

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

Felipe Curtolo

**Reaction Mechanisms of Flavins and Flavoproteins from
an Electronic-Structure Perspective**

Versão original da tese defendida

São Paulo
Data de Depósito na SPG
07/10/2022

Felipe Curtolo

Reaction Mechanisms of Flavins and Flavoproteins from an Electronic-Structure Perspective

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

2022

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

C978r Curtolo, Felipe
 Reaction Mechanisms of Flavins and Flavoproteins
 from an Electronic-Structure Perspective / Felipe
 Curtolo. - São Paulo, 2022.
 133 p.

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

Orientador: Arantes, Guilherme Menegon

1. flavoprotein. 2. flavin. 3. fumarate
reductase. 4. reaction mechanism. 5. electronic-
structure calculation. I. T. II. Arantes, Guilherme
Menegon, orientador.

To my father, Antonio

Acknowledgements

I am deeply grateful to my supervisor Prof. Guilherme Menegon Arantes. During these years, his ideas and suggestions have taught me not only how to do good science, but also how to communicate and write better. Under his mentorship, I developed the skills that have made the professional I am today.

I would also like to thank my colleagues, Ariane Nunes-Alves, Murilo H. Teixeira, André A. O. Reis and Sofia R. G. Camilo. Since the beginning, they have provided me a friendly and intellectually stimulating environment. I am especially grateful to Ariane and Murilo for all our scientific discussions and for the numerous times they have helped me.

I owe my partner Vitor a huge debt of gratitude for having supported me tirelessly over these years. He was responsible for keeping me on track during some very difficult times, always giving me valuable advice. I also need to thank him for all the times he watched my presentations and gave me tips. His knowledge of public speaking and English conversation has greatly helped me.

I must also express my thanks to my mother Denise, my sister Lauren and my aunt Vera for supporting me since my beginning in São Carlos. I am also grateful to Paula, my mother-in-law, for welcoming me into her family and making me feel at home in São Paulo.

I am also grateful to the friends I have made here, especially Bea, my roommate during quarantine. Her friendship was fundamental during the almost two years of social isolation and remote work we spent together. All the laughs and moments shared made life easier during that time.

I also want to acknowledge the Graduate Program and the Department of Biochemistry for all the lectures and seminars I attended during my PhD. I am equally grateful to IQ-USP for providing the amazing infrastructure and staff.

Finally, I acknowledge São Paulo Research Foundation (FAPESP) for funding the project that originated this thesis (grants 2017/26109-0 and 2016/23525-0) and for allowing me to participate in various conferences, schools and workshops in Brazil and abroad.

“Life is nothing but an electron looking for a place to rest”

Albert Szent-Györgyi

Abstract

Curtolo, F. **Reaction Mechanisms of Flavins and Flavoproteins from an Electronic-Structure Perspective** 2022. 133p. PhD thesis - Graduate Program in Biochemistry. Instituto de Química, Universidade de São Paulo, São Paulo.

Proteins equipped with flavin adenine dinucleotides (FAD) or flavin mononucleotides (FMN) are named flavoproteins and constitute about 1% of all existing proteins. They catalyze redox, acid-base and photochemical reactions in a variety of biochemical phenomena that goes from energy metabolism to DNA repair and light sensing. The versatility observed in flavoproteins is ultimately a balance of flavin intrinsic properties modulated by a protein environment. This thesis aims to investigate how flavoproteins work by systematic evaluating flavin properties and reactivity. In particular, the mechanism of fumarate reduction by the flavoenzyme fumarate reductase Fcc₃ was determined. Electronic-structure calculations were used for this task based on rigorous calibration with experimental data and error assessment. Flavin properties at chemical accuracy were obtained with single reference coupled-cluster CCSD(T) calculations at the complete basis set limit. Density functional theory was demonstrated an excellent alternative with lower computational costs and slightly less accuracy. Flavin protonation and tautomerism were shown to be important modulators of flavin properties and reactivity, with the possibility of various tautomers existing at neutral pH. Regarding flavin redox properties, an analysis based on multiconfigurational wave function weights was proposed for categorizing flavin redox reactions as hydride or hydrogen-atom transfers. This analysis is an upgrade over traditional partial charges methods and can be applied not only to flavin reactions but to any proton-coupled electron transfer. In the investigation of the enzymatic mechanism of fumarate reduction, the reaction was determined as a nucleophilic addition by hydride transfer with carbanion formation. Fumarate reductase employs electrostatic catalysis in contrast to previous proposals of substrate straining and general-acid catalysis. Also, hydride transfer was shown to be vibronically adiabatic with low tunneling contribution. These findings give new insights into the mechanisms of fumarate reductases and provide a framework

for future computational studies of flavoproteins in general. The analyses and benchmark studies presented can be used to build better models of properties and reactivity of flavins and flavoproteins.

Keywords: flavoprotein, flavin, fumarate reductase, reaction mechanism, electronic-structure calculation, benchmark study

Resumo

Curtolo, F. **Mecanismos de Reação de Flavinas e Flavoproteínas por uma Perspectiva de Estrutura Eletrônica** 2022. 133p. Tese - Programa de Pós-Graduação em Bioquímica. Instituto de Química, Universidade de São Paulo, São Paulo.

Proteínas equipadas com dinucleotídeos de flavina-adenina (FAD) e mononucleotídeos de flavina (FMN) são chamadas flavoproteínas e constituem cerca de 1% de todas as proteínas existentes. Elas catalisam reações redox, ácido-base e fotoquímicas numa variedade de fenômenos bioquímicos que vão desde o metabolismo energético até reparo de DNA e captação de luz. A versatilidade observada em flavoproteínas é em última instância um balanço das propriedades intrínsecas de flavinas moduladas por um ambiente proteico. Esta tese busca investigar como flavoproteínas funcionam através de avaliações sistemáticas de propriedades e reatividade de flavinas. Em particular, o mecanismo de redução de fumarato pela flavoenzima fumarato redutase Fcc₃ foi determinado. Cálculos de estrutura eletrônica foram usados para esta tarefa com base em rigorosa calibração com dados experimentais e avaliação de erros. As propriedades de flavinas foram determinadas com acurácia química com cálculos monoconfiguracionais de *coupled-cluster* CCSD(T) no limite de conjunto base completo. A teoria do funcional da densidade mostrou-se uma alternativa excelente com menor custo computacional e um pouco menos de acurácia. Protonação e tautomerismo de flavinas mostraram-se moduladores importantes de suas propriedades e reatividade, com a possibilidade de vários tautômeros existirem em pH neutro. Em relação às propriedades redox de flavinas, uma análise baseada nos pesos de funções de onda multiconfiguracionais foi proposta para categorizar as reações redox de flavinas como transferências de hidreto ou hidrogênio. Esta análise é uma melhoria em relação aos métodos tradicionais de cargas parciais e pode ser aplicada não apenas para reações de flavinas mas para qualquer transferência de próton acoplada a elétrons. Na investigação do mecanismo enzimático de redução de fumarato, a reação foi designada como uma adição nucleofílica por transferência de hidreto e formação de carbânion. A fumarato redutase

usa catálise eletrostática diferentemente de propostas anteriores envolvendo distorção do substrato e catálise ácida geral. Além disso, a transferência de hidreto mostrou-se vibronicamente adiabática com pouca contribuição de tunelamento. Estas descobertas abrem novas perspectivas sobre os mecanismos de fumarato redutases e fornecem uma base para estudos computacionais futuros sobre flavoproteínas em geral. As análises e estudos comparativos apresentados podem ser usados para construir melhores modelos para propriedades e reatividade de flavinas e flavoproteínas.

Palavras-chave: flavoproteína, flavina, fumarato redutase, mecanismo de reação, cálculos de estrutura eletrônica, estudo comparativo

List of abbreviations and acronyms

CASSCF	Complete Active-space Self-consistent Field
CCSD	Coupled-cluster Singles and Doubles
CCSD(T)	Coupled-cluster Singles and Doubles with Perturbative Triples
CI	Configuration Interaction
DFT	Density Functional Theory
DLPNO	Domain-based Local-pair Natural Orbital
FAD	Flavin Adenine Dinucleotide
Fcc ₃	Flavocytochrome <i>c</i> ₃
FCI	Full-configuration Interaction
FMN	Flavin Mononucleotide
GGA	Generalized Gradient Approximation
LCAO	Linear Combination of Atomic Orbitals
MCSCF	Multiconfigurational Self-consistent Field
MP2	Second-order Møller–Plesset Perturbation Theory
NEVPT2	<i>n</i> -electron Valence State Perturbation Theory at Second-order
PES	Potential Energy Surface
RHF	Restricted Hartree–Fock
RRHO	Rigid-Rotor-Harmonic-Oscillator
SCF	Self-consistent Field
UHF	Unrestricted Hartree–Fock
ZPVE	Zero-point Vibrational Energy

Contents

1	INTRODUCTION	19
1.1	Flavin coenzymes and flavoproteins	20
1.1.1	Flavin chemical properties	21
1.1.2	Redox flavoproteins	23
1.1.3	Flavocytochrome c_3 – A soluble fumarate reductase	26
1.2	Electronic-structure calculations	32
1.2.1	Electronic Hamiltonian	32
1.2.2	Hartree–Fock theory	34
1.2.3	Electron correlation	38
1.2.4	Multiconfigurational quantum chemistry	38
1.2.5	Perturbation theory	40
1.2.6	Coupled-cluster theory	43
1.2.7	Density functional theory	46
1.2.8	Connection to thermodynamics	48
1.3	Objectives	52
2	MOLECULAR PROPERTIES AND TAUTOMERIC EQUILIBRIA OF ISOLATED FLAVINS	55
3	MECHANISMS FOR FLAVIN-MEDIATED OXIDATION: HYDRIDE OR HYDROGEN-ATOM TRANSFER?	83
4	MECHANISM OF FLAVIN OXIDATION CATALYZED BY FUMARATE REDUCTASES	93
5	CONCLUSIONS	115
	BIBLIOGRAPHY	119
	APPENDIX	129
	APPENDIX A – CASSCF CALCULATIONS SETUP	131
	APPENDIX B – ADDITIONAL PRODUCTION	133

1 Introduction

Flavins are extremely versatile coenzymes. Since their first isolation and characterization in the 19th century,¹ their catalytic roles have been surprising to scientists in various fields. As a catalyst, flavin is not only a biologically important coenzyme, but also a promising organic catalyst for synthetic and photochemical applications.^{2,3}

Due to its versatility, understanding the molecular properties of flavins and the mechanisms of flavoenzymes at molecular level remains a challenge. Computational chemistry methods, especially highly accurate quantum chemistry calculations were used in this thesis to respond some open questions regarding general chemistry of flavins and the mechanism of the flavoenzyme flavocytochrome c_3 (Fcc₃).⁴

This is an introduction chapter designed to contextualize the reader to these two subjects: flavins and computational chemistry. Section 1.1 provides an overview of the properties and reactivity of flavins and of Fcc₃ enzyme. Section 1.2 describes the computational methods used here and section 1.3 finishes the chapter presenting the specific objectives of the thesis.

The other chapters are organized in the form of articles that address the objectives described in section 1.3. Chapter 2 is a research article published in the Journal of Computational Chemistry that presents a benchmark study of flavin properties. The accuracy of various quantum chemistry approximations are discussed and the importance of different electronic and thermodynamic effects are assessed. Chapter 3 was published in the Journal of Chemical Information and Modeling, where a new analysis to quantify the number of electrons transferred in flavin proton-coupled electron transfers is presented. Finally, in chapter 4, the mechanism of Fcc₃ is determined. The contribution of different effects such as proton tunneling, substrate protonation, involvement of excited states and long-range electrostatic interactions are assessed. In the end, a general conclusion is presented in chapter 5.

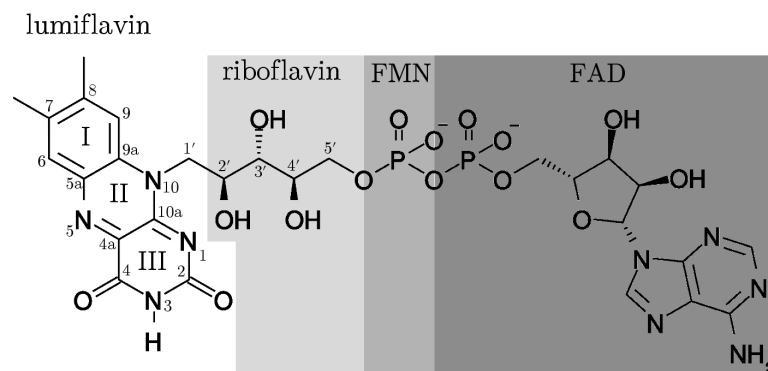


Figure 1 – Structure and atom numbering of lumiflavin, riboflavin, FMN and FAD. The isoalloxazine moiety is formed by rings I (xylene), II (pyrazine) and III (pyrimidine).

1.1 Flavin coenzymes and flavoproteins

The name flavin is derived from the Latin *flavus*, meaning yellow,⁵ the color of crystals obtained from purified flavins.⁶ The chemical structure of this class of compounds contains the tricyclic ring isoalloxazine,^{7,8} which consists in fused xylene, pyrazine and pyrimidine rings (Fig. 1). The isoalloxazine ring provides very unique properties to flavins, allowing them to be used by proteins in a variety of roles, ranging from redox coenzymes⁹ to light-absorbing pigments.⁸

Mammals are unable to biosynthesize isoalloxazine,¹⁰ so they obtain their flavins by dietary sources and by absorbing flavins produced by gut microbiota.¹¹ In the free form, flavins are found mainly as riboflavins (also named vitamin B₂), while the catalytic forms flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD), which comprise the majority of flavins, are found within the cells as prosthetic groups tightly or covalently bound to proteins.^{10,12,13}

The only difference between flavins is the group bound to nitrogen 10: in riboflavin it is a ribityl, in FMN a ribityl-phosphate and in FAD an adenine-diphosphate-ribityl moiety (Fig. 1). The part responsible for their catalytic and optical properties is the isoalloxazine ring, where groups bound to nitrogen 10 have predominantly a structural role.^{6–8} Flavoproteins, i.e. proteins that use flavin as cofactors, bind to FAD or FMN specifically by recognizing their pyrophosphate or phosphate groups.¹⁴ Therefore, riboflavin

cannot bind in flavoproteins and is considered simply a precursor to the biosynthesis of FAD and FMN.¹⁰

In cells, the interconversion of riboflavin to FMN and FAD is catalyzed by riboflavin kinase^{15,16} and FAD synthase.¹⁷ The first uses adenosine triphosphate (ATP) to phosphorylate riboflavin at position 5'-hydroxyl to generate FMN, and the second uses ATP to adenylate FMN to FAD, forming a pyrophosphate bond. These reactions are coordinated to maintain flavin homeostasis.¹⁸

Lumiflavin (Fig. 1), which is a product of photodegradation of riboflavin,⁷ has a methyl at position 10 and is an important nonbiological flavin, usually employed as a model system to study flavin reactivity and photochemistry, both *in vitro*^{19–22} and *in silico*.^{23–27}

1.1.1 Flavin chemical properties

The different groups bound to nitrogen 10 of the flavins are generally not reactive. For FAD and FMN, the hydrogen-bond networks formed between these groups and the flavoprotein, provides a large enthalpic contribution to the binding free energy of the flavin coenzyme,^{28,29} which is the main reason they are kept tightly associated with their apoprotein.

The chemical and optical properties of flavins result from a balance of isoalloxazine individual-ring properties and the conjugation between electrons throughout its constituent rings. The xylene ring is hydrophobic and virtually unreactive,⁶ whereas the pyrimidine ring is hydrophilic and participate in various hydrogen-bonds that can polarize flavin in different ways.^{7,20,30} Pyrazine, is strongly influenced by the adjacent pyrimidine ring that has an electron-withdrawing inductive effect that makes nitrogen 5 and carbon 4a susceptible to nucleophilic attacks.^{6,7,9}

Electron conjugation between isoalloxazine rings, allow flavin to exist in three stable redox states: the oxidized state (flavoquinone), the one-electron reduced state (flavosemiquinone), and the two-electron reduced state (flavoquinol).^{6,31} This property is

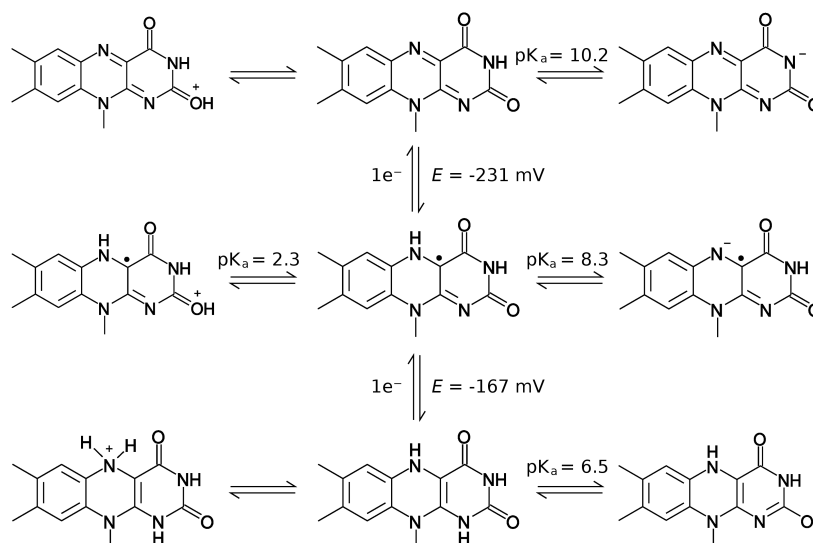


Figure 2 – Scheme of the different redox and charge states of flavins. Each row is a different redox state: flavoquinone (first row), flavosemiquinone (second row) and flavoquinol (third row). The columns represent different charge states: cationic (first column), neutral (second column) and anionic (third column). pK_a and E are experimental acid dissociation constants and reduction potentials measured in water.^{6,8}

the basis for FAD and FMN biological function as redox coenzymes, because it allows them to participate in both one- and two-electron transfer reactions, differently from other redox coenzymes that can only participate in two-electron redox processes, like nicotinamide coenzymes $NAD(P)H$,³¹ or strict one-electron redox processes, like metallic cofactors.^{32,33}

Another important property of flavins is their mild acid dissociation constants (pK_a) that allow them to exist not only in different redox states, but also in different protonation states (Fig. 2). Additionally, isoalloxazine has a high polarizability and the equilibria in Fig. 2 can be easily shifted inside proteins, allowing flavins to exist in different protonation states.^{7,34}

Tautomerism adds another layer of complexity to flavin chemistry. Atoms N1, N3, N5, O2 and O4 have lone pairs of electrons that can receive hydrogen-bonds or bind protons, giving alternative tautomeric forms to flavins than those shown in Fig. 2. The pyrimidine ring of isoalloxazine (Fig. 1), can assume different tautomeric forms in the same way as pyrimidines from nucleotides.³⁵ For example, the neutral flavoquinol presents

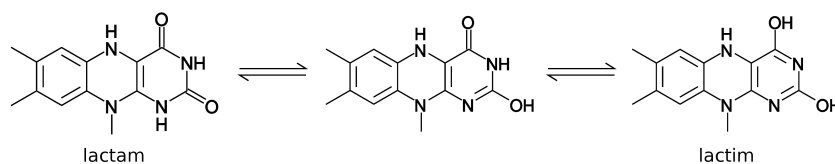


Figure 3 – Lactam-lactim tautomerism of neutral flavoquinol.

lactam-lactim tautomerism (Fig. 3), changing from a cyclic amide to a cyclic imidic acid that possess distinct properties.

