZEWSKI, M. Iron-sulfur cluster proteins: Electron transfer and beyond. *Acta Biochim. Pol.*, v. 53, p. 685–691, 2006.
- 77 NATH, K.; WESSENDORF, R. L.; LU, Y. A nitrogen-fixing subunit essential for accumulating 4Fe-4S-containing photosystem I core proteins. *Plant Physiol.*, v. 172, p. 2459–2470, 2016.
- 78 NELSON, D. L.; COX, M. M. *Lehninger Principles of Biochemistry*. 6<sup>th</sup>. ed. New York: W. H. Freeman, 2014.
- 79 FLINT, D. H.; ALLEN, R. M. Iron-sulfur proteins with nonredox functions. *Chem. Rev.*, v. 96, p. 2315–2334, 1996.

- 80 CHAN, M. K.; KIM, J.; REES, D. C. The nitrogenase FeMo-cofactor and P-cluster pair: 2.2 Å resolution structures. *Science*, v. 260, p. 792–794, 1993.
- 81 PETERS, J. W. et al. Redox-dependent structural changes in the nitrogenase P-cluster. *Biochemistry*, v. 36, p. 1181–1187, 1997.
- 82 BEINERT, H.; HOLM, R. H.; MÜNCK, E. Iron-sulfur clusters: Nature’s modular, multipurpose structures. *Science*, v. 277, p. 653–659, 1997.
- 83 MARTIC, M. et al. Heterometallic [AgFe<sub>3</sub>S<sub>4</sub>] ferredoxin variants: Synthesis, characterization, and the first crystal structure of an engineered heterometallic iron-sulfur protein. *J. Biol. Inorg. Chem.*, v. 18, p. 261–276, 2013.
- 84 LIU, J. et al. Metalloproteins containing cytochrome, iron-sulfur, or copper redox centers. *Chem. Rev.*, v. 114, p. 4366–4369, 2014.
- 85 BAK, D. W.; ELLIOTT, S. J. Alternative fes cluster ligands: Tuning redox potentials and chemistry. *Curr. Opin. Chem. Biol.*, v. 19, p. 50–58, 2014.
- 86 ZURIS, J. A. et al. Engineering the redox potential over a wide range within a new class of FES proteins. *J. Am. Chem. Soc.*, v. 132, p. 13120–13122, 2010.
- 87 HOSSEINZADEH, P. et al. Design of a single protein that spans the entire 2-V range of physiological redox potentials. *Proc. Natl. Acad. Sci. U. S. A.*, v. 113, p. 262–267, 2016.
- 88 LEAL, S.; GOMES, C. Iron-sulfur clusters, protein folds, and ferredoxin stability. In: GOMES, C. M.; WITTUNG-STAFSHEDE, P. (Ed.). *Protein Folding and Metal Ions: Mechanisms, Biology and Disease*. New York: CRC Press, 2016. cap. 5, p. 81–96.
- 89 ZHENG, P. et al. Force-induced chemical reactions on the metal centre in a single metalloprotein molecule. *Nat. Commun.*, v. 6, p. 1–9, 2015.
- 90 FROMME, J. C.; VERDINE, G. L. Structure of a trapped endonuclease III-DNA covalent intermediate. *EMBO J.*, v. 22, p. 3461–3471, 2003.
- 91 CRACK, J. C. et al. Signal perception by FNR: The role of the iron-sulfur cluster. *Biochem. Soc. Trans.*, v. 36, p. 1144–1148, 2008.
- 92 LIPINSKI, P. et al. Induction of iron regulatory protein 1 RNA-binding activity by nitric oxide is associated with a concomitant increase in the labile iron pool: Implications for DNA damage. *Biochem. Biophys. Res. Commun.*, v. 327, p. 349–355, 2005.
- 93 WINSBERG, E. Computer simulations in science. In: ZALTA, E. N. (Ed.). *The Stanford Encyclopedia of Philosophy*. Winter 2019. [S.l.]: Metaphysics Research Lab, Stanford University, 2019.
- 94 FIELD, M. J. *A Practical Introduction to the Simulation of Molecular Systems*. 2<sup>nd</sup>. ed. Cambridge: Cambridge University Press, 2007.
- 95 ZUCKERMAN, D. M. *Statistical Physics of Biomolecules: An Introduction*. 1<sup>st</sup>. ed. New York: CRC Press, 2010.
- 96 HUANG, J.; JR, A. D. M. Charmm36 all-atom additive protein force field: Validation based on comparison to nmr data. *J. Comp. Chem.*, v. 34, p. 2135–2145, 2013.

- 
- 97 THOMAS, A. et al. Is there a covalent intermediate in the viral neuraminidase reaction? A hybrid potential free-energy study. *J. Am. Chem. Soc.*, v. 121, p. 9693–9702, 1999.
- 98 FEYNMAN, R. Forces in molecules. *Phys. Rev.*, v. 56, p. 340–343, 1939.
- 99 SZABO, A.; OSTLUND, N. *Modern Quantum Chemistry*. 1<sup>st</sup>. ed. New York: Dover, 1989.
- 100 CREMER, D. Møller-Plesset perturbation theory: From small molecule methods to methods for thousands of atoms. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, v. 1, p. 509–530, 2011.
- 101 BARTLETT, R. J.; MUSIAŁ, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.*, v. 79, p. 291–352, 2007.
- 102 JONES, R. O. Density functional theory: Its origins, rise to prominence, and future. *Rev. Mod. Phys.*, v. 87, 2015.
- 103 FIELD, M. J. Simulating enzyme reactions: Challenges and perspectives. *J. Comp. Chem.*, v. 23, p. 48–58, 2002.
- 104 LIN, H.; TRUHLAR, D. G. QM/MM: What have we learned, where are we, and where do we go from here? *Theor. Chem. Acc.*, v. 117, p. 185–199, 2007.
- 105 ROUX, B. The calculation of the potential of mean force using computer simulations. *Comp. Phys. Comm.*, v. 91, p. 275–282, 1995.
- 106 KUMAR, S. et al. The weighted histogram analysis method for free-energy calculations on biomolecules. i. the method. *J. Comput. Chem.*, v. 13, p. 1011–1021, 1992.
- 107 GROSSFIELD, A. *WHAM: the weighted histogram analysis method*. <<http://membrane.urmc.rochester.edu/content/wham>>.
- 108 MOORE, C. C. Ergodic theorem, ergodic theory, and statistical mechanics. *Proc. Natl. Acad. Sci. U. S. A.*, v. 112, p. 1907–1911, 2015.
- 109 FIORIN, G.; KLEIN, M. L.; HÉNIN, J. Using collective variables to drive molecular dynamics simulations. *Mol. Phys.*, v. 111, p. 3345–3362, 2013.
- 110 BRANDUARDI, D.; GERVASIO, F. L.; PARRINELLO, M. From A to B in free energy space. *J. Chem. Phys.*, v. 126, p. 054103, 2007.