The existence of flavins in different tautomeric, redox and charge states is ultimately a result of their intrinsic molecular properties modulated by an environment that can shift the equilibria between states by polarizing flavin differently. In chapter 2 of this thesis we investigate the intrinsic properties of lumiflavin without the environment. This way, we were able to map the contributions of different electronic effects to the observable properties of flavins. Additionally, we performed an extensive benchmark study with nearly all possible tautomeric, redox and charge states of lumiflavin to estimate the accuracy of different computational chemistry methods (Section 1.2) in predicting flavin properties.

Finally, despite the ground-state properties of flavins discussed so far, it is noteworthy to mention flavins can also absorb blue light (300 to 500 nm) to generate stable electronic excited states,⁸ whose properties differ significantly from ground state flavins. Excited-state flavins are used by enzymes to catalyze otherwise difficult reactions, like the pyrimidine dimer repair in DNA by DNA photolyase.³⁶

1.1.2 Redox flavoproteins

Redox flavoproteins, also known as flavoenzymes, catalyze some of the most remarkable reactions in biochemistry, such as the desaturation of aliphatic carbon chains,³⁷ the epoxidation of carbon-carbon double bonds³⁸ and the hydroxylation of aromatic compounds.³⁹ This versatility is a result of the chemical properties of the flavin cofactor, that inside flavoenzymes, might assume reduction potentials ranging from -400 mV to $+60$ mV⁷!

The catalytic cycle of any flavoenzyme can be thought as two separate redox

reactions.⁹ In the first, the reductive reaction, the flavin coenzyme is reduced by a substrate or by an electron donor protein that becomes oxidized. Whereas in the second, the oxidative reaction, the reduced flavin is reoxidized by another substrate or protein that becomes reduced. Both redox reactions are necessary to regenerate the original state of the enzyme and complete enzyme turnover.

Because flavins can exist in three different redox states (Fig. 2), flavoenzymes are classified based on the number of electrons transferred in each of its redox reaction.⁴⁰ For example, if m is the number of electrons transferred to the flavin in the reductive reaction and n the number of electrons transferred from the flavin in the oxidative reaction, the flavoenzymes can be classified by the index m/n in four possible categories: 1/1, 2/2, $(1+1)/2$ or $2/(1+1)$. Chart 1 illustrates these categories schematizing the reductive and oxidative reactions catalyzed by flavin.

Chart 1 – Types of flavoenzyme catalyzed reactions

Reductive reaction	Oxidative reaction	Overall reaction
<i>Class 2/2 – Dehydrogenase/Oxidase</i>		
$\text{FAD} + \text{AH}_2 \rightarrow \text{FADH}_2 + \text{A}$	$\text{FADH}_2 + \text{O}_2 \rightarrow \text{FAD} + \text{H}_2\text{O}_2$	$\text{AH}_2 + \text{O}_2 \rightarrow \text{A} + \text{H}_2\text{O}_2$
<i>Class 2/2 – Dehydrogenase/Monooxygenase</i>		
$\text{FAD} + \text{AH}_2 \rightarrow \text{FADH}_2 + \text{A}$	$\text{FADH}_2 + \text{B} + \text{O}_2 \rightarrow \text{FAD} + \text{BO} + \text{H}_2\text{O}$	$\text{AH}_2 + \text{B} + \text{O}_2 \rightarrow \text{A} + \text{BO} + \text{H}_2\text{O}$
<i>Class 2/2 – Transhydrogenase</i>		
$\text{FAD} + \text{AH}_2 \rightarrow \text{FADH}_2 + \text{A}$	$\text{FADH}_2 + \text{B} \rightarrow \text{FAD} + \text{BH}_2$	$\text{AH}_2 + \text{B} \rightarrow \text{A} + \text{BH}_2$
<i>Class 1/1 – Electron transferase</i>		
$\text{FAD} + \text{A}^- + \text{H}^+ \rightarrow \text{FADH}^\bullet + \text{A}$	$\text{FADH}^\bullet + \text{B} \rightarrow \text{FAD} + \text{B}^- + \text{H}^+$	$\text{A}^- + \text{B} \rightarrow \text{A} + \text{B}^-$
<i>Class 2/(1+1) – Dehydrogenase/Electron transferase</i>		
$\text{FAD} + \text{AH}_2 \rightarrow \text{FADH}_2 + \text{A}$	$\text{FADH}_2 + \text{B} \rightarrow \text{FADH}^\bullet + \text{B}^- + \text{H}^+$ $\text{FADH}^\bullet + \text{B} \rightarrow \text{FAD} + \text{B}^- + \text{H}^+$	$\text{AH}_2 + 2\text{B} \rightarrow \text{A} + 2\text{B}^- + 2\text{H}^+$
<i>Class (1+1)/2 – Electron transferase/Reductase</i>		
$\text{FAD} + \text{B}^- + \text{H}^+ \rightarrow \text{FADH}^\bullet + \text{B}$ $\text{FADH}^\bullet + \text{B}^- + \text{H}^+ \rightarrow \text{FADH}_2 + \text{B}$	$\text{FADH}_2 + \text{A} \rightarrow \text{FAD} + \text{AH}_2$	$\text{A} + 2\text{B}^- + 2\text{H}^+ \rightarrow \text{AH}_2 + 2\text{B}$

Source: Own authorship.

Note: Some flavoenzymes use FMN instead of FAD. A and B in the chart represent general substrates. For simplification, the reactions here are shown for neutral flavins FAD, FADH^\bullet and FADH_2 , but charged flavin species are equally possible.

Flavoenzymes 2/2 are involved strictly in two-electron transfer reactions, with the flavin cofactor transitioning from the fully oxidized flavoquinone to the fully reduced flavoquinol. In this class, the reductive reaction is always a dehydrogenation, with two electrons and two protons being removed from the substrate to reduce flavoquinone to

flavoquinol⁴⁰ (FAD to FADH₂ in Chart 1). In the oxidative reaction, flavoenzymes 2/2 might assume different roles, being categorized as oxidases, monooxygenases or transhydrogenases. The oxidases regenerate flavoquinone by reducing molecular oxygen O₂ to hydrogen peroxide H₂O₂, that is readily consumed by catalase to avoid oxidative damage.⁴¹ Monooxygenases also reduces O₂, but forms water instead of H₂O₂, with the second oxygen used to oxygenate a different substrate.⁴² Finally, transhydrogenases perform in the oxidative reaction a second dehydrogenation, where electrons and protons from the flavoquinol are transferred to an oxidized substrate.⁷ Examples of flavoenzymes of this class are D-amino acid oxidases,⁴³ bacterial luciferase (a monooxygenase)⁴⁴ and glutathione reductase (a transhydrogenase).⁴⁵

Flavoenzymes 1/1 are electron transferases, with both reductive and oxidative reactions consisting in pure electron transfers. Differently from dehydrogenases, they do not couple the electron transfer to a proton abstraction from substrate. Representatives of this class of flavoenzymes are flavodoxins⁴⁶ and the mitochondrial electron-transferring flavoprotein (ETF) involved in β -oxidation of fatty acids.⁴⁷

Flavoenzymes 2/1 couple two-electron transfer reactions to one-electron transfer reactions in the reductive and oxidative parts respectively. They are commonly found in large complexes associated with cytochromes and iron-sulfur proteins, like in respiratory complexes I and II.^{48,49} They always act as dehydrogenases in the reductive reaction and as electron transferases in the oxidative reaction. Two sequential steps of single-electron transfers take place in the oxidative reaction, where first flavoquinol is converted to flavosemiquinone and then to flavoquinone (FADH₂ \rightarrow FADH \cdot \rightarrow FAD, in Chart 1). Examples of 2/1 flavoenzymes are NADH dehydrogenase,⁴⁸ succinate dehydrogenase⁴⁹ and acyl-CoA dehydrogenase.⁵⁰

Flavoenzymes 1/2 are found mainly in anaerobic organisms, working in the reverse direction of 2/1 flavoenzymes. In the reductive reaction, they receive two electrons sequentially by single-electron transfers and in the oxidative reaction these electrons are transferred directly to an oxidized substrate. An example of this class is flavocytochrome c₃,⁴ that receives electrons from heme cofactors in the reductive reaction and reduces

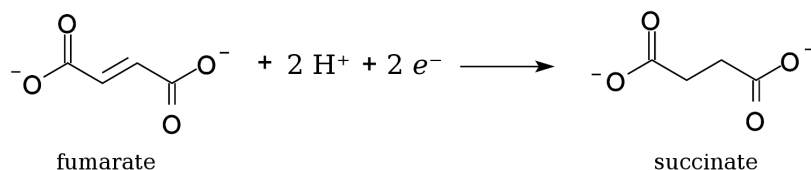


Figure 4 – Reaction of fumarate reduction to succinate.

fumarate to succinate in the oxidative reaction. This enzyme is discussed in detail in section 1.1.3 and chapter 4.

More than 90% of all flavoproteins catalyze redox reactions⁵¹ as dehydrogenases, monooxygenases, oxidases or electron transferases. The remaining nonredox flavoproteins are involved in processes like circadian clock regulation;⁵² control of plant development induced by light;⁵³ DNA damage repair employing nonoxidative bond breaks;³⁶ and even acid-base catalysis in some isomerizations,^{54,55} to name a few.

In redox flavoproteins, knowing the number of electrons transferred between the substrate and the flavin coenzyme is essential to characterize them mechanistically, so we proposed in chapter 3 a new analysis based on multiconfigurational quantum chemistry calculations (section 1.2.4) to quantify the number of electrons transferred in dehydrogenation reactions. In particular, we used NADH and succinate oxidations as model systems to test our proposed analysis and answer some disputing mechanistic views.

1.1.3 Flavocytochrome c_3 – A soluble fumarate reductase

In anaerobic respiration, ATP is synthesized from ADP and inorganic phosphate in a process that resembles that of oxidative phosphorylation,³¹ except that the terminal electron acceptor is not O_2 . Among the many possible molecules used in anaerobic respiration as terminal electron acceptor, the most common is fumarate,⁵⁶ which is reduced to succinate by the reaction shown in Fig. 4.

Most organisms that perform fumarate respiration uses the membrane-bound complex quinol:fumarate reductase to catalyze fumarate reduction.^{57–59} Some exceptions are the Gram-negative bacteria of genus *Shewanella*,⁴ trypanosomatids⁶⁰ and *S. cerevisiae*⁶¹

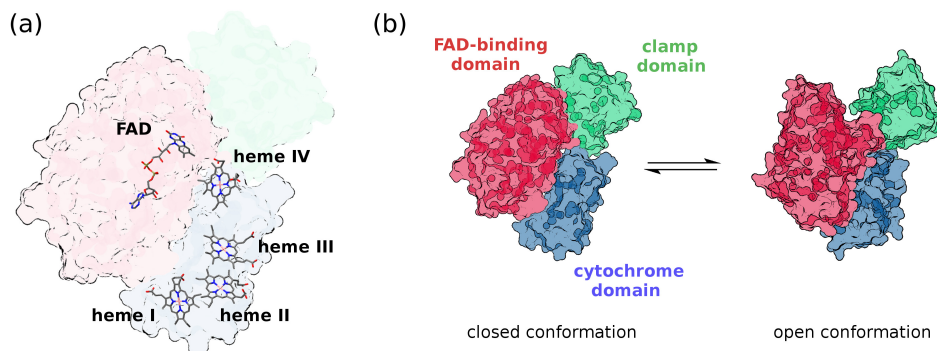


Figure 5 – Structural model of flavocytochrome c_3 (Fcc₃) of *Shewanella frigidimarina*. (a) Position of the five prosthetic groups of Fcc₃ forming the “electron wire” that brings electrons to FAD in the reductive reaction. Hemes I to IV are type-*c* hemes and FAD is not covalently bound to the protein. (b) The closed and open conformation of Fcc₃ determined by crystallography.^{4,62} The cytochrome domain, FAD-binding domain and clamp-domain are represented in blue, red and green, respectively.

that employ soluble fumarate reductases homologs. In common, all fumarate reductases are flavoenzymes that use FADH^- to catalyze the reaction of Fig. 4 in their oxidative cycle. The biggest difference among them is the reductive reaction that might employ electron transfers from iron-sulfur clusters,^{57–59} heme cofactors⁴ or other flavin coenzymes.⁶¹

In *Shewanella* sp., the soluble fumarate reductase is a periplasmic cytochrome named flavocytochrome c_3 (Fcc₃) composed of a single polypeptide chain organized in three domains⁴ (Fig. 5). The N-terminal domain is a cytochrome domain bound to four type-*c* hemes, named from I to IV, that are axially coordinated by histidine residues. The FAD-binding domain is the largest and binds FAD non-covalently. The clamp domain forms a “hinge” with the FAD-binding domain and is able to transition between a closed⁴ and open conformation⁶² (Fig. 5b). The active site of fumarate reduction is localized in the cleft between the clamp and FAD-binding domain, suggesting a large domain movement must be associated with substrate binding.^{4,62}

In the reductive reaction, Fcc₃ reduces FAD to anionic FADH^- ⁶³ by acting as an electron transferase. Heme I is the entrance site of electrons, that are transferred sequentially from heme I, II, III, IV and then to FAD.⁶⁴ The prosthetic groups are organized in an almost linear “wire” with distances less than 8 Å that allow fast electron transfer.⁶⁵ The proton necessary to form FADH^- most likely comes from solvent when Fcc₃

assumes the open conformation,⁶³ while the electrons come from another membrane-bound cytochrome *c* named CymA that removes electrons from the bacterial quinol pool in the inner membrane and passes them to heme I of Fcc₃.⁶⁶ Mechanistic details regarding the order of electron and proton transfer in the reductive reaction are currently unknown.

The oxidative reaction of Fcc₃ was extensively studied.^{4,27,62,63,67–70} The FAD-binding domain and the clamp domain are conserved in all fumarate reductases and succinate dehydrogenases, soluble and membrane-bound, suggesting the enzymatic mechanism of oxidoreduction of fumarate/succinate (Fig. 4) is ubiquitous.⁴⁹ In fact, enzymes with resolved structures show an absolute conservation of active site residues^{4,57–59,61,71–74} and the only difference is that membrane-bound fumarate reductases and succinate dehydrogenases have a covalently bound FAD^{57–59,71–74} with higher reduction potential (e.g. -55 mV in *E. coli*⁷⁵) than that of soluble fumarate reductases (-152 mV in the Fcc₃ of *Shewanella frigidimarina*⁶³). Ultimately, the lower reduction potential of soluble fumarate reductases make them unidirectional enzymes that are unable to catalyze the reverse reaction of succinate oxidation.⁶³

The reduction of fumarate in Fig. 4 is an addition reaction where two protons and two electrons are added to the C=C double-bond of fumarate. In the enzymatic reaction, the two electrons are transferred with a proton from FADH[−] nitrogen 5 to fumarate carbon 2 as a hydride equivalent^{4,67} (see Fig. 6a for residues and atom numbering). The other proton is added to the carbon 3 of fumarate from catalytic R402.^{68,69,76} Despite the high pK_a of arginines,⁷⁷ fumarate reductases and succinate dehydrogenases developed a mechanism where the conserved residues R402, E378 and R381 (Fig. 6b) form a proton path that connects the active site to the surface of the protein, allowing fast proton exchange with solvent to reprotonate R402 without having to transition to the open conformation.⁷⁰

Although, there are no doubts about the identities of R402 and FADH[−] as proton and hydride donors, *the sequence of proton and hydride transfers is questionable*. Kinetic isotope effects data show divergent conclusions,^{78–80} with some pointing to a concerted mechanism where proton and hydride equivalents are transferred to fumarate in a single

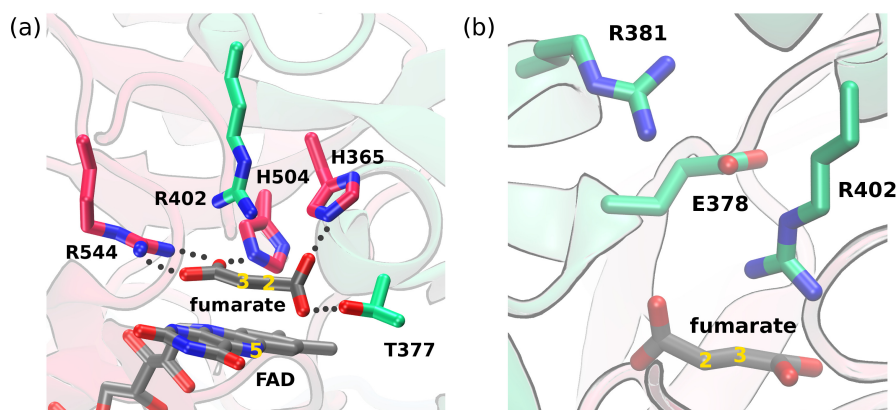


Figure 6 – Enzymatic reduction of fumarate. (a) Active site of Fcc₃ bound to fumarate in the closed conformation. The position and atom numbering of FAD, fumarate and some residues involved in substrate binding and catalysis are shown. The dotted lines represent hydrogen-bonds between fumarate and the enzyme. Residues in red belong to the FAD-binding domain and residues in green belong to the clamp domain. (b) Residues that form the proton path connecting the active site and the protein surface, allowing rapid proton exchange during the reaction.

step⁷⁸ and others pointing to a stepwise mechanism.⁸⁰

Figure 7 describes three possible mechanisms for fumarate reduction with respect to the sequence of hydride (H^-) and proton (H^+) transfer to fumarate. In nucleophilic addition, the hydride from FADH^- is transferred first forming a carbanion intermediate, whereas in the electrophilic addition, a proton from R402 is transferred first to generate a carbocation. In the concerted addition, the hydride from FADH^- and the proton from R402 are transferred together and no reaction intermediate is formed. An alternative mechanism, not shown in Figure 7, is the radical addition mechanism, where FADH^- would transfer H^\bullet and e^- separately, instead of H^- in a single step.

When Fcc₃ with oxidized FAD is co-crystallized with fumarate, an “oxygenated” intermediate is observed in the active site instead of fumarate,⁴ which was attributed to a nucleophilic attack of water to fumarate. This was taken as evidence that a stable intermediate is formed during fumarate reduction and that in a scenario with Fcc₃ bound to FADH^- instead of FAD, a hydride and not a water would perform this nucleophilic attack. Consequently, the most accepted view is that the enzymatic reduction of fumarate takes place in two steps with hydride transfer first, via nucleophilic addition. However,

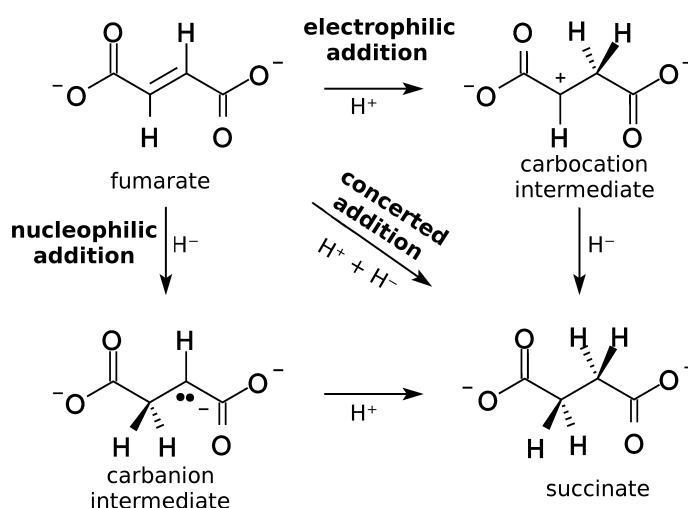


Figure 7 – Scheme with three possible mechanisms for fumarate reduction. In nucleophilic addition a hydride (H^-) is transferred first; in electrophilic addition a proton (H^+) is transferred first; and in the concerted addition H^- and H^+ are transferred together. In the enzymatic reaction, H^- comes from FADH^- and H^+ from R402.

this is not a very strong evidence and this hypothesis still needs validation.

Furthermore, the transfer of a hydride equivalent ($\text{H}^+ + 2e^-$) from FADH^- to fumarate can be synchronous or asynchronous (Fig. 8). In the synchronous transfer, the proton and the two electrons from FADH^- are transferred to fumarate at the same time,⁸¹ whereas in the asynchronous transfer either the proton or the electrons are transferred in advance. Because asynchronous transfers involve large charge separation, they are commonly associated with adduct formation,^{82,83} which is a common mechanism employed by flavoenzymes.⁸⁴

Fcc₃ catalyze fumarate reduction employing several strategies. Positively charged residues R544 and H504 are responsible for polarizing fumarate to facilitate hydride transfer.^{4,68} The residue H504, whose pK_a is 7.4,⁶⁸ could even work as a Brønsted acid protonating fumarate before its reduction.^{27,68,76} Active site topology also seems to be important, for example, fumarate has one of its carboxylates twisted in 90° because of hydrogen bonds with T377 and H365 (see Fig. 6a), which destabilize the electron conjugation throughout its double bonds, making it more susceptible to reduction.^{4,85–87} More importantly, the interaction between T377 and fumarate seems to be fundamental for

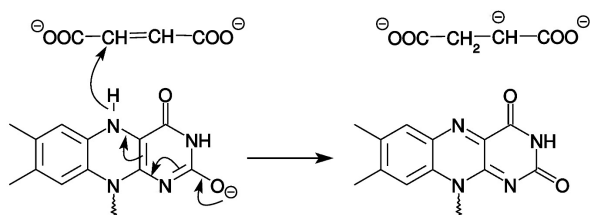
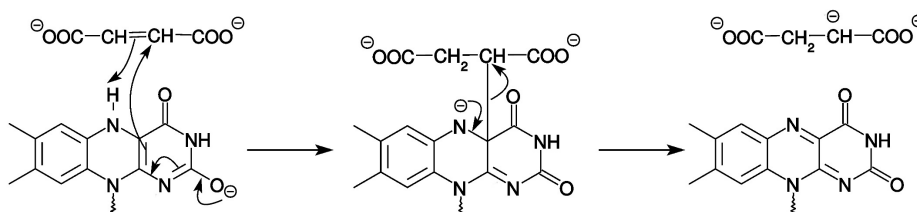
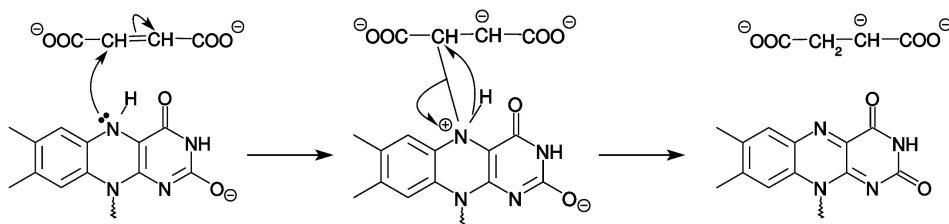
Synchronous transfer**Asynchronous transfer with advanced proton****Asynchronous transfer with advanced electrons**

Figure 8 – Possibilities of hydride ($\text{H}^+ + 2e^-$) transfer regarding the sequence of individual H^+ and e^- transfers. Synchronous proton and electron transfers; asynchronous transfer with advanced proton; and asynchronous transfer with advanced electrons. In the asynchronous mechanisms adducts with the flavin coenzyme are usually formed.

active site closure via clamp domain,^{86,87} expelling water molecules and aligning fumarate to an ideal geometry that ultimately facilitates the reaction.

To our knowledge, there is only one computational work addressing some of these questions.²⁷ They used a minimal model of the active site containing protonated fumaric acid, the isoalloxazine ring of FADH^- , a guanidinium as R402 and an imidazolium as H504. They conclude the reaction is a nucleophilic addition and that H504 protonates fumarate. However, as they have used a very small model, they could not answer many of the mechanistic questions raised here.

A larger model of the active site of Fcc₃ with 465 atoms was investigated in this thesis with electronic-structure calculations. Chapter 4 describes these calculations and

answers to the mechanistic questions after careful calibration and assessment of errors.

1.2 Electronic-structure calculations

In computational chemistry, theories are applied to obtain information relevant to understand chemical phenomena.⁸⁸ It is particularly important when the questions to be answered cannot be determined or are difficult to treat experimentally. In this thesis, in order to achieve the objectives set out in section 1.3, computational chemistry calculations were used.

The biggest challenge of computational chemistry is to select a suitable theoretical level to deal with a given chemical problem, so that the final results have small errors. In the works presented here, estimating the uncertainty of the calculations is a major goal.

Because understanding the limitations of the various methods is fundamental to interpreting the data generated by computational chemistry calculations, this section was designed to provide a general overview of the theories used here. First, the basic principles of electronic-structure theory are presented in sections 1.2.1 to 1.2.3, followed by descriptions of multiconfigurational quantum chemistry theory (section 1.2.4), perturbation theory (section 1.2.5), coupled-cluster theory (section 1.2.6) and density functional theory (section 1.2.7). Finally, a discussion is presented on how to incorporate thermodynamic contributions to quantum chemical calculations (section 1.2.8).

1.2.1 Electronic Hamiltonian

The most exact approach to treat chemical phenomena computationally is by employing *ab initio* quantum chemistry methods.⁸⁹ Because chemical systems are composed essentially of electrons and atomic nuclei, all of their physical and chemical properties can in principle be computed by solving the time-independent Schrödinger equation:⁹⁰

$$\hat{H}\Psi = E\Psi \quad (1.1)$$

where \hat{H} is the Hamiltonian operator, Ψ is the wave function that describes the chemical system and E is the total energy.

For a system of N particles, Ψ depends simultaneously on the position of each particle, making it impossible to solve the Schrödinger equation 1.1 analytically.⁹¹ Therefore, approximations are necessary and can be applied by either simplifying the wave function or by ignoring some physical interactions in the Hamiltonian. In addition, as *ab initio* methods do not apply empirical parameters, all approximations must rely on theoretical foundations.

The Born–Oppenheimer approximation,⁹² which assumes that electrons respond instantaneously to the motion of the nuclei, is commonly employed in *ab initio* quantum chemistry and introduces a simplification to \hat{H} . The mass of a proton is 1836 times the mass of an electron and in average the kinetic energy of the system must be distributed evenly among each particle according to the equipartition theorem. So, electrons will carry more momenta than protons and other heavier nuclei, moving much faster.

Fundamentally, it implies that considering atomic nuclei fixed in space and without kinetic energy is a good approximation. In this regime, \hat{H} becomes an electronic Hamiltonian \hat{H}_{el} , given by

$$\hat{H}_{\text{el}} = \hat{T}_{\text{e}} + \hat{V}_{\text{ne}} + \hat{V}_{\text{ee}} + \hat{V}_{\text{nn}} \quad (1.2)$$

where \hat{T}_{e} is the kinetic energy of the electrons, \hat{V}_{ne} is the Coulombic attraction potential between electrons and nuclei, \hat{V}_{ee} is the repulsion between electrons and \hat{V}_{nn} is the repulsion between nuclei. In atomic units:

$$\hat{T}_{\text{e}} = \frac{1}{2} \sum_i \nabla_i^2 \quad (1.3)$$

$$\hat{V}_{\text{ne}} = - \sum_i \sum_I \frac{Z_I}{r_{iI}} \quad (1.4)$$

$$\hat{V}_{\text{ee}} = \sum_{i < j} \frac{1}{r_{ij}} \quad (1.5)$$

$$\hat{V}_{\text{nn}} = \sum_{I < J} \frac{Z_I Z_J}{r_{IJ}} \quad (1.6)$$

with indices i, j labeling the electrons and I, J labeling the nuclei. Equation 1.3 is a sum over single electron contributions, whereas eqs. 1.4 to 1.6 are sums over pairs of particles.

The term \hat{V}_{nn} is independent of the positions of the electrons and contributes only with a constant energy term. ∇^2 is the Laplacian operator, r is the distance between two particles, and Z is the atomic number.

More commonly, \hat{H}_{el} is expressed in terms of the one-electron operator \hat{h}_i and the two-electron operator \hat{g}_{ij} :

$$\hat{H}_{\text{el}} = \sum_i \hat{h}_i + \sum_{i < j} \hat{g}_{ij} + \sum_{I < J} \frac{Z_I Z_J}{r_{IJ}} \quad (1.7)$$

where

$$\hat{h}_i = \frac{1}{2} \nabla_i^2 - \sum_I \frac{Z_I}{r_{iI}} \quad (1.8)$$

and

$$\hat{g}_{ij} = \frac{1}{r_{ij}} \quad (1.9)$$

The solution of the Schrödinger equation 1.1 for \hat{H}_{el} is known as an electronic-structure calculation and provides an electronic wave function $\Psi_{\text{el}}(\mathbf{r}_i; \mathbf{r}_I)$ that depends directly on the position of the electrons \mathbf{r}_i , but only parametrically on the position of the nuclei \mathbf{r}_I . The total energy $E(\mathbf{r}_I)$ is a function of the nuclei positions and defines the potential energy surface (PES), which is discussed in more detail in section 1.2.8.

1.2.2 Hartree–Fock theory

Within Born–Oppenheimer approximation, electronic-structure calculations consist basically in solving the Schrödinger equation for a system of N electrons in an external potential generated by the nuclei. In Hartree–Fock theory,⁹³ it is assumed that each electron does not interact directly with each other, but with a mean electrostatic field generated by the other $N - 1$ electrons. Hartree proposed⁹⁴ that within this approximation the N -electronic wave function $\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ can be written as product of one-electron orthonormal functions $\{\psi_p(\mathbf{x}_1), \psi_q(\mathbf{x}_2), \dots\}$, named molecular orbitals, that depend only on a single-electron spatial and spin coordinate \mathbf{x}_i . The indices p and q are arbitrary quantum numbers that characterize these orbitals.

As electrons are indistinguishable and have spin 1/2, they must follow Pauli exclusion principle.⁹⁵ Thus, an N -electronic wave function must also be antisymmetric

with respect to permutation of two electrons. A simple product of one-electron molecular orbitals $\psi_p(\mathbf{x}_i)$ as proposed by Hartree⁹⁴ would not respect this condition, so Fock and Slater noted that the antisymmetry requirement could be fulfilled if the wave function takes the form of a determinant instead of a simple product.^{96–98}

$$\Phi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) & \cdots & \psi_N(\mathbf{x}_1) \\ \psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) & \cdots & \psi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\mathbf{x}_N) & \psi_2(\mathbf{x}_N) & \cdots & \psi_N(\mathbf{x}_N) \end{vmatrix} \quad (1.10)$$

Here, the prefactor $1/\sqrt{N!}$ is the normalization constant and Φ is called a Slater determinant. In Hartree–Fock theory, the N -electronic system is assumed to be completely described by a wave function in the form of a single Slater determinant (i.e., $\Psi \equiv \Phi$). Alternatively, when Hartree–Fock fails, Φ can also be used as a starting point for more advanced calculations⁹⁹ (sections 1.2.4 to 1.2.6).

Starting with a Slater determinant, the task is then to find the set of molecular orbitals $\{\psi_p(\mathbf{x}_i)\}$. Considering Φ a trial function, the variational principle states that by adjusting $\{\psi_p(\mathbf{x}_i)\}$, the expectation value $E[\Phi] = \langle \Phi | \hat{H} | \Phi \rangle$ can be minimized to provide the best possible energy for the ground state, within the trial function limitations, with E an upper bound to the real energy.⁹⁰ By using the Hamiltonian (eq. 1.7) and the condition that $\{\psi_p(\mathbf{x}_i)\}$ forms an orthonormal set, the functional $E[\Phi]$ can be written as

$$E[\Phi] = \sum_p h_{pp} + \frac{1}{2} \sum_p \sum_q \langle pq || pq \rangle \quad (1.11)$$

where

$$h_{pp} = \langle \psi_p(\mathbf{x}_i) | \hat{h}_i | \psi_p(\mathbf{x}_i) \rangle \quad (1.12)$$

and

$$\langle pq || pq \rangle = \langle \psi_p(\mathbf{x}_i) \psi_q(\mathbf{x}_j) | \hat{g}_{ij} | \psi_p(\mathbf{x}_i) \psi_q(\mathbf{x}_j) \rangle - \langle \psi_p(\mathbf{x}_i) \psi_q(\mathbf{x}_j) | \hat{g}_{ij} | \psi_p(\mathbf{x}_j) \psi_q(\mathbf{x}_i) \rangle \quad (1.13)$$

are one-electron and two-electron integrals respectively.

Equation 1.11 is the final expression for the energy of an N -electron system with wave function given by a Slater determinant. Using Lagrange multipliers (ε_p) to minimize

$E[\Phi]$ with the constraint of orthonormality $\langle \psi_p | \psi_q \rangle = \delta_{pq}$, a system of Hartree–Fock equations is generated:

$$\hat{f}_i \psi_p(\mathbf{x}_i) = \varepsilon_p \psi_p(\mathbf{x}_i) \quad (1.14)$$

with

$$\hat{f}_i = \hat{h}_i + \hat{v}_i^{\text{HF}} \quad (1.15)$$

where \hat{f}_i is called Fock operator and it acts like an “effective Hamiltonian” for one-electron molecular orbitals $\psi_p(\mathbf{x}_i)$. Here, \hat{h}_i is the same of eq. 1.8 and \hat{v}_i^{HF} , given by

$$\hat{v}_i^{\text{HF}} \psi_p(\mathbf{x}_i) = \sum_{q \neq p} \left[\int \frac{\psi_q^*(\mathbf{x}_j) \psi_q(\mathbf{x}_j)}{r_{ij}} d\mathbf{x}_j \right] \psi_p(\mathbf{x}_i) - \sum_{q \neq p} \left[\int \frac{\psi_q^*(\mathbf{x}_j) \psi_p(\mathbf{x}_j)}{r_{ij}} d\mathbf{x}_j \right] \psi_q(\mathbf{x}_i) \quad (1.16)$$

is a mean-field operator, that describes the interaction of electron i with all other $N - 1$ electrons. In the first summation, the integral describes the average repulsion caused by the charge density of one electron in orbital ψ_q interacting with electron i . As the sum is over all molecular orbitals, the first term gives a mean electrostatic repulsion felt by electron i due to the other electrons. The second term is the exchange interaction and the integral is over the overlap of orbitals ψ_p and ψ_q . This is a pure quantum mechanical interaction and is a consequence of Pauli exclusion principle.

The molecular orbitals $\{\psi_p(\mathbf{x}_i)\}$ depend on both spatial and spin coordinate of the electron and each $\psi_p(\mathbf{x}_i)$ can be written as a product of a spatial function $\phi_p(\mathbf{r}_i)$ and a spin function $\sigma(\omega)$. The spatial coordinate \mathbf{r}_i is a three-dimensional vector position, whereas ω is a spin coordinate and can assume only two values: $\frac{1}{2}$ or $-\frac{1}{2}$. To solve the Hartree–Fock equations 1.14, restrictions can be applied to $\psi_p(\mathbf{x}_i)$ to keep its spatial part $\phi_p(\mathbf{r}_i)$ constant for both $\sigma(\frac{1}{2})$ and $\sigma(-\frac{1}{2})$ (Restricted Hartree–Fock, RHF) or $\phi_p(\mathbf{r}_i)$ can be left free to change (Unrestricted Hartree–Fock, UHF).

In general, $\{\phi_p(\mathbf{r}_i)\}$ cannot be obtained directly by solving the Fock equations 1.14. So, a linear combination of atomic orbitals (LCAO) is used:

$$\phi_p(\mathbf{r}_i) = \sum_{\mu=1}^M c_{\mu p} \chi_{\mu}(\mathbf{r}_i) \quad (1.17)$$

where $c_{\mu p}$ are the coefficients of the expansion and $\{\chi_{\mu}(\mathbf{r}_i)\}$ is a basis set of atomic orbitals with dimension M . Each atomic orbital $\chi_{\mu}(\mathbf{r}_i = r, \theta, \phi)$ is shaped like hydrogen-atom

orbitals,⁹⁰ with the angular part given by spherical harmonics $Y(\theta, \varphi)$ and the radial part by a Gaussian function:

$$\chi_\mu(r, \theta, \varphi) \propto Y(\theta, \varphi) e^{-\alpha r^2} \quad (1.18)$$

By replacing eq. 1.17 into 1.14, one can arrive at the Roothaan–Hall equations:^{100, 101}

$$\mathbf{F}\mathbf{C}_p = \mathbf{S}\mathbf{C}_p\varepsilon_p \quad (1.19)$$

with \mathbf{C}_p a column vector containing the coefficients of the expansion of orbital $\phi_p(\mathbf{r}_i)$ with energy ε_p , and \mathbf{F} , \mathbf{S} square matrices of dimension $M \times M$, named Fock matrix and overlap matrix, with elements given by

$$F_{\mu\nu} = \langle \chi_\mu(\mathbf{r}_i) | \hat{f}_i | \chi_\nu(\mathbf{r}_i) \rangle \quad (1.20)$$

$$S_{\mu\nu} = \langle \chi_\mu(\mathbf{r}_i) | \chi_\nu(\mathbf{r}_i) \rangle \quad (1.21)$$

The use of LCAO introduces a large simplification to the Hartree–Fock theory, by transforming the complex integro-differential equations 1.14 into the simpler eigenvalue equation 1.19, that is solved by a self-consistent field (SCF) method,¹⁰² where the coefficients $c_{\mu p}$ are optimized iteratively while the energy is minimized.

In summary, SCF is a numerical procedure used to solve the Hartree–Fock equations by introducing a finite basis set of atomic orbitals. In a system with N electrons, the SCF calculation can be performed with a basis set of size $M \geq N$ to generate M spatial molecular orbitals $\{\phi_p(\mathbf{r}_i)\}$ in RHF or $2M$ spin orbitals $\{\psi_p(\mathbf{x}_i)\}$ in UHF. However, as the final wave function is given by a Slater determinant (eq. 1.10), only the N molecular orbitals with the lowest energies ε_p will be occupied by electrons and contribute to the final energy (aufbau principle). The remaining $M - N$ nonoccupied orbitals are named virtual orbitals and are important for methods beyond the mean-field approximation of Hartree–Fock theory as discussed in sections 1.2.4 to 1.2.6.

1.2.3 Electron correlation

There are two main limitations in Hartree–Fock theory. The first is that the motion of electrons are not correlated. The operator \hat{v}_i^{HF} (eq. 1.16) used to describe the electron–electron interaction is an average potential where each electron feels a repulsion from the other occupied orbitals and not from the electrons themselves. The second limitation is that Hartree–Fock theory provides a wave function with a single Slater determinant, which might be a problem when the system under study allows multiple electronic configurations to coexist. To overcome the first limitation, dynamical electron correlation must be recovered, whereas for the second limitation static electron correlation must be calculated.¹⁰³

Static correlation effects only arise when near-degenerate states are possible, for example, in transition metal compounds, excited molecules, transition states and radicals. In these examples, the Hartree–Fock solution fails qualitatively, because a single Slater determinant is unable to capture evenly the effects of the multiple configurations.

Dynamical electron correlation are always present in multielectronic systems, because they arise from the direct interaction of electrons, that avoid coming too close to each other and are ignored in Hartree–Fock theory. So, recovering dynamical correlation is always necessary to obtain quantitative results with electronic-structure calculations.

1.2.4 Multiconfigurational quantum chemistry

Multiconfigurational quantum chemistry comprises a range of methods that perform electronic-structure calculations taking into account the effect of static electron correlation, i.e for systems that show nearly degenerate configurations and cannot be described by a single Slater determinant. The most natural way to overcome this problem is by treating the wave function Ψ as a linear combination of Slater determinants Φ_k (eq. 1.10),⁹⁹ instead of using only one determinant as in Hartree–Fock theory:

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \sum_k C_k \Phi_k(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \quad (1.22)$$

Here $\{\Phi_k\}$ form a basis of Slater determinants with different molecular orbitals $\{\psi_p(\mathbf{x}_i)\}$ being occupied, and C_k are the expansion coefficients. For example, the Hartree–Fock

determinant, here denoted Φ_{HF} (eq. 1.10), has the N lowest energy molecular orbitals $\{\psi_p(\mathbf{x}_i)\}$ occupied, but others Φ_k can be generated by occupying higher energy orbitals, that is, the virtual orbitals generated by the LCAO expansion.

The method of variationally optimizing the coefficients C_k of the expansion in eq. 1.22 is called configuration interaction (CI). If the set $\{\Phi_k\}$ is composed of all possible excitation patterns, Ψ is the full configuration interaction (FCI) wave function and will give the best possible solution for the Schrödinger equation. However, the expansion in eq. 1.22 for the FCI case grows factorially with the number of electrons N and molecular orbitals M , making it impossible to solve for all but very small systems. Instead, in multiconfigurational quantum chemistry, only some determinants are usually employed, the ones that describe the nearly degenerate configurations.¹⁰³

The spatial part of the molecular orbitals $\{\phi_p(\mathbf{r}_i)\}$ that form $\{\Phi_k\}$ are themselves expanded in basis functions of atomic orbitals (eq. 1.17). So, there are two sets of parameters to be optimized: the CI coefficients C_k of eq. 1.22 and the $c_{\mu p}$ coefficients of the molecular orbitals of eq. 1.17. In multiconfigurational self-consistent field (MCSCF) calculations, both sets of coefficients are optimized.

Among MCSCF methods, the most popular and robust is the complete active-space self-consistent field (CASSCF). In this method, a subset of the spatial molecular orbitals $\{\phi_p(\mathbf{r}_i)\}$ that might be involved with the near-degeneracy effects, are selected and a number of electrons are attributed to them. This is the active space and its orbitals are denominated active orbitals, while the remaining orbitals are classified as inactive orbitals or external orbitals. By definition, inactive orbitals will always be doubly occupied ($n = 2$) and external orbitals empty ($n = 0$), whereas the active orbitals are allowed to have occupation of $n = 0, 1$ or 2 electrons to form the subset of Slater determinants $\{\Phi_k\}$ used in eq. 1.22.

The CASSCF method is fully variational and can be solved iteratively by a SCF procedure that adjusts the coefficients C_k and $c_{\mu p}$ to minimize the energy. At convergence, the gradient of energy with respect to variations on the coefficients must be zero and the energy obtained will be an upper bound to the true energy. However, the optimization of a

CASSCF wave function is considerable more difficult than in a Hartree–Fock calculation, because the CI and orbital coefficients are very interdependent and the space of parameters are much larger than in Hartree–Fock.

The development of fast and robust algorithms to perform CASSCF calculations is an area of intense research.^{104,105} Also, the calculation of CASSCF wave functions can be very difficult and requires strategies in order to achieve convergence. In Appendix A, a discussion is presented about the approaches used to set and converge the CASSCF calculations reported in this thesis.

1.2.5 Perturbation theory

Rayleigh–Schrödinger perturbation theory is one of the theories that can be employed to recover dynamical electron correlation from a mean-field wave function like Hartree–Fock or CASSCF.^{99,106} It assumes these wave functions are exact solutions of an approximate “unperturbed” Hamiltonian $\hat{H}^{(0)}$ that can be mapped to the exact N -electronic Hamiltonian \hat{H} with increasing order of accuracy.

Mathematically, a perturbed Hamiltonian \hat{H}^λ can be written as:

$$\hat{H}^\lambda = \hat{H}^{(0)} + \lambda \hat{V} \quad (1.23)$$

where $\hat{H}^{(0)}$ is the inexact zeroth-order Hamiltonian with known wave functions Φ_k , \hat{V} is a perturbation operator that describes the contribution of electron correlation and λ is a coupling constant, which becomes 0 in the limit of no electron correlation ($\hat{H}^\lambda = \hat{H}^{(0)}$) and 1 when the Hamiltonian is exact, i.e. $\hat{H}^\lambda = \hat{H}$.

The corresponding exact ground-state energy E_0 and wave function Ψ_0 can be expanded as power series in terms of λ as:

$$E_0 = E_0^{(0)} + \lambda E_0^{(1)} + \lambda^2 E_0^{(2)} + \dots \quad (1.24)$$

$$\Psi_0 = \Psi_0^{(0)} + \lambda \Psi_0^{(1)} + \lambda^2 \Psi_0^{(2)} + \dots \quad (1.25)$$

with the zeroth-order wave function $\Psi_0^{(0)}$ equal to the self-consistent field wave function Φ_0 obtained with the unperturbed $\hat{H}^{(0)}$.

By inserting eqs. 1.24 and 1.25 into the Schrödinger equation 1.1 and grouping the terms by order of λ , a set of equations is generated. For each equation, an energy term correction can be obtained by left multiplying with Φ_0 and assuming the normalization condition $\langle \Phi_0 | \Psi_0^{(i)} \rangle = \delta_{0i}$:

$$E_0^{(0)} = \langle \Phi_0 | \hat{H}^{(0)} | \Phi_0 \rangle \quad (1.26)$$

$$E_0^{(1)} = \langle \Phi_0 | \hat{V} | \Phi_0 \rangle \quad (1.27)$$

$$E_0^{(2)} = \langle \Phi_0 | \hat{V} | \Psi_0^{(1)} \rangle \quad (1.28)$$

The zeroth- and first-order corrections to the energy (eqs. 1.26 and 1.27) only use integrals calculated over the zeroth-order wave function Φ_0 , but for the second-order correction $E_0^{(2)}$ (eq. 1.28), the perturbed wave function $\Psi_0^{(1)}$ has to be calculated. It can be obtained by left multiplying the first-order terms by Φ_k , instead of Φ_0 :

$$\Psi_0^{(1)} = - \sum_{k \neq 0} \frac{\langle \Phi_k | \hat{V} | \Phi_0 \rangle}{E_k^{(0)} - E_0^{(0)}} \Phi_k \quad (1.29)$$

where, $E_k^{(0)}$ are the energies of the excited states calculated with $\hat{H}^{(0)}$.

Finally, by replacing eq. 1.29 into eq. 1.28, the expression for the second-order correction to the energy is obtained:

$$E_0^{(2)} = - \sum_{k \neq 0} \frac{|\langle \Phi_0 | \hat{V} | \Phi_k \rangle|^2}{E_k^{(0)} - E_0^{(0)}} \quad (1.30)$$

A special case of Rayleigh–Schrödinger perturbation theory for systems with negligible static correlation is the Møller–Plesset perturbation theory (MP),^{107,108} where $\hat{H}^{(0)}$ is chosen to be the N -electron Hartree–Fock operator:

$$\hat{F}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \sum_i^N \hat{f}_i(\mathbf{x}_i) \quad (1.31)$$

with $\hat{f}_i(\mathbf{x}_i)$, the one-electron Fock operator that appears in the Hartree–Fock equations 1.15. Thus, the perturber \hat{V} must be the difference between the actual electron–electron repulsion given by the Coulomb potential and the mean-field Hartree–Fock repulsion \hat{v}_i^{HF} (eq. 1.16):

$$\hat{V}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \frac{1}{2} \sum_i^N \sum_{i \neq j}^N \frac{1}{r_{ij}} - \sum_i^N \hat{v}_i^{\text{HF}} \quad (1.32)$$

The ground-state zeroth-order wave function Φ_0 will be the Hartree–Fock determinant Φ_{HF} and Φ_k can be the single, double, triple, etc, excited determinants, i.e. $\Phi_k = \Phi_p^a, \Phi_{pq}^{ab}, \dots$, with p and q the occupied molecular orbitals in Φ_{HF} that are excited to virtual molecular orbitals a and b .

Because $\Phi_0 = \Phi_{\text{HF}}$ in MP perturbation theory, the zeroth-order contribution to the ground-state energy $E_0^{(0)}$ (eq. 1.26) is simply the sum of occupied molecular orbitals energies ε_p :

$$E_0^{(0)} = \sum_p^{\text{occ}} \varepsilon_p \quad (1.33)$$

whereas, the first- and second-order corrections can be calculated with eqs. 1.27 and 1.30, by replacing the wave functions Φ_{HF} and Φ_k and applying Slater–Condon rules^{91,99} to simplify the integrals:

$$E_0^{(1)} = -\frac{1}{2} \sum_{p,q}^{\text{occ}} \langle pq || pq \rangle \quad (1.34)$$

$$E_0^{(2)} = -\frac{1}{4} \sum_{p,q}^{\text{occ}} \sum_{a,b}^{\text{vir}} \frac{|\langle pq || ab \rangle|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_p - \varepsilon_q} \quad (1.35)$$

The sum of $E_0^{(0)}$ and $E_0^{(1)}$ gives exactly the Hartree–Fock energy and the first contribution due to electron correlation appears at second order with $E_0^{(2)}$ correction. The sum of these three contributions is the MP2 energy (Second-order Møller–Plesset energy), which can be improved by higher-order perturbations. However, it is common to stop the correction at second-order, because calculations become computationally expensive at higher-orders and the computed energies start to show oscillatory behavior.¹⁰⁸

For multiconfigurational calculations, dynamical correlation can also be computed by Rayleigh–Schrödinger perturbation theory, but differently from Møller–Plesset, the zeroth-order wave function will not be the Hartree–Fock determinant, but the multireference wave function (eq. 1.22). This introduces additional difficulties both theoretically and computationally, which led to the development of several approaches that differ basically in how to select the zeroth-order wave function Φ_0 and how to partition the Hamiltonian in $\hat{H}^{(0)}$ and \hat{V} .¹⁰⁹

In this thesis, only the n -electron valence state perturbation theory (NEVPT2) was employed.¹¹⁰ It provides second-order corrections to CASSCF energy similar to MP2 for Hartree–Fock. It has two formulations,¹¹¹ the partially contracted NEVPT2 (PC-NEVPT2) and the strongly contracted NEVPT2 (SC-NEVPT2), which differ in the basis Φ_k used to expand the wave function. SC-NEVPT2 uses a more compact expansion to be less computationally demanding in exchange of being slightly less accurate.

The main characteristic of NEVPT2, that differs it from other multiconfigurational perturbation theory methods, is that it employs the Dyall Hamiltonian¹¹² as $\hat{H}^{(0)}$. This Hamiltonian is exact for active orbitals and equivalent to the Møller–Plesset Hamiltonian \hat{F} for inactive and external orbitals, avoiding some technical difficulties like the appearance of intruder states.¹¹⁰

1.2.6 Coupled-cluster theory

Even though, methods like MP2 and NEVPT2 are able to recover up to 80% of dynamical electron correlation,^{99,108} chemical accuracy, i.e. results calculated with errors less than 1 kcal.mol⁻¹, is usually not achieved with perturbation theory. Coupled-cluster theory, in particular the coupled-cluster singles doubles with perturbative triples – CCSD(T) provide much more accurate results,¹¹³ being able to predict chemical and physical properties at chemical accuracy.

In coupled-cluster theory the correlated wave function Ψ_{CC} takes the form of an exponential ansatz:

$$\Psi_{CC} = e^{\hat{T}} \Phi_0 \quad (1.36)$$

where Φ_0 is the reference wave function, typically the Hartree–Fock determinant Φ_{HF} and \hat{T} is the cluster operator defined as:

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \dots + \hat{T}_N \quad (1.37)$$

Here, \hat{T}_1 , \hat{T}_2 , etc, are excitation operators responsible to generate determinants single-excited, double-excited and so on. They are better defined in the formalism of second

quantization¹¹⁴ using the creation (\hat{a}_a^\dagger) and annihilation (\hat{a}_p) operators:

$$\hat{T}_1 = \sum_p^{\text{occ}} \sum_a^{\text{vir}} t_p^a \hat{a}_a^\dagger \hat{a}_p \quad (1.38)$$

$$\hat{T}_2 = \sum_{p < q}^{\text{occ}} \sum_{a < b}^{\text{vir}} t_{pq}^{ab} \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_q \hat{a}_p \quad (1.39)$$

Again indices p and q refer to occupied orbitals and a and b to virtual orbitals. The coefficients t_p^a and t_{pq}^{ab} are named single-excitation cluster amplitudes and double-excitation cluster amplitudes. When $\hat{a}_a^\dagger \hat{a}_p$ act on Φ_0 , the single-excited determinant Φ_p^a is generated, and $\hat{a}_{ab}^\dagger \hat{a}_{pq}$ operating on Φ_0 generates the double-excited determinant Φ_{pq}^{ab} . The same happens for triple excitations with \hat{T}_3 , quadruple excitations with \hat{T}_4 , etc.

Considering the truncation of \hat{T} for coupled-cluster singles doubles (CCSD): $\hat{T} = \hat{T}_1 + \hat{T}_2$. The exponential ansatz can be expanded as:

$$\begin{aligned} e^{\hat{T}} &= 1 + \hat{T} + \frac{\hat{T}^2}{2!} + \frac{\hat{T}^3}{3!} + \dots \\ e^{\hat{T}} &= 1 + (\hat{T}_1 + \hat{T}_2) + \left(\frac{\hat{T}_1^2 + 2\hat{T}_1\hat{T}_2 + \hat{T}_2^2}{2} \right) + \left(\frac{\hat{T}_1^3 + 3\hat{T}_1^2\hat{T}_2 + 3\hat{T}_1\hat{T}_2^2 + \hat{T}_2^3}{6} \right) + \dots \quad (1.40) \\ e^{\hat{T}} &= 1 + \underbrace{\hat{T}_1}_{\text{singles}} + \underbrace{\left(\hat{T}_2 + \frac{1}{2}\hat{T}_1^2 \right)}_{\text{doubles}} + \underbrace{\left(\hat{T}_1\hat{T}_2 + \frac{1}{6}\hat{T}_1^3 \right)}_{\text{triples}} + \underbrace{\left(\frac{1}{2}\hat{T}_2^2 + \frac{1}{2}\hat{T}_1^2\hat{T}_2 + \frac{1}{24}\hat{T}_1^4 \right)}_{\text{quadruples}} + \dots \end{aligned}$$

where each term in parenthesis correspond to a different excitation level. For example, \hat{T}_1 generates single excitations. Double excitations can be generated by two ways: by \hat{T}_2 that excites two electrons “simultaneously” or by \hat{T}_1^2 that excites two electrons independently. Triple excitations can be generated by $\hat{T}_1\hat{T}_2$ or \hat{T}_1^3 . Whereas quadruple excitations can be generated by \hat{T}_2^2 , $\hat{T}_1^2\hat{T}_2$ or \hat{T}_1^4 .

By substituting eq. 1.40 into eq. 1.36, an expression for the wave function Ψ in terms of the determinants is generated:

$$\begin{aligned} \Psi_{\text{CCSD}} &= e^{\hat{T}} \Phi_{\text{HF}} \\ \Psi_{\text{CCSD}} &= \left[1 + \hat{T}_1 + \left(\hat{T}_2 + \frac{1}{2}\hat{T}_1^2 \right) + \left(\hat{T}_1\hat{T}_2 + \frac{1}{6}\hat{T}_1^3 \right) + \dots \right] \Phi_{\text{HF}} \\ \Psi_{\text{CCSD}} &= \Phi_{\text{HF}} + \hat{T}_1 \Phi_{\text{HF}} + \left(\hat{T}_2 + \frac{1}{2}\hat{T}_1^2 \right) \Phi_{\text{HF}} + \left(\hat{T}_1\hat{T}_2 + \frac{1}{6}\hat{T}_1^3 \right) \Phi_{\text{HF}} + \dots \quad (1.41) \\ \Psi_{\text{CCSD}} &= \Phi_{\text{HF}} + \sum_p^{\text{occ}} \sum_a^{\text{vir}} c_p^a \Phi_p^a + \sum_{p < q}^{\text{occ}} \sum_{a < b}^{\text{vir}} c_{pq}^{ab} \Phi_{pq}^{ab} + \sum_{p < q < r}^{\text{occ}} \sum_{a < b < c}^{\text{vir}} c_{pqr}^{abc} \Phi_{pqr}^{abc} + \dots \end{aligned}$$

with coefficients given by:

$$\begin{aligned}
 c_p^a &= t_p^a \\
 c_{pq}^{ab} &= t_{pq}^{ab} + t_p^a * t_q^b \\
 c_{pqr}^{abc} &= t_p^a * t_{qr}^{bc} + t_p^a * t_q^b * t_r^c \\
 &\vdots
 \end{aligned} \tag{1.42}$$

Here, the symbol $*$ represents an antisymmetric product, for example $t_p^a * t_q^b = t_p^a t_q^b - t_q^a t_p^b$.

The advantage of couple-cluster theory over other methods become clear here. In CCSD, despite the truncation of \hat{T} to include only contributions from the singles and doubles cluster operators \hat{T}_1 and \hat{T}_2 , higher-order excitations are also generated as “by-product” due to the exponential ansatz. In perturbation theory, all types of corrections (singles, doubles, triples, etc.) are included under the restriction imposed by the perturbation order, while coupled-cluster includes all corrections of a given type to infinite order. For example, CCSD includes all possible corrections formed by \hat{T}_1 and \hat{T}_2 excited determinants, and their respective product combinations, recovering more than 90% of dynamical electron correlation.

The solution of the Schrödinger equation for the coupled-cluster wave function (eq. 1.41) by the variational method would be too complex due to the N -fold excitations. So, coupled-cluster implementations use the projection method,¹¹³ providing the coupled-cluster correlation energy E_{CC} :

$$E_{CC} = \frac{1}{4} \sum_{p,q}^{\text{occ}} \sum_{a,b}^{\text{vir}} (t_{pq}^{ab} + 2t_p^a t_q^b) \langle pq || ab \rangle \tag{1.43}$$

No discussions about how to obtain the cluster amplitudes t_p^a and t_{pq}^{ab} will be made here, because they involve complex mathematical operations that employ Feynman diagrams and can be consulted elsewhere.¹¹³ Yet, an important observation is that eq 1.43 is not only valid for CCSD, but also for larger truncations like CCSDT and CCSDTQ for instance.¹¹³ The contributions of \hat{T}_3 and \hat{T}_4 appear indirectly when determining the amplitudes t_p^a and t_{pq}^{ab} .

Calculations beyond CCSD are extremely demanding computationally, which led to alternatives to include \hat{T}_3 contribution approximately.^{99,113} The most famous is CCSD(T),

which calculates the additional triples contributions perturbatively. This calculation is able to recover almost all dynamical correlation⁹⁹ and is considered the gold-standard method of quantum chemistry.

The biggest limitation of coupled-cluster, both CCSD and CCSD(T) is their computational cost, while the former scales as $\mathcal{O}(\mathcal{N}^6)$, the second scales as $\mathcal{O}(\mathcal{N}^7)$, with \mathcal{N} the number of molecular orbitals, limiting their application only to small systems. Recently, alternatives have been developed to reduce the computational cost of coupled-cluster calculations. Among them, the domain-based local-pair natural orbital (DLPNO) approximation developed by Neese et al,¹¹⁵ which assumes that dynamical correlation occurs locally, allowing a simplification that reduces the computational cost of coupled-cluster calculations to $\mathcal{O}(\mathcal{N})$, in exchange of a slightly loss of accuracy.

1.2.7 Density functional theory

Density functional theory (DFT) takes a different approach to treat chemical systems at quantum level and do not employ a wave function formalism. The basic idea behind DFT is that instead of solving the Schrödinger equation (eq. 1.1) to obtain the wave function Ψ , which depends on the position of each electron, one just need to calculate the electron density $\rho(\mathbf{r})$, which is a spatial function that depends on three coordinates only.¹¹⁶

This approach is possible because Hohenberg and Kohn proved with a theorem that the ground-state energy of an N -electronic system is a functional of its electron density $\rho(\mathbf{r})$.¹¹⁷ The exact form of this functional $E[\rho(\mathbf{r})]$ is not known, but it can be split into individual contributions:

$$E[\rho(\mathbf{r})] = T[\rho(\mathbf{r})] + E_{\text{ne}}[\rho(\mathbf{r})] + E_{\text{H}}[\rho(\mathbf{r})] + E_{\text{XC}}[\rho(\mathbf{r})] \quad (1.44)$$

where $T[\rho(\mathbf{r})]$ is the kinetic energy of the electrons, $E_{\text{ne}}[\rho(\mathbf{r})]$ is the energy contribution due to the interaction between electrons and nuclei, whereas $E_{\text{H}}[\rho(\mathbf{r})]$ and $E_{\text{XC}}[\rho(\mathbf{r})]$ describe the electron-electron interactions and are named Hartree energy and exchange-correlation energy, respectively. Coulombic repulsion between nuclei are omitted here, because they

contribute to a constant term within Born–Oppenheimer approximation (see section 1.2.1).

The functionals $E_{\text{ne}}[\rho(\mathbf{r})]$ and $E_{\text{H}}[\rho(\mathbf{r})]$ can be developed using classical electrostatics:¹¹⁶

$$E_{\text{ne}}[\rho(\mathbf{r})] = \sum_I \int \frac{Z_I \rho(\mathbf{r})}{|\mathbf{R}_I - \mathbf{r}|} d\mathbf{r} \quad (1.45)$$

$$E_{\text{H}}[\rho(\mathbf{r})] = \frac{1}{2} \int \int \frac{\rho(\mathbf{r}) \rho'(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \quad (1.46)$$

with the sum in eq. 1.45 over all nuclei I and the factor $1/2$ in eq. 1.46 appearing because $\rho(\mathbf{r})$ is integrated twice.

The deduction of $T[\rho(\mathbf{r})]$ is not straightforward and was only established by the pivotal work of Kohn and Sham that build the foundations of current DFT.¹¹⁸ They consider $T[\rho(\mathbf{r})]$ to be the kinetic energy of a non-interacting electron system:

$$T[\rho(\mathbf{r})] \approx -\frac{1}{2} \sum_p \langle \phi_p | \nabla^2 | \phi_p \rangle \quad (1.47)$$

with

$$\rho(\mathbf{r}) = \sum_p |\phi_p|^2 \quad (1.48)$$

where ϕ_p are one-electron functions analogous to the ones in Hartree–Fock theory, which can be mapped to the electron density $\rho(\mathbf{r})$ by eq. 1.48

More importantly, Kohn–Sham formalism provided a set of equations that can be used to determine $\rho(\mathbf{r})$ and the system’s energy self-consistently in the same way as in Hartree–Fock theory (eq. 1.14).

The determination of $E_{\text{ne}}[\rho(\mathbf{r})]$, $E_{\text{H}}[\rho(\mathbf{r})]$ and $T[\rho(\mathbf{r})]$ were all based on theoretical foundations. However, $E_{\text{XC}}[\rho(\mathbf{r})]$ does not have an analytic form and approximations have to be made to solve Kohn–Sham equations.

Many density functional approximations to $E_{\text{XC}}[\rho(\mathbf{r})]$ have been developed and it is not uncommon to employ empirical parameters or even fit the different contributions of $E_{\text{XC}}[\rho(\mathbf{r})]$ based on experimental data or high-level wave function calculations. Yet, DFT is considered an *ab initio* theory, because $E_{\text{XC}}[\rho(\mathbf{r})]$ can be improved based on a hierarchy of approximations.

The simplest approximation to $E_{\text{XC}}[\rho(\mathbf{r})]$ is the local density approximation (LDA), that assumes $\rho(\mathbf{r})$ will be identical to the electron density of a uniform electron gas, whose $E_{\text{XC}}[\rho(\mathbf{r})]$ is known exactly.¹¹⁹

The second hierarchy of approximation is the generalized gradient approximation (GGA), that assumes $E_{\text{XC}}[\rho(\mathbf{r})]$ will depend also on the gradient of the electron density $\nabla\rho$ and that the electron density is not uniform. This is the first level of approximation able to yield accurate electronic structures of chemical systems. Some density functional approximations at this level are: BLYP,^{120,121} PBE,¹²² OLYP^{121,123} and OPBE.^{122,123}

The third approximation level is called meta-GGA, and assumes $E_{\text{XC}}[\rho(\mathbf{r})]$ will also depend on the second derivatives of the electron density $\nabla^2\rho$. Some examples of methods at this level are TPSS¹²⁴ and M06-L¹²⁵ functionals.

The fourth level of hierarchy is the inclusion of non-local exchange information into GGA or meta-GGA functionals, using the Hartree–Fock formalism (eq. 1.10) to generate hybrid functionals. Examples of this approximation level are B3LYP,^{126,127} PBE0,¹²⁸ M06,¹²⁹ M06-2X,¹²⁹ O3LYP,¹³⁰ TPSSH.¹³¹

At the fifth level, range-separated hybrid functionals were developed, where Hartree–Fock exchange inclusion is dependent on the inter-electronic distance, opposing simple hybrids that include a fixed amount of Hartree–Fock exchange. Examples are CAM-B3LYP¹³² and ω B97X¹³³ approximations.

The final hierarchy level are the double-hybrid functionals, that also includes non-local correlation information. For example, B2PLYP¹³⁴ uses second-order perturbation theory with the virtual Kohn–Sham orbitals to recover additional dynamic correlation.

1.2.8 Connection to thermodynamics

Born–Oppenheimer approximation⁹² considers nuclei fixed in space, hence the resultant energy from an electronic-structure calculation in this regime (sections 1.2.2 to 1.2.7) will depend on the nuclei position. For a molecular system with N atoms, its energy $E(\mathbf{r}_1, \dots, \mathbf{r}_N)$ will be a function of N vector positions \mathbf{r}_I describing the location of nuclei.

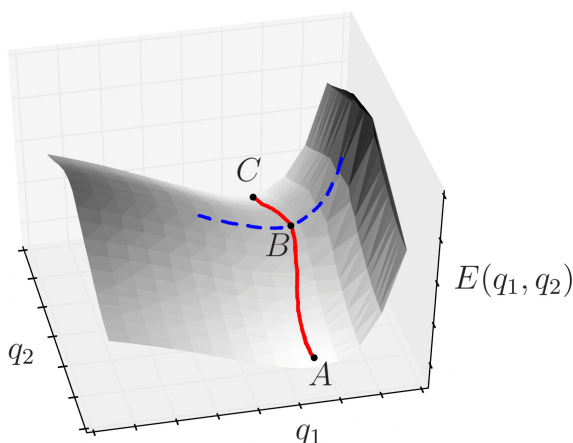


Figure 9 – Illustrative potential energy surface (PES) with energy $E(q_1, q_2)$ as a function of two arbitrary coordinates q_1 and q_2 . Points A , B and C are stationary points: A and C are minima; whereas B is a saddle point, i.e a minimum along q_1 (blue dashed line) and a maximum along q_2 . The red line shows the minimum energy path that connects A to C .

The direct mapping of nuclei positions to a single energy value, introduces the concept of potential energy surface (PES) to quantum chemistry formalism.¹³⁵

Figure 9 shows an example of a PES that depends on two arbitrary coordinates q_1 and q_2 . Points A , B and C correspond to stationary points on the PES ($\nabla E = 0$) and have a physical interpretation. A and C are local minima and represent stable chemical species, whereas B is a saddle point and corresponds to a transition state for the reaction $A \rightarrow C$.¹³⁵ The red line in Fig. 9 is the reaction path connecting species A and C through transition state B .

The information obtained from the PES can be interpreted in the context of transition state theory¹³⁶ to calculate reaction rates k and equilibrium constants K_{eq} by:

$$k = \frac{k_B T}{h} e^{-\Delta G^\ddagger / RT} \quad (1.49)$$

and

$$K_{eq} = e^{-\Delta G_0 / RT} \quad (1.50)$$

where k_B is the Boltzmann constant, h is the Planck constant, R is the universal gas constant, T is the absolute temperature, ΔG^\ddagger is the difference in Gibbs free energy between the transition state and reactant state and ΔG_0 is the difference in Gibbs free energy between product and reactant states.

The connection between microscopic potential energies E , obtained by solving the Schrödinger equation, and macroscopic thermodynamic observables, like ΔG , is in the scope of statistical mechanics.⁹⁵ This formalism assumes that each macroscopic observable is an ensemble average of a related microscopic function that depends only on the position and momenta of the particles that compose the system.

To obtain an ensemble to calculate macroscopic observables (measured experimentally), computational chemists usually employ molecular dynamics or Monte Carlo simulations¹³⁷ to sample the molecular system of interest. However, performing extensive sampling at quantum chemistry level is extremely demanding computationally, so it is usual to employ the rigid-rotor-harmonic-oscillator (RRHO)¹³⁶ approximation to include thermal effects in quantum chemistry calculations.

The RRHO approximation assumes a non-interacting system where the canonical partition function Q can be calculated exactly. Considering a system of N non-interacting identical molecules, its partition function Q can be written in terms of the partition function of a single molecule q using the relation $Q = q^N/N!$.¹³⁸ Because energies are additive, the calculation of q , can be split into calculating individual contributions and then multiplying them:

$$q = q_{\text{trans}} \times q_{\text{rot}} \times q_{\text{vib}} \times q_{\text{elec}} \quad (1.51)$$

with q_{trans} the partition function due to translation of the molecule, q_{rot} the rotation partition function, q_{vib} the vibration partition function and q_{elec} the electronic partition function. Each representing different degrees of freedom of the molecule.

The translational partition function q_{trans} is assumed to be the partition function of a particle in a box:¹³⁸

$$q_{\text{trans}} = \left(\frac{2\pi m k_{\text{B}} T}{h^2} \right)^{3/2} V \quad (1.52)$$

with m the mass of the molecule and V usually assumed to be the volume of one mol of an ideal gas.

For the molecular rotation, the partition function q_{rot} is approximated to be the

partition function of a rigid rotor:¹³⁸

$$q_{\text{rot}} = \frac{\sqrt{\pi}}{\sigma} \left(\frac{8\pi^2 k_B T}{h^2} \right)^{3/2} \sqrt{I_1 I_2 I_3} \quad (1.53)$$

Here, I_1 , I_2 and I_3 are the three moments of inertia of the molecule and σ is the symmetry index, whose value is given by the order of the molecular point group.

Molecular vibrations are approximated as harmonic oscillators, with frequencies ν given in terms of force constants and atomic masses. The force constants can be obtained by the second derivatives of the energy with respect to the nuclei positions \mathbf{r}_I . In a polynuclear molecule, with N atoms, the force constants are given by the Hessian matrix, that when diagonalized produce normal mode coordinates that can be used to solve the vibrational wave function as $3N$ one-dimensional harmonic oscillators.⁸⁸ The Schrödinger equation for the harmonic oscillator gives as solution the energies $\varepsilon_{\text{vib}} = (n + \frac{1}{2})h\nu$,⁹⁰ with n the vibrational quantum number. If n is set to zero, the ground vibrational state energy can be obtained, which is called zero-point vibrational energy (ZPVE) and describes the minimum energy the system can have at quantum level.

The vibrational partition function q_{vib} is then a product over each individual vibration i , that contributes with a sum over all its quantum states n :⁸⁸

$$q_{\text{vib}} = \prod_{i=1}^{3N-6} \left(\sum_{n=0}^{\infty} e^{-(n+\frac{1}{2})\frac{h\nu_i}{k_B T}} \right) = \prod_{i=1}^{3N-6} \frac{e^{-\frac{h\nu_i}{2k_B T}}}{1 - e^{-\frac{h\nu_i}{k_B T}}} \quad (1.54)$$

Here, the product over $3N - 6$ vibrations is because at normal mode coordinates, the six lowest energy modes correspond to the three translation and three rotation degrees of freedom, that are treated by eq. 1.52 and 1.53. Additionally, for transition states, one of the modes will have an imaginary frequency because it is a saddle point in the PES and corresponds to a vibration along the reaction path, thus the product in eq. 1.54 has to be calculated over $3N - 7$ vibrations.

At low frequencies the harmonic oscillator approximation usually breaks down, so an alternative approximation named quasi-rigid rotor harmonic oscillator (quasi-RRHO)¹³⁹ can be employed. Where low-frequency vibrations ($< 50 \text{ cm}^{-1}$) are treated as rotations.

The electronic partition function q_{elec} has to count all possible electronic quantum states. However, as excited states usually have high energy, the molecule remains in the

ground electronic state at normal temperatures and $q_{\text{elec}} = 1$, for singlet molecules. For doublet molecules, $q_{\text{elec}} = 2$ because of the spin degeneracy, for triplet molecules $q_{\text{elec}} = 3$ and so on.⁸⁸

After the calculation of each contribution to q , the canonical partition function for the ensemble $Q = q^N/N!$ can be calculated for 1 mol of molecules to derive the Helmholtz free energy A and the internal energy U of the system by:¹³⁸

$$A = -k_B T \ln Q \quad (1.55)$$

and

$$U = k_B T^2 \left(\frac{\partial \ln Q}{\partial T} \right)_{N,V} \quad (1.56)$$

However, the calculation of Gibbs free energy, to determine rates and equilibrium constants (eq. 1.49 and 1.50), depends on the system's enthalpy H and entropy S , that can be derived directly by eqs. 1.55 and 1.56.¹³⁸

$$H = U + PV = k_B T^2 \left(\frac{\partial \ln Q}{\partial T} \right)_{N,V} + k_B T V \left(\frac{\partial \ln Q}{\partial V} \right)_{N,T} \quad (1.57)$$

$$S = \frac{U - A}{T} = k_B T \left(\frac{\partial \ln Q}{\partial T} \right)_{N,V} + k_B \ln Q \quad (1.58)$$

where the relation $P = -(\frac{\partial A}{\partial V})_{N,T}$ was used to develop eq. 1.57.

Finally, using eq. 1.57 and 1.58, an expression for Gibbs free energy can be derived:

$$G = H - TS = k_B T V \left(\frac{\partial \ln Q}{\partial V} \right)_{N,T} - k_B T \ln Q \quad (1.59)$$

1.3 Objectives

The general goal of this thesis is to understand flavin and flavoenzyme reaction mechanisms in detail with electronic-structure calculations. Chapters 2 to 4 deal with different topics related to the reactivity of flavins and flavoenzymes.

In chapter 2, the main objective is to understand the limitations of different quantum chemistry methods in predicting ground-state properties of flavins like electron

affinity, gas-phase basicity, torsion energy, dipole moment and molecular geometry. A second aim is to determine the equilibrium composition of flavin tautomers in various redox and charge states. Ultimately, these objectives can help building better models to study flavin reactions.

The general objective of chapter 3 is to characterize proton-coupled electron transfers mediated by flavins. The first objective is to develop a diagnostic tool able to identify if a hydride or a hydrogen-atom is transferred to flavin during a given reaction. Then, the second objective is to apply this formalism to characterize the NADH and succinate oxidation reactions, both involving proton-coupled electron transfers to flavin.

In chapter 4 the main objective is determine the reaction mechanism of fumarate reduction in the Fcc₃ flavoenzyme. Additional objectives include assessing which factors aid catalysis and to what extent proton tunneling and excited states contribute to the reaction.

2 Molecular properties and tautomeric equilibria of isolated flavins

Felipe Curtolo and Guilherme M. Arantes

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

Reproduced with permission from Curtolo, F.; Arantes, G. M. *J. Comput. Chem.* **2022**, 43(23), 1561–1572. DOI: [10.1002/jcc.26957](https://doi.org/10.1002/jcc.26957). Copyright 2022 Wiley & Sons, Inc.

RESEARCH ARTICLE

Molecular properties and tautomeric equilibria of isolated flavins

Felipe Curtolo  | Guilherme M. Arantes 

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

Correspondence

Felipe Curtolo 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.
Email: felipe.curtolo@usp.br and garantes@iq.usp.br

Funding information

Fundação de Amparo à Pesquisa do Estado de São Paulo, Grant/Award Numbers: 2017/26109-0, 2019/21856-7

Abstract

Flavins are employed as redox cofactors and chromophores in a plethora of flavoenzymes. Their versatility is an outcome of intrinsic molecular properties of the isoalloxazine ring modulated by the protein scaffold and surrounding solvent. Thus, an investigation of isolated flavins with high-level electronic-structure methods and with error assessment of the calculated properties will contribute to building better models of flavin reactivity. Here, we benchmarked ground-state properties such as electron affinity, gas-phase basicity, dipole moment, torsion energy, and tautomer stability for lumiflavins in all biologically relevant oxidation and charge states. Overall, multiconfigurational effects are small and chemical accuracy is achieved by coupled-cluster treatments of energetic properties. Augmented basis sets and extrapolations to the complete basis-set limit are necessary for consistent agreement with experimental energetics. Among DFT functionals tested, M06-2X shows the best performance for most properties, except gas-phase basicity, in which M06 and CAM-B3LYP perform better. Moreover, dipole moments of radical flavins show large deviations for all functionals studied. Tautomers with noncanonical protonation states are significantly populated at normal temperatures, adding to the complexity of modeling flavins. These results will guide future computational studies of flavoproteins and flavin chemistry by indicating the limitations of electronic-structure methodologies and the contributions of multiple tautomeric states.

KEYWORDS

electronic-structure, flavoprotein, isoalloxazine, quantum chemistry, tautomerism

1 | INTRODUCTION

Flavins are prosthetic groups composed by the fused tricyclic ring isoalloxazine (benzo[g]pteridine, Figure 1). All natural flavins are methylated at carbons C7 and C8, but the group bound to nitrogen N10 varies.^[1] Riboflavin, or vitamin B₂, has a ribityl group bound to N10, while its photoproduct, lumiflavin, has a methyl bound to N10.^[2] The ubiquitous coenzymes flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD) have ribityl phosphate and ribityl-adenosine diphosphate moieties respectively attached to N10.

Proteins equipped with flavins, known as flavoproteins, are involved in a wide range of catalytic and signaling processes.^[2,3]

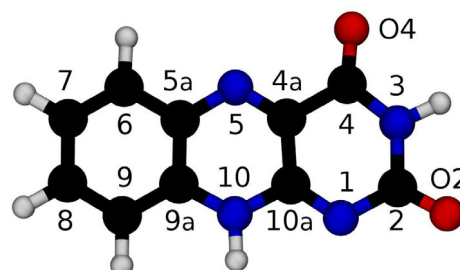


FIGURE 1 Structure and heavy-atom numbering for the isoalloxazine ring, the core group of flavins responsible for their redox and photophysical properties

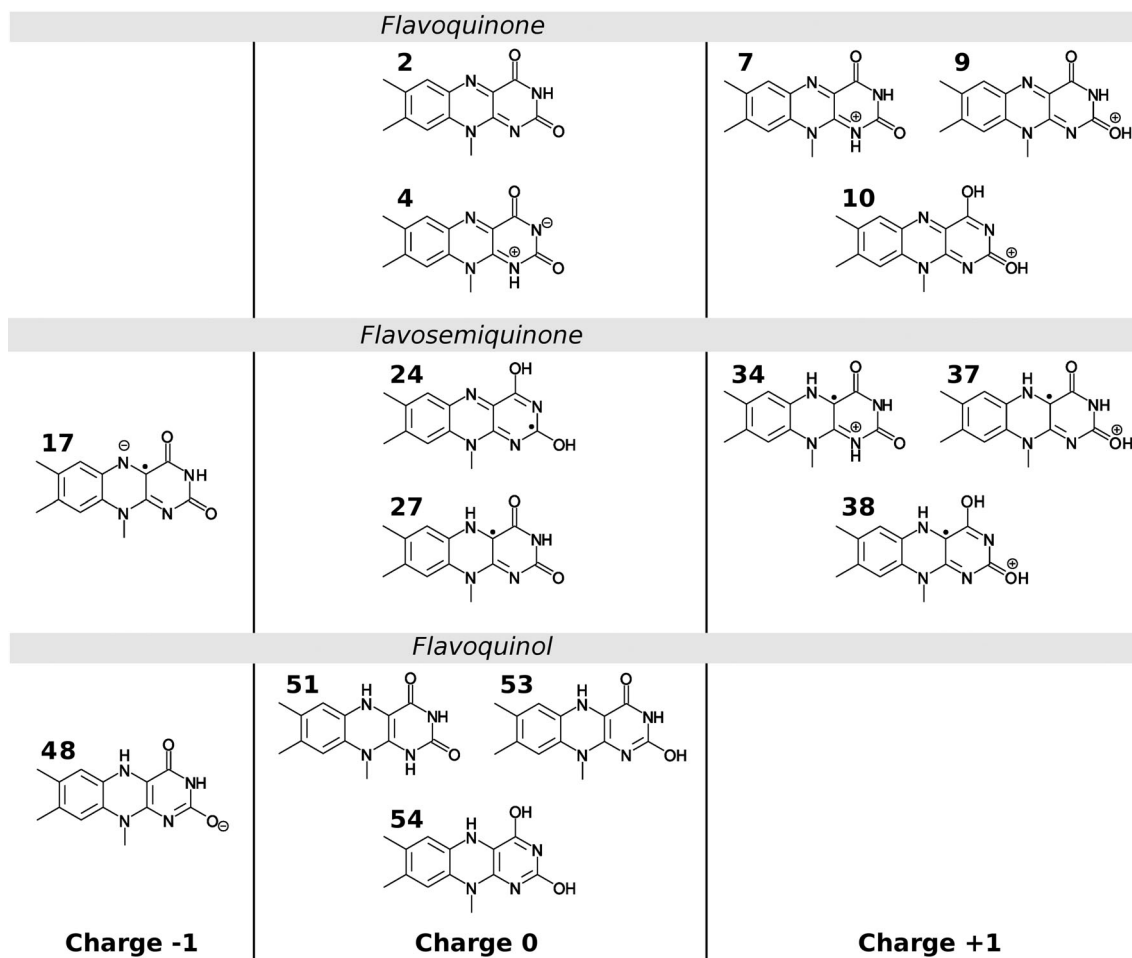


FIGURE 2 Structural formula of the main lumiflavin tautomers discussed here with their respective numbering, redox and charge state

Redox flavoproteins participate in single-electron transfer reactions,^[4] in oxygen-dependent oxidation,^[5] and in proton-coupled electron transfers (PCET) with varying number of electrons.^[2,6,7] Flavin chromophores are found in blue-light protein receptors^[3] and in photo-dependent enzymes.^[8] Recently, the biological role of flavins as general acid-base catalysts has also been demonstrated.^[9,10]

This versatility of flavoprotein function results from their ability to assume different redox, protonation and electronic states. Flavins are found in fully oxidized flavoquinone, 1-electron reduced flavosemiquinone radical, and two-electron reduced flavoquinol forms (Figure 2).^[1] Charge states with anionic, neutral, zwitterionic or cationic character are also possible, with their relative stability depending on the microenvironment and protein scaffold.^[2,11]

An additional layer of variability concerns flavin tautomerism.^[12–18] N3, N5, and N1 are canonical protonation sites and are often found protonated in flavoproteins under neutral pH.^[1] However, alternative tautomers with protonations in O2 and O4 were already observed.^[9,18,19] In flavosemiquinones, spin delocalization allows even carbons C4_a and C10_a as possible protonation sites.^[16]

Computational methods have been used to study flavins for more than four decades.^[20–23] While early works were based on semiempirical models, the first ab initio and density functional theory (DFT)

studies of flavins were performed in the late 1990s^[12–15] followed by a number of computational investigations (Table S1).^[16–18,24–56] The vast majority were based on DFT for ground-state properties^[14,16–18,24–36] or time-dependent DFT (TD-DFT) for excited-state calculations,^[36–50] mostly focused on the neutral flavoquinone form using lumiflavin as a model system. High-level ab initio coupled-cluster EOM-CCSD^[56] and approximate CC2^[55] were used recently to calculate vertical ionization potentials and flavin excited states, respectively. Multireference CASPT2 calculations were also reported^[51] and even relativistic contributions were obtained with 4-component TD-DFT.^[54]

However, none of these studies addressed systematically the approximations employed in calculations for flavins of various forms. Here, we provide a benchmark evaluation for ground-state properties of isolated lumiflavin in all possible oxidation and protonation states (a total of 59 species, Figures S1, S2, and S3). After description of the computational methods, we first investigate the effects of static and dynamic electron correlation. The coupled-cluster level is found as an appropriate reference after validation in comparison to experimental structures, electron affinities, and gas-phase basicities. Then, we analyze in detail the performance of various electronic-structure methods, including a ladder DFT functionals, for the calculation of

molecular-structural, energetic, electrical, and chemical properties. We note that alternative protonation sites may be intrinsically stabilized and challenge the canonical assignment of tautomeric equilibria in flavins.

2 | COMPUTATIONAL METHODS

2.1 | Geometry optimization

All flavins studied here had their geometry optimized with the B3LYP functional^[57,58] and def2-TZVP^[59] basis set. Optimizations of closed-shell flavins were performed with restricted Kohn-Sham formalism and flavosemiquinones with an open shell were optimized with unrestricted calculations. Spin contamination in the latter was small and the average deviation from the expected value ($\langle S^2 \rangle$) was 0.013, with the largest being 0.023 for **36**. All calculations were conducted with ORCA 4.1.1,^[60,61] using tight SCF convergence criteria and increased integration grids. All optimized structures (lumiflavins **1–59**) are available online.^[62]

The DFT accuracy in geometry optimizations (Section 3.2) was evaluated for functionals BLYP,^[63,64] PBE,^[65] M06-L,^[66] PBE0,^[67] and M06-2X.^[68] Among ab initio wave-function methods, coupled cluster CCSD(T) with the domain-based local-pair natural orbital (DLPNO) approximation^[69] and the augmented basis set aug-cc-pVTZ^[70] was used to optimize the geometry of the isoalloxazine ring in comparison to DFT and experimental geometries. The aug-cc-pVTZ/C^[71] auxiliary basis set was employed, with tight DLPNO thresholds. Also, the DLPNO-CCSD(T) geometry optimization was considered converged when root-mean-square gradient was less than $0.0015 \text{ E}_h \text{ a}_0^{-1}$ and the maximum component of the gradient was less than $0.0050 \text{ E}_h \text{ a}$. This looser convergence criteria had to be employed here due to the computational demands of the method.

Flavins can bend when N5 and N10 move above the isoalloxazine ring plane (Figure 3). This “butterfly” motion is observed mainly in flavoquinols and is relevant for flavin reactivity and dynamics. The associated torsion energy E_{tor} varies with the flavin redox and protonation state. Here, E_{tor} is defined as the difference in electronic potential energy between flavin in its planar and bent optimized conformations.

Geometries with isoalloxazine ring torsion were generated by constrained optimizations with B3LYP/def2-TZVP. The dihedral angles between atoms C4–N5–N10–C9 and N1–N10–N5–C6 (Figure 1) were restrained. Stable geometries for flavoquinols show both dihedrals close to 150° (Figure 3). In flavoquinones and flavosemiquinones, these dihedrals were constrained to 150.0° to generate bent conformations, and in flavoquinols they were constrained to 180.0° to obtain planar geometries.

2.2 | Model chemistry

Wave-function single-reference calculations with the MP2, CCSD, and CCSD(T) methods^[72] employed Dunning's correlation-consistent

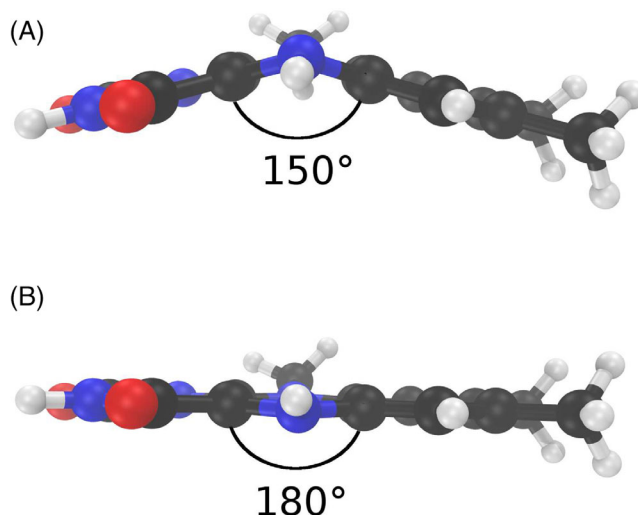


FIGURE 3 Planarity of the isoalloxazine ring. (A) Bent conformation found in stable reduced flavoquinol and (B) planar conformation found in stable oxidized flavoquinone

basis-set families, cc-pVnZ^[73] and aug-cc-pVnZ,^[70] with $n = \text{D, T, Q}$. The frozen-core approximation was employed in all post-Hartree-Fock calculations. To avoid spin contamination in flavosemiquinones, restricted open-shell Hartree-Fock was used as the zeroth-order reference wave function. DLPNO-CCSD and DLPNO-CCSD(T) approximations^[69] were also tested. For flavosemiquinones, unrestricted Hartree-Fock was performed first, and quasi-restricted orbitals were generated prior to the coupled-cluster excitations. This is the default approach employed by ORCA. The DLPNO approximation required the use of auxiliary cc-pVnZ/C or aug-cc-pVnZ/C^[71] basis sets. Tight DLPNO thresholds (“TightPNO” keyword in ORCA) were employed.^[74]

Multiconfigurational wave functions (CASSCF)^[72] were built with all isoalloxazine π orbitals in the active space, that is, a CAS(14,14) for flavoquinones or a CAS(15,14) for flavosemiquinones. The initial set of molecular orbitals were MP2 natural orbitals and the composition of the active space was checked before and after convergence. The contribution of dynamical correlation was calculated with the partially contracted version of NEVPT2.^[75] The frozen-core approximation was employed with all electrons occupying internal orbitals 1 s removed from the perturbation treatment.

For DFT calculations, the minimally augmented version of Karlsruhe def2 basis-set family,^[59] ma-TZVP,^[76] was employed. This basis set is efficient and accurate in DFT calculations. The usage of an augmented version is due to the presence of anionic flavins.

Seventeen functionals were tested: BLYP,^[63,64] OLYP,^[64,77] PBE,^[65] BPBE,^[63,65] OPBE,^[65,77] TPSS,^[78] M06-L,^[66] B3LYP,^[57,58] B3LYP + D3,^[79,80] PBE0,^[67] O3LYP,^[81] M06,^[68] M06-2X,^[68] CAM-B3LYP,^[82] LC-BLYP,^[83] ω B97X,^[84] and B2PLYP.^[85]

The resolution of identity (RI) was employed with auxiliary basis def2/J^[86] to approximate the Coulomb integrals. For hybrid functionals, chain-of-spheres approximations (COSX)^[87,88] was also used in the calculation of the Hartree-Fock exchange. For

flavosemiquinones, unrestricted Kohn-Sham was used. All calculations were performed with tight SCF convergence criteria and increased integration grids.

Zero-point vibrational energies (ZPV E) were calculated for stationary geometries at the B3LYP/def2-TZVP level without empirical scaling factors. Entropic and enthalpic contributions were calculated at 298.15 K using the quasi rigid-rotor-harmonic-oscillator approximation (QRRHO).^[89,90] No imaginary frequencies were observed for all optimized flavins, indicating they were true minima.

The conductor-like polarizable continuum model (CPCM)^[91] was used as an implicit-solvent model. Water solvation was mimicked with a dielectric constant and a refractive index set to 80.4 and 1.33, respectively. Nevertheless, we refrain from extensively including environmental effects and calculating condensed-phase properties here. These effects represent additional modeling difficulties, which may be tackled in future studies.

2.3 | Basis-set extrapolation

Wave-function calculations are particularly sensitive to basis-set incompleteness. Thus, the complete basis-set limit (CBS) was approximated by two-point extrapolations. For single-point energies, the mean-field SCF energy $E_{\text{SCF}}^{\text{CBS}}$ and the correlation energy $E_{\text{corr}}^{\text{CBS}}$ are extrapolated separately, following Equations (1)^[92] and (2),^[93,94] respectively. Electric dipole moments μ^{CBS} are extrapolated using Equation (3)^[95]:

$$E_{\text{SCF}}^{(X)} = E_{\text{SCF}}^{\text{CBS}} + Ae^{-\alpha\sqrt{X}}, \quad (1)$$

$$E_{\text{corr}}^{\text{CBS}} = \frac{X^\beta E_{\text{corr}}^{(X)} - Y^\beta E_{\text{corr}}^{(Y)}}{X^\beta - Y^\beta}, \quad (2)$$

$$|\mu|^{\text{CBS}} = \frac{X^3 |\mu|^{(X)} - Y^3 |\mu|^{(Y)}}{X^3 - Y^3}, \quad (3)$$

where X and Y are the basis-set cardinal numbers: 2 for double-zeta (D), 3 for triple-zeta (T) and 4 for quadruple-zeta (Q) extrapolations. For cc-pV[D/T]Z, $\alpha = 4.42$ and $\beta = 2.46$; for aug-cc-pV[D/T]Z, $\alpha = 4.30$ and $\beta = 2.51$ and for aug-cc-pV[T/Q]Z, $\alpha = 5.79$ and $\beta = 3.05$.^[96] In Equation (1), A is a constant determined by the two-point extrapolation.

2.4 | Molecular properties

The electric dipole moment (μ), a first-order electrical property, was obtained from DFT calculations directly from the SCF solution. For MP2 and B2PLYP methods, the relaxed electron density was calculated to account for orbital relaxation effects. For the calculation of dipole moments with DLPNO-CCSD, the parameter T_{CutPNO} was tightened to 10^{-8} because the use of DLPNO default truncation parameters would lead to inaccurate values.^[97] All dipoles were calculated at

equilibrium geometries, and the dipole origin was set to the center of mass.

The electron affinity (EA) of a molecule is defined as the negative of the energy difference of ground vibrational/rotational states before (M) and after ($M^{\bullet-}$) electron attachment. This process may induce a geometry change in the molecule. Electron affinities obtained from the energy difference of equilibrium geometries and associated ZPV E contributions are called adiabatic (EA_{adiab}):

$$EA_{\text{adiab}} = (E_{\text{M}} + \text{ZPVE}_{\text{M}}) - (E_{\text{M}^{\bullet-}} + \text{ZPVE}_{\text{M}^{\bullet-}}), \quad (4)$$

where E_{M} and $E_{\text{M}^{\bullet-}}$ are single-point energies for optimized geometries of the oxidized and reduced species. However, experimental measurements probe a different, thermalized condition and often consider EA (here called EA_{therm}) as the negative of the enthalpy of the reaction $\text{M} + e^- \rightleftharpoons \text{M}^{\bullet-}$ ^[34,98]:

$$EA_{\text{therm}} = H_{\text{M}} + H_{e^-} - H_{\text{M}^{\bullet-}} \quad (5)$$

where H is the enthalpy of formation and $H_{e^-} \equiv 0$.^[98] Here, EA_{therm} was calculated by including enthalpic contributions obtained with the QRRHO approximation at the B3LYP/def2-TZVP level.

The gas-phase basicity (GB) of a molecule M at temperature T is the negative of the Gibbs free energy for the protonation reaction $\text{M} + \text{H}^+ \rightleftharpoons \text{MH}^+$. Flavins have multiple protonation sites and the calculated GB was determined by using a population-weighted average (Equation (6)) with each tautomer MH_i^+ having a weight w_i given by its canonical probability (Equation (7)).^[99] For example, if M is **2**, then $\text{MH}^+ = \{\mathbf{7}, \mathbf{9}, \mathbf{12}, \mathbf{13}\}$, and the GB of M is an average over contributions from this set of flavins:

$$\text{GB}(T) = - \sum_i w_i [G(\text{MH}_i^+) - G(M) - G(\text{H}^+)], \quad (6)$$

$$w_i = \frac{e^{-\Delta G_i/k_B T}}{\sum_i e^{-\Delta G_i/k_B T}}, \quad (7)$$

where k_B is the Boltzmann constant and ΔG_i the relative free energy between tautomer MH_i^+ and the tautomer with the lowest free energy.

To calculate GB with different model chemistries, entropic and enthalpic contributions were obtained with B3LYP/def2-TZVP except for the electronic contribution, which was replaced by the energy calculated with the method of choice. The Gibbs free energy of the proton was considered to be $-0.010 E_h$, the free energy of a monoatomic ideal gas at 298.15 K.

3 | RESULTS AND DISCUSSION

3.1 | Electron correlation and reference model chemistry

Static electron correlation is important for the description of flavins in excited states,^[51] but it is unclear whether it also plays a role in

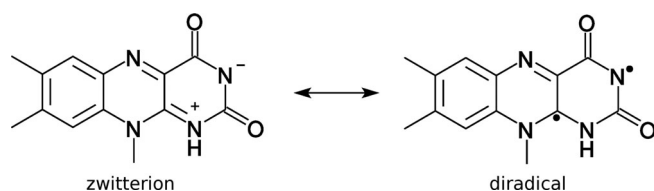


FIGURE 4 Resonance structures of **4**. When N1 (see Figure 1 for atom-numbering) is protonated, zwitterion or diradical (note unpaired electrons in C10_a and N3) resonance structures are possible

ground-state properties. We first applied the τ_1 diagnostics^[100] to all flavins studied here (Tables S2, S3, and S4) and checked that they show τ_1 values smaller than 0.020, indicating negligible static correlation. The only borderline case was **17** in bent conformation with τ_1 of 0.020 (Table S3) and discussed in more detail below.

Multiconfigurational calculations in selected flavosemiquinones and zwitterionic flavoquinones were conducted. The unpaired electron in the former may delocalize over the π -conjugated system, while the latter have competing resonance structures (Figure 4), possibly amounting for sizable multiconfigurational effects.

In Table 1, it is shown for semiquinones **17**, **27**, and **37** that the total energy difference between single-reference MP2 and multireference NEVPT2 is less than 1 mE_h, even for the bent **17** which displayed the borderline τ_1 value mentioned above, indicating that static correlation is not important. The correlation energy recovered in CASSCF is mainly dynamic correlation obtained by the full-CI treatment within the active space. The main electronic configuration in the CASSCF wave function for the three planar flavosemiquinones has a weight of 0.81 while all other configurations have weights less than 0.02. This configuration has only one single-occupied molecular orbital (MO), spatially delocalized over the isoalloxazine ring. The same configuration is observed in the Hartree-Fock wave function, and thus the effect of spin delocalization is properly accounted with the mean-field treatment.

For zwitterionic **4**, a state-average (SA) CASSCF calculation with two roots (50:50 weights) was performed to account for both resonance structures (Figure 4). Analysis of the obtained localized orbitals shows that the first (R1) and second (R2) roots correspond respectively to zwitterionic and diradical states. For natural orbitals, the principal configuration of root R1 contributes with 78% of the total CI expansion, while all remaining configurations have weights lower than 2%. For the root R2, the two principal configurations contribute with 48% and 20% of the total CI expansion. This is an indication of multiconfigurational character for the excited diradical state. It suggests that appropriate multiconfigurational methods should be used to calculate properties and reactivity which may involve flavin diradical states.

In Table 2, it is shown that the NEVPT2 energy for R1 is 3 mE_h smaller than the single-reference MP2 solution with a zwitterionic resonance structure. Thus, the multiconfigurational contribution for the ground-state R1 is small and the energetics is described well by a single-reference method.

Additionally, the energy difference found between the two roots for **4** (R1 and R2, 0.11 E_h with NEVPT2 in Table 2) exemplify that excited electronic states in flavins have considerably high energies and their nonadiabatic crossings may be safely neglected for calculation of ground-state properties.

In the condensed phase, however, dielectric and specific contacts may stabilize zwitterionic flavins and increase the importance of their resonance states and the associated multiconfigurational character. For instance, although the MP2/cc-pVDZ energy for the isolated neutral tautomer **2** is 54 mE_h lower than **4**, solvation in a polar dielectric (estimated by the CPCM model for water)^[91,101] stabilizes preferentially the zwitterion and the energy difference drops to 29 mE_h, still in favor of the neutral **2**. Yet, specific interactions in enzymes may stabilize N1 protonation (Figure 2), shifting even further this energy difference and the associated tautomeric equilibria (**2** \rightleftharpoons **4**).

Results shown above indicate that a single-reference electronic-structure method should be sufficient for the description of ground-state properties and reactions of flavins studied here. Dynamic electron correlation should also be recovered. Applying the gold-standard single-reference CCSD(T) method to flavins is only computationally feasible using a double-zeta basis set, which compromises the amount of correlation recovered. Thus, we first compared the full CCSD(T) within this limited basis set and the DLPNO local-pair approximation,^[69] which provides linear scaling with system size and allows coupled-cluster calculations of flavins with a larger basis.

In Table 3, both CCSD(T) and DLPNO-CCSD(T) calculations employed quasi-restricted orbitals. The error due to the DLPNO approximation is significant in the triples correction for two representative flavins and reaches 11 mE_h for **17** even using tight calculation thresholds (TightPNO keyword). This error is much higher than often observed in smaller molecules^[74] and similar to deviations found for transition metal complexes.^[102] For the flavins, the error may be attributed to the delocalized and resonant nature of the electron distribution in the tricyclic isoalloxazine ring. Nevertheless, the error in relative energies due to the DLPNO approximation decreases to 1.3 mE_h, as in the reaction **2** \rightleftharpoons **17** + e⁻. This is the expected upper limit in the accuracy of relative energies for flavin reactions obtained here from DLPNO-CCSD(T) calculations with larger (triple-zeta and extrapolated) basis sets. Nevertheless, the DLPNO-CCSD(T) method was chosen here as the reference model chemistry.

3.2 | Accuracy of calculated molecular structure, dipole, and torsion energy

The calculated equilibrium geometry of isoalloxazine ring may be assessed by comparison with solid-state crystallography of related flavins. The structure for 3-methyl-lumiflavin has been determined by X-ray diffraction^[103] and Table 4 shows a comparison of bond lengths with the gas-phase optimized geometry of isoalloxazine. The uncertainty in bond lengths of the crystal structure is 0.003 Å, so there is no significant difference between calculated and experimental bond lengths for C=C, C=N, and C=O. The difference of 0.014 and

Method	17	17(bent)	27	37
HF	−866.938886	−866.929755	−867.484272	−867.881032
CASSCF ^a	−867.082975	−867.073693	−867.635744	−868.026266
MP2	−869.654678	−869.644219	−870.196964	−870.595394
NEVPT2 ^a	−869.653852	−869.645101	−870.197265	−870.595219

^aThe active space (15 e[−] in 14 MO) contains the complete resonant π -system.

TABLE 1 Electronic energy (in E_h) obtained for **17**, **27**, and **37** with the cc-pVDZ basis set

TABLE 2 Electronic energy (in E_h) obtained for zwitterionic **4** with the cc-pVDZ basis set

Method	Total energy
SA-CASSCF R1 ^a	−866.993155
SA-CASSCF R2 ^b	−866.908477
NEVPT2 R1	−869.562887
NEVPT2 R2	−869.449373
MP2	−869.559812

^aFirst root with zwitterionic character.

^bSecond root with diradical character.

TABLE 3 Electron correlation energies (E_{Method}−E_{HF}, in E_h) for **2** and the reduced flavosemiquinone **17**

Method	2	17
CCSD/cc-pVDZ	−2.772852	−2.796138
DLPNO-CCSD/cc-pVDZ	−2.772857	−2.793963
CCSD(T)/cc-pVDZ	−2.893783	−2.916723
DLPNO-CCSD(T)/cc-pVDZ	−2.883622	−2.905234

TABLE 4 Average lengths (in Å) for different bond types in isoalloxazine ring

Bond	Experiment ^a	Calculated ^b	Δ
C=C	1.396	1.398	0.002
C=N	1.302	1.298	0.004
C=O	1.213	1.212	0.001
C–N	1.375	1.383	0.008
C–C	1.465	1.479	0.014

^a3-methyl-lumiflavin crystal structure.^[103]

^bDLPNO-CCSD(T)/aug-cc-pVTZ geometry for isoalloxazine.

0.008 Å observed for the more flexible C–C and C–N bonds may be attributed to packing effects in the crystal environment, which may shorten the bond lengths. In conclusion, the similarity between calculated and experimental bond lengths corroborates that DLPNO-CCSD(T)/aug-cc-pVTZ provides accurate geometries.

Table 5 shows mean unsigned errors (MUE) and maximum absolute errors (MAX) in bond lengths and angles for selected DFT functionals tested in comparison to reference geometries optimized at the DLPNO-CCSD(T)/aug-cc-pVTZ level for the isoalloxazine molecule (Figure 1). B3LYP gives the smallest errors regarding bond lengths

TABLE 6 DLPNO-CCSD(T)/aug-cc-pVTZ energy (in E_h) for **2** and **17** with geometries optimized by different DFT functionals

Functional	E ₂	E ₁₇	ΔE
PBE	−870.70018	−870.76169	0.06151
M06-L	−870.69987	−870.76142	0.06155
B3LYP	−870.70076	−870.76216	0.06140

TABLE 5 Mean unsigned errors (MUE) and maximum absolute errors (MAX) of isoalloxazine bond lengths and angles obtained from DFT optimizations

Functional	Length (Å)		Angle (°)	
	MUE	MAX	MUE	MAX
BLYP	0.009	0.020	0.32	1.12
PBE	0.007	0.016	0.27	0.78
M06-L	0.005	0.012	0.32	0.93
B3LYP	0.003	0.0103	0.35	1.38
PBE0	0.006	0.014	0.29	1.08
M06-2X	0.006	0.016	0.45	1.64

Note: Reference values comes from DLPNO-CCSD(T)/aug-cc-pVTZ geometry.

(both MUE and MAX), while PBE gives the smallest errors in bond angles.

Table 6 shows total energies obtained with DLPNO-CCSD(T)/aug-cc-pVTZ after optimization of **2** and **17** with functionals PBE, M06-L, and B3LYP. The smallest energies for both flavins are found for the B3LYP geometry, indicating a better performance. But, differences between relative energies ΔE are negligible (<0.15 mE_h) suggesting that optimizations with any of these three functionals would result in geometries with equivalent quality.

It is unlikely that B3LYP performance is an artifact of error cancellation because of the favorable comparison to DLPNO-CCSD(T)/aug-cc-pVTZ geometries (Table 5). Thus, the B3LYP/def2-TZVP model was chosen for geometry optimizations of all equilibrium and bent structures studied here.

For flavoquinols **40**, **41**, and **56**, that show protonations in both N1 and O2/N3, pirimidization of N10 is more pronounced than in the other flavoquinols, and their methyl group remains completely below the isoalloxazine ring. This change in methyl position make N10 more negative, which stabilizes electrostatically the adjacent positive

TABLE 7 Magnitude of μ (in D) calculated for **2** with different methods

	MP2	DLPNO-CCSD
aug-cc-pVDZ	8.97	8.96
aug-cc-pVTZ	9.01	9.02
aug-cc-pV[D/T]Z	9.03	9.04

sites {N1, O2, N3} that are bound to two or three protons in **40**, **41**, and **56**.

Flavins may undergo a butterfly motion bending the isoalloxazine ring (Figure 3). When fully reduced, electronic repulsions favor the pirimidization of N5 and N10, and lead to ring bending.^[24] Thus, it is relevant to check the associated torsion energy (E_{tor}) for which DLPNO-CCSD(T)/aug-cc-pV[D/T]Z was used as reference.

It was suggested experimentally for several flavoquinol derivatives in solution that E_{tor} is lower than 20 kJ.mol⁻¹.^[104] For isolated flavoquinols calculated here, E_{tor} has a similar magnitude (Table S19). For flavoquinones and flavosemiquinones, experimental data show that these flavins assume planar or quasi-planar conformations,^[18,24,103] again in agreement with results here.

The electric dipole moment μ was evaluated here because this property is relevant to flavin interactions in condensed phase. Its calculation by coupled-cluster methods is very demanding, so the accuracy of μ obtained with MP2 was assessed first (Table 7). Regardless of the basis set used, MP2 and DLPNO-CCSD dipoles differ on their magnitudes by only 0.01 D for **2** and MP2/aug-cc-pV[D/T]Z dipoles were chosen as a reference here.

3.3 | Comparison with experimental gas-phase reactions

Experimental electron affinity and gas-phase basicity are available for **2**^[34] and are used here for comparison with various levels of theory and basis set. Tables 8 and S5 show an opposed trend for EA_{therm} increasing with the basis-set size while GB decreases, particularly in the DLPNO-CCSD(T) level.

Employing augmented basis sets is essential for a correct description of the diffuse anionic product (**17**). For instance, the cc-pVTZ set gives less accurate EA than aug-cc-pVDZ for all methods tested. Augmented sets also give a better description for the GB calculation. MP2 calculations employing any CBS extrapolation provides results in agreement with both experimental properties within their uncertainties, probably due to fortuitous error cancellation. For more reliable coupled-cluster methods, the DLPNO-CCSD(T) level with the aug-cc-pVQZ set or with the extrapolated CBS limit (either [T/Q]Z or [D/T]Z) provide excellent results.

It should also be noted that EA_{adiab} is systematically smaller than EA_{therm} for all calculation levels (Table S6). The latter was used for comparisons here because it is a better approximation to EA measured experimentally in thermalized conditions. However, it is unclear which

TABLE 8 Electron affinity (EA, in eV) and gas-phase basicity (GB, in kJ.mol⁻¹) for **2**

Method	EA_{therm}	GB
MP2		
cc-pV[D/T]Z	1.88	926
aug-cc-pV[D/T]Z	1.93	923
aug-cc-pV[T/Q]Z	1.94	924
DLPNO-CCSD		
cc-pV[D/T]Z	1.77	938
aug-cc-pV[D/T]Z	1.81	935
aug-cc-pV[T/Q]Z	1.80	932
DLPNO-CCSD(T)		
cc-pVDZ	1.21	950
cc-pVTZ	1.59	940
cc-pV[D/T]Z	1.78	934
aug-cc-pVDZ	1.75	925
aug-cc-pVTZ	1.79	929
aug-cc-pVQZ	1.80	928
aug-cc-pV[D/T]Z	1.82	930
aug-cc-pV[T/Q]Z	1.81	928
Experimental ^[34]	1.86 ± 0.12	919 ± 9

temperature the GB experiments were conducted at.^[34] Here $T = 298.15$ K is assumed, but this may be an additional source of uncertainty for our calculated values in relation to the experiment.

By neglecting thermal and entropic contributions, EA and GB calculated with DLPNO-CCSD(T)/aug-cc-pV[T/Q]Z would yield 1.77 eV and 954 kJ.mol⁻¹, respectively. While EA would still agree with experiment within uncertainty, GB would disagree in 26 kJ.mol⁻¹. By also neglecting ZPV E corrections, EA and GB would become 1.69 eV and 988 kJ.mol⁻¹, largely degrading the calculated accuracy. Therefore, including zero-point vibrational (ZPV E) and thermal contributions is essential for quantitative agreement with experiment.

Due to the favorable comparisons shown in this section, reasonable computational cost and the more systematic behavior of coupled-cluster methods in recovering electron correlation, the DLPNO-CCSD(T)/aug-cc-pV[D/T]Z level was chosen here to produce reference energetic values for other flavin tautomers. Although other approximations are evoked here (harmonic vibrational modes, frozen-core energies, etc.), it may be concluded that electron and proton affinities for isolated flavins are well described with this level of theory and B3LYP/def2-TZVP geometry optimizations.

3.4 | Tautomer thermochemistry

As shown in Figure 2, flavins with given redox state and total charge may assume different tautomeric forms. For example, three tautomers (**6**, **7**, and **9**) are found for a cationic flavoquinone, by protonating two atoms in the set {N1, N3, O2}. Table 9 shows the calculated free

TABLE 9 Calculated^a relative free energies ($\Delta G[298.15\text{ K}]$, in $\text{kJ}\cdot\text{mol}^{-1}$) and equilibrium composition (in %) of isolated flavins with multiple tautomeric states

	ΔG	Composition
Flavoquinone neutral		
2	0.0	100.00
4	120.8	0.00
Flavoquinone cation		
7	11.2	1.07
9	0.0	98.82
10	17.0	0.11
Flavosemiquinone neutral		
24	13.2	0.48
27	0.0	99.52
Flavosemiquinone cation		
34	9.1	2.47
37	0.0	97.51
38	20.7	0.02
Flavoquinol neutral		
51	0.0	99.66
53	16.0	0.15
54	15.6	0.19

^aDLPNO-CCSD(T)/aug-cc-pv[D/T]Z//B3LYP/def2-TZVP.

energy and equilibrium composition at 298.15 K for possible tautomers.

For flavoquinones, the most stable tautomers are **2** for neutral and **9** for cationic forms, in accordance with previous qualitative predictions.^[12,14,18] However, the current results highlight that tautomers such as **10** are more stable than previously proposed. For instance, Dopfer et al.^[18] obtained $\Delta G = 24.2\text{ kJ}\cdot\text{mol}^{-1}$ for **10** using B3LYP/cc-pVDZ calculations. Furthermore, Ridge et al.^[34] performed B3LYP/6-31G* and M06-L/6-31G* calculations and suggested that **7** was more stable among the cationic forms. This disagreement with results reported here (Tables 9 and S7) may be caused by the geometry found by Ridge et al. after optimization which shows the dihedral N1-C2-O2-H rotated by 180° in comparison to the structures found here and previously.^[18]

For flavosemiquinone, the most stable tautomer are **27** for neutral and **37** for cationic forms, again in qualitative agreement with previous calculations.^[12,14,16] The relative energy of **34** was predicted to be 5.4 and 4.1 $\text{kJ}\cdot\text{mol}^{-1}$ higher than **37** by earlier calculations with HF/6-31G** and B3LYP/6-31G* methods, respectively.^[12,14] These are potential energy values, but are similar to the free energy of 9.1 $\text{kJ}\cdot\text{mol}^{-1}$ reported here.

The work of Zheng and Ornstein^[13] with HF/6-31G* calculations suggested that **27** was not the most stable neutral flavosemiquinone. Their conclusion disagrees qualitatively with the results shown here and in earlier publications.^[14,16] We suspect that the approximate

nature of the unrestricted HF calculation performed^[13] lead to the wrong conclusion.

Hadad and Platz et al.^[16] tested stabilities of alternative tautomers of neutral flavosemiquinones. They considered not only oxygens and nitrogens as possible protonation sites (See Figure S2), but also the carbons C4_a and C10_a. Their B3LYP/6-31+G** calculations added by ZPV E agrees with ours and predict **27** as the most stable tautomer. Protonation at C4_a and C10_a resulted in tautomers 105.9 and 213.8 $\text{kJ}\cdot\text{mol}^{-1}$ more energetic than **27**, respectively.^[16] This large energy difference indicates that flavin radicals with protonation at carbon will be unstable and irrelevant to tautomeric thermochemistry. Nevertheless, these species may still play a role as alternative reaction mechanisms and transient intermediates in flavoenzymes.

For flavoquinols, **51** is the most stable in the neutral form. Meyer et al. in an early study reported heats of formation calculated with the semiempirical method PM3 for different flavoquinol tautomers.^[12] They found **51** as the most stable, **54** was 45.2 $\text{kJ}\cdot\text{mol}^{-1}$ more energetic and **53** was 25.9 $\text{kJ}\cdot\text{mol}^{-1}$ more energetic. These results disagree with the energies of 15.6 and 16.0 $\text{kJ}\cdot\text{mol}^{-1}$ reported here respectively, showing that the PM3 method (and possibly other semiempirical methods based in the NDDO approximation)^[105] is inadequate to predict tautomeric equilibria of flavins.

Thermochemical analysis at normal temperature may indicate if more than one tautomer can exist at equilibrium. For anionic flavins and neutral flavoquinone, there is a significantly lower energetic tautomer, comprising 100% of the ensemble: **17** is the tautomeric form adopted by anionic flavosemiquinone, **48** for anionic flavoquinol and **2** for neutral flavoquinone (See Tables S7, S8, and S9). In gas phase, only one tautomer needs to be accounted for when calculating thermodynamic properties for these molecules. Alternatively, cationic flavins, neutral flavosemiquinone, and neutral flavoquinol have more than one tautomer significantly populated (Table 9).

Only nitrogen sites (N1, N3, or N5) are protonated in the most stable tautomer of each redox and charge, except for the cationic flavins, **9** and **37**, which have O2 protonated. This unusual oxygen protonation is also found in **10**, **24**, **38**, **53**, and **54**, which are within 20 $\text{kJ}\cdot\text{mol}^{-1}$ of the most stable tautomer in their respective redox and charge state. For instance, flavoquinol **51** has all three nitrogens protonated and the **54** tautomer, with O2 and O4 protonated, is 15.6 $\text{kJ}\cdot\text{mol}^{-1}$ more energetic. Other naturally occurring quinols such as ubiquinol and plastoquinol always have their oxygens protonated.^[106,107] But the resonance structure in flavins results in stabilization of nitrogen protonation and (flavo)quinols without a phenol group.

The isolated GB of stable anionic flavosemiquinone **17** and stable anionic flavoquinol **48** are almost identical, 1336 and 1335 $\text{kJ}\cdot\text{mol}^{-1}$, respectively (Tables S15 and S17). However, in aqueous solution, the flavosemiquinone is significantly more basic than the flavoquinol, with $\text{pK}_a \approx 8$ and 6, respectively for their conjugated acids.^[1] For the stable neutral flavoquinone **2** and stable neutral flavosemiquinone **27**, although their GB = 963 and 930 $\text{kJ}\cdot\text{mol}^{-1}$, respectively (Tables S14 and S16), their cationic conjugated acids in aqueous solution exist only under very low pH.^[1]

TABLE 10 Mean unsigned errors (MUE) and maximum absolute errors (MAX) of properties calculated by DFT functionals for lumiflavin forms studied here

Functional	E_{adiab} (eV)		GB (kJ.mol ⁻¹)		$ \mu $ (D)		E_{tor} (kJ.mol ⁻¹)	
	MUE	MAX	MUE	MAX	MUE	MAX	MUE	MAX
<i>GGA and meta-GGA</i>								
BLYP	0.30	0.74	13	34	0.47	1.45	3.3	6.2
OLYP	0.35	0.78	9	25	0.50	1.53	3.7	6.6
PBE	0.20	0.61	13	35	0.49	1.48	3.5	7.0
BPBE	0.22	0.65	9	25	0.50	1.48	3.5	7.0
OPBE	0.26	0.68	13	30	0.53	1.57	3.8	6.8
TPSS	0.25	0.65	8	19	0.50	1.49	3.0	6.7
M06-L	0.35	0.73	12	22	0.59	1.52	3.0	8.5
<i>Hybrid</i>								
B3LYP	0.22	0.53	5	14	0.48	1.32	2.3	7.2
B3LYP + D3	0.22	0.52	6	17	0.48	1.32	2.2	6.7
PBE0	0.22	0.44	5	13	0.49	1.33	2.3	7.9
O3LYP	0.21	0.56	6	21	0.46	1.36	3.1	8.1
M06	0.17	0.38	4	13	0.49	1.32	2.6	7.2
M06-2X	0.14	0.23	7	12	0.48	1.31	1.4	5.6
<i>Range-separated</i>								
CAM-B3LYP	0.22	0.31	4	12	0.53	1.25	2.1	7.0
LC-BLYP	0.20	0.33	22	32	0.54	1.27	2.0	6.6
ω B97X	0.23	0.37	7	15	0.55	1.34	1.5	6.1
<i>Double-hybrid</i>								
B2PLYP	0.17	0.38	8	18	0.44	1.43	1.6	5.0

Note: Reference values were obtained with DLPNO-CCSD(T) for energies and MP2 for dipoles, both with the aug-cc-pV[D/T]Z basis extrapolation.

This comparison reasserts that the gas-phase basicity and electron affinity cannot be used directly as probes of condensed-phase basicity and redox potentials, as solvation modulates the ability to bind protons and electrons slightly differently for each flavin form. Nevertheless, the intrinsic tautomeric equilibria of isolated flavins described here builds the foundation for modeling their condensed-phase behavior by including the influence of microenvironments and light absorption as a perturbation of intrinsic equilibria.^[12,14,17] For instance, the **9** \rightleftharpoons **7** equilibrium was suggested previously^[12,14] and above (Section 3.1) to shift toward **7** formation in implicit solvent models. Explicit hydrogen bonds provided by flavoprotein scaffolds are expected to further affect flavin tautomerism.^[11,33,108]

3.5 | Performance of DFT for molecular properties

As DFT is often employed for ground-state calculations of flavins, an appraisal of the performance of selected functionals in the computation of their molecular properties is useful and shown in Table 10. Properties for each flavin species and functional are shown in Tables S10–S19. MUE and MAX were calculated over different sets for each property. For E_{adiab} , $N = 34$, comprising all flavoquinones and the neutral and cationic flavosemiquinones (**1–14** and **20–39**).

Flavoquinols and anionic flavosemiquinone were removed as they would require calculation of unstable three-electron reduced species or generate unstable dianions, respectively. For GB , $N = 30$, comprising all anionic flavins, neutral flavoquinones, and flavosemiquinones (**1–5**, **15–19**, and **40–49**). Calculations of GB for neutral flavoquinols would result in high-energy four-time protonated flavins. For μ and E_{tor} , $N = 14$, comprising the most stable tautomers in Figure 2 and excluding **4**.

The functional showing the lowest deviation varies with the calculated property. Overall, M06-2X shows the best performance, with the lowest MAX errors among GGA and hybrid functionals for all properties, and the lowest MUE for EA and E_{tor} properties among all functionals. Thus, M06-2X may be indicated for calculation of redox potentials of flavins. Surprisingly, this functional shows the highest MUE for GB among hybrid functionals. Alternatively, M06 and CAM-B3LYP perform particularly well for GB and dipole moments, so they may be indicated for the computation of tautomer thermochemistry in condensed phase, for instance using hybrid QC/MM potentials.^[109,110] Among computationally cheaper GGA functionals, TPSS, BPBE, and pure PBE show the best overall performance. However, all GGA functionals except OPBE make qualitatively wrong predictions on the relative stability of flavosemiquinone tautomers.

The popular B3LYP functional performs well for all properties (maybe except for EA), while inclusion of D3 dispersion correction does not improve its accuracy. The higher computational demand of B2PLYP calculations does not increase its performance significantly, which is similar to the cheaper functional M06.

E_{tor} is much easier to predict and all functionals tested have MUE below chemical accuracy. MAX errors always correspond to **17**, which bent geometry shows increased multiconfigurational character (Table S3 and Section 3.1). Calculated components for μ are similar among functionals, with the MUE of their magnitude ~ 0.5 D. However, the errors in open-shell flavins are significantly higher for all functionals. For instance, M06-2X shows MUE of 0.25 D for closed-shell and 0.84 D for open-shell flavins. Hence, calculations of flavosemiquinones using DFT for dipole moments and other response properties that depend on electronic polarization by the environment (such as QM/MM simulations in condensed phase) must be seen with caution.

4 | CONCLUSION

A systematic evaluation of approximations in the computation of the molecular electronic structure of isolated flavins was presented here. Multiconfigurational effects are negligible for ground-state properties, except for diradical species in resonance to zwitterionic forms. Including complete basis-set extrapolation and augmented functions in the calculations is necessary for agreement with experimental electron affinities and gas-phase basicities without resorting to error cancellation. The uncertainty expected here for DLPNO-CCSD(T) relative energies is 1.3 mE_h and for MP2 dipole moments it is 0.01 D, suggesting these methods can be used as references for flavin energetic and electrical properties, respectively. Nevertheless, it should be noted that the DLPNO approximation presented high errors for the triples correction to absolute energies, probably due to the delocalized electron distribution in the isoalloxazine ring.

For structural properties, DFT calculations with B3LYP or PBE functionals had similar quality. The popular B3LYP, most often used in previous computational studies of flavins, performs reasonably well for energetic properties but other functionals with similar computational cost are better recommended. For electron affinity calculations, Minnesota hybrid functional M06-2X should be used and for gas-phase basicity, M06 and CAM-B3LYP give the most accurate results. DFT should be carefully used when studying flavosemiquinones and higher level calculations like DLPNO-CCSD(T) are recommended. Pure- and meta-GGA functionals give qualitatively wrong gas-phase basicities and all functionals tested result in large errors for dipole moments of radical flavin forms.

The present analysis of tautomeric equilibria is qualitatively similar to previous studies,^[12,14,16] but quantitatively more accurate and embraces a larger and complete set of flavin oxidation and charge states. Tautomeric compositions described here corroborate that nitrogens sites are preferentially protonated, but alternative tautomers are possible and might be explored during catalysis in flavoenzymes.^[9] In particular, cationic flavins have unusual protonation at

oxygen. Finally, this benchmark study should guide future work in the calculation of redox potentials, pK_a , and other condensed-phase properties for flavins.

ACKNOWLEDGMENT

Funding from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, scholarship grant 2017/26109-0 to Felipe Curtolo and research grant 2019/21856-7 to Guilherme M. Arantes) is gratefully acknowledged.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Felipe Curtolo  <https://orcid.org/0000-0002-4459-0968>

Guilherme M. Arantes  <https://orcid.org/0000-0001-5356-7703>

REFERENCES

- [1] P. F. Heelis, *Chem. Soc. Rev.* **1982**, 11, 15.
- [2] A. M. Edwards, *R. Soc. Chem* **2006**, 6, 1.
- [3] J. M. Christie, L. Blackwood, J. Petersen, S. Sullivan, *Plant Cell Physiol.* **2014**, 56, 401.
- [4] M. A. Vanoni, *Open Biol.* **2021**, 11, 210010.
- [5] E. Romero, J. R. G. Castellanos, G. Gadda, M. W. Fraaije, A. Mattevi, *Chem. Rev.* **2018**, 118, 1742.
- [6] P. F. Fitzpatrick, *Bioorg. Chem.* **2004**, 32, 125.
- [7] F. Curtolo, G. M. Arantes, *J. Chem. Inf. Model.* **2020**, 60, 6282.
- [8] D. Zhong, *Annu. Rev. Phys. Chem.* **2015**, 66, 691.
- [9] H. Unno, S. Yamashita, Y. Ikeda, S. y. Sekiguchi, N. Yoshida, T. Yoshimura, M. Kusunoki, T. Nakayama, T. Nishino, H. Hemmi, *J. Biol. Chem.* **2009**, 284, 9160.
- [10] C. J. Thibodeaux, H.-W. Liu, *Arch. Biochem. Biophys.* **2017**, 632, 47.
- [11] R. K. Kar, A.-F. Miller, M.-A. Mrogiński, *WIREs Comput. Mol. Sci.* **2022**, 12, e1541.
- [12] M. Meyer, H. Hartwig, D. Schomburg, *THEOCHEM J. Mol. Struct.* **1996**, 364, 139.
- [13] Y.-L. Zheng, R. L. Ornstein, *J. Am. Chem. Soc.* **1996**, 118, 9402.
- [14] M. Meyer, *THEOCHEM J. Mol. Struct.* **1997**, 417, 163.
- [15] J. Wouters, F. Durant, B. Champagne, J.-M. André, *Int. J. Quantum Chem.* **1997**, 64, 721.
- [16] C. B. Martin, M.-L. Tsao, C. M. Hadad, M. S. Platz, *J. Am. Chem. Soc.* **2002**, 124, 7226.
- [17] S. Salzmann, C. M. Marian, *Chem. Phys. Lett.* **2008**, 463, 400.
- [18] J. Langer, A. Günther, S. Seidenbecher, G. Berden, J. Oomens, O. Dopfer, *ChemPhysChem* **2014**, 15, 2550.
- [19] D. Müller, O. Dopfer, *Phys. Chem. Chem. Phys.* **2020**, 22, 18328.
- [20] D. A. Dixon, D. L. Lindner, B. Branchaud, W. N. Lipscomb, *Biochemistry* **1979**, 18, 5770.
- [21] M. H. Palmer, I. Simpson, R. J. Platenkamp, *J. Mol. Struct.* **1980**, 66, 243.
- [22] L. H. Hall, B. J. Orchard, S. K. Tripathy, *Int. J. Quantum Chem.* **1987**, 31, 195.
- [23] L. H. Hall, B. J. Orchard, S. K. Tripathy, *Int. J. Quantum Chem.* **1987**, 31, 217.
- [24] J. Rodríguez-Otero, E. Martínez-Núñez, A. Peña-Gallego, S. A. Vázquez, *J. Org. Chem.* **2002**, 67, 6347.
- [25] J. D. Walsh, A.-F. Miller, *THEOCHEM J. Mol. Struct.* **2003**, 623, 185.
- [26] Y. Zheng, P. R. Carey, B. A. Palfey, *J. Raman Spectrosc.* **2004**, 35, 521.

- [27] M. Kondo, J. Nappa, K. L. Ronayne, A. L. Stelling, P. J. Tonge, S. R. Meech, *J. Phys. Chem. B* **2006**, *110*, 20107.
- [28] S. Bhattacharyya, M. T. Stankovich, D. G. Truhlar, J. Gao, *J. Phys. Chem. A* **2007**, *111*, 5729.
- [29] X.-L. Li, Y. Fu, *THEOCHEM J. Mol. Struct.* **2008**, *856*, 112.
- [30] M. M. N. Wolf, S. Schumann, R. Gross, T. Domratheva, R. Diller, *J. Phys. Chem. B* **2008**, *112*, 13424.
- [31] M. A. North, S. Bhattacharyya, D. G. Truhlar, *J. Phys. Chem. B* **2010**, *114*, 14907.
- [32] R.-K. Zhao, A. Lukacs, A. Haigney, R. Brust, G. M. Greetham, M. Towrie, P. J. Tonge, S. R. Meech, *Phys. Chem. Chem. Phys.* **2011**, *13*, 17642.
- [33] M. Kılıç, B. Ensing, *J. Chem. Theory Comput.* **2013**, *9*, 3889.
- [34] T. Zhang, K. Papson, R. Ochran, D. P. Ridge, *J. Phys. Chem. A* **2013**, *117*, 11136.
- [35] A. Aleksandrov, *J. Comput. Chem.* **2019**, *40*, 2834.
- [36] P. Mondal, K. Schwinn, M. Huix-Rotllant, *J. Photochem. Photobiol. A* **2020**, *387*, 112164.
- [37] C. Neiss, P. Saalfrank, M. Parac, S. Grimme, *J. Phys. Chem. A* **2003**, *107*, 140.
- [38] E. Sikorska, I. V. Khmelinskii, D. R. Worrall, J. Koput, M. Sikorski, *J. Fluoresc.* **2004**, *14*, 57.
- [39] E. Sikorska, I. V. Khmelinskii, W. Prukała, S. L. Williams, M. Patel, D. R. Worrall, J. L. Bourdelande, J. Koput, M. Sikorski, *J. Phys. Chem. A* **2004**, *108*, 1501.
- [40] E. Sikorska, I. V. Khmelinskii, J. Koput, M. Sikorski, *THEOCHEM J. Mol. Struct.* **2004**, *676*, 155.
- [41] E. Sikorska, I. V. Khmelinskii, J. Koput, J. L. Bourdelande, M. Sikorski, *J. Mol. Struct.* **2004**, *697*, 137.
- [42] E. Sikorska, J. R. Herance, J. L. Bourdelande, I. V. Khmelinskii, S. L. Williams, D. R. Worrall, G. Nowacka, A. Komasa, M. Sikorski, *J. Photochem. Photobiol. A* **2005**, *170*, 267.
- [43] E. Sikorska, I. Khmelinskii, A. Komasa, J. Koput, L. F. V. Ferreira, J. R. Herance, J. L. Bourdelande, S. L. Williams, D. R. Worrall, M. Insińska-Rak, M. Sikorski, *Chem. Phys.* **2005**, *314*, 239.
- [44] M. Kowalczyk, E. Sikorska, I. V. Khmelinskii, J. Komasa, M. Insińska-Rak, M. Sikorski, *THEOCHEM J. Mol. Struct.* **2005**, *756*, 47.
- [45] M. Insińska-Rak, E. Sikorska, J. L. Bourdelande, I. V. Khmelinskii, W. Prukała, K. Dobek, J. Karolczak, I. F. Machado, L. F. V. Ferreira, A. Komasa, D. R. Worrall, M. Sikorski, *J. Mol. Struct.* **2006**, *783*, 184.
- [46] Y.-K. Choe, S. Nagase, K. Nishimoto, *J. Comput. Chem.* **2007**, *28*, 727.
- [47] S. Salzmann, J. Tatchen, C. M. Marian, *J. Photochem. Photobiol. A* **2008**, *198*, 221.
- [48] M. Wu, L. A. Eriksson, *J. Phys. Chem. A* **2010**, *114*, 10234.
- [49] A. Vdovin, A. Slenczka, B. Dick, *Chem. Phys.* **2013**, *422*, 195.
- [50] B. Karasulu, W. Thiel, *J. Phys. Chem. B* **2015**, *119*, 928.
- [51] T. Climent, R. González-Luque, M. Merchán, L. Serrano-Andrés, *J. Phys. Chem. A* **2006**, *110*, 13584.
- [52] J. y Hasegawa, S. Bureekaew, H. Nakatsuji, *J. Photochem. Photobiol. A* **2007**, *189*, 205.
- [53] S. Salzmann, V. V. Martinez-Junza, B. Zorn, S. E. Braslavsky, M. M. Mansurova, C. M. Marian, W. Gärtner, *J. Phys. Chem. A* **2009**, *113*, 9365.
- [54] O. Falklöf, B. Durbeej, P. Norman, *J. Phys. Chem. A* **2015**, *119*, 11911.
- [55] R. K. Kar, V. A. Borin, Y. Ding, J. Matysik, I. Schapiro, *Photochem. Photobiol.* **2019**, *95*, 662.
- [56] F. Abyar, I. Novak, *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2022**, *264*, 120268.
- [57] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648.
- [58] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, *J. Phys. Chem.* **1994**, *98*, 11623.
- [59] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.
- [60] F. Neese, *WIREs Comput. Mol. Sci.* **2012**, *2*, 73.
- [61] F. Neese, *WIREs Comput. Mol. Sci.* **2018**, *8*, e1327.
- [62] Curtolo, F.; Arantes, G. M. Lumiflavin molecular structures and energies [Data set]. <https://doi.org/10.5281/zenodo.6348046>, **2022**.
- [63] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098.
- [64] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785.
- [65] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865.
- [66] Y. Zhao, D. G. Truhlar, *J. Chem. Phys.* **2006**, *125*, 194101.
- [67] C. Adamo, V. Barone, V., *J. Chem. Phys.* **1999**, *110*, 6158.
- [68] Y. Zhao, D. G. Truhlar, *Theor. Chem. Accounts* **2008**, *120*, 215.
- [69] C. Riplinger, P. Pinski, U. Becker, E. F. Valeev, F. Neese, *J. Chem. Phys.* **2016**, *144*, 024109.
- [70] R. A. Kendall, T. H. Dunning, R. J. Harrison, *J. Chem. Phys.* **1992**, *96*, 6796.
- [71] F. Weigend, A. Köhn, C. Hättig, *J. Chem. Phys.* **2002**, *116*, 3175.
- [72] T. Helgaker, P. Jørgensen, J. Olsen, *Molecular Electronic-Structure Theory*, John Wiley & Sons, Ltd, Chichester, England **2000**; Ch.5, p. 142.
- [73] T. H. Dunning, *J. Chem. Phys.* **1989**, *90*, 1007.
- [74] D. G. Liakos, M. Sparta, M. K. Kesharwani, J. M. L. Martin, F. Neese, *J. Chem. Theory Comput.* **2015**, *11*, 1525.
- [75] C. Angeli, R. Cimiraglia, J.-P. Malrieu, *J. Chem. Phys.* **2002**, *117*, 9138.
- [76] J. Zheng, X. Xu, D. G. Truhlar, *Theor. Chem. Accounts* **2011**, *128*, 295.
- [77] N. C. Handy, A. J. Cohen, *Mol. Phys.* **2001**, *99*, 403.
- [78] J. Tao, J. P. Perdew, V. N. Staroverov, G. E. Scuseria, *Phys. Rev. Lett.* **2003**, *91*, 146401.
- [79] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.
- [80] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.* **2011**, *32*, 1456.
- [81] W.-M. Hoe, A. J. Cohen, N. C. Handy, *Chem. Phys. Lett.* **2001**, *341*, 319.
- [82] T. Yanai, D. P. Tew, N. C. Handy, *Chem. Phys. Lett.* **2004**, *393*, 51.
- [83] Y. Tawada, T. Tsuneda, S. Yanagisawa, T. Yanai, K. Hirao, *J. Chem. Phys.* **2004**, *120*, 8425.
- [84] J.-D. Chai, M. Head-Gordon, *J. Chem. Phys.* **2008**, *128*, 084106.
- [85] S. Grimme, *J. Chem. Phys.* **2006**, *124*, 034108.
- [86] F. Weigend, *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057.
- [87] F. Neese, F. Wennmohs, A. Hansen, U. Becker, *Chem. Phys.* **2009**, *356*, 98.
- [88] R. Izsák, F. Neese, *J. Chem. Phys.* **2011**, *135*, 144105.
- [89] S. Grimme, *Chem. Eur. J.* **2012**, *18*, 9955.
- [90] G. M. Arantes, *J. Phys. Chem. B* **2008**, *112*, 15244.
- [91] V. Barone, M. Cossi, *J. Phys. Chem. A* **1998**, *102*, 1995.
- [92] A. Karton, J. M. L. Martin, *Theor. Chem. Accounts* **2006**, *115*, 330.
- [93] T. Helgaker, W. Klopper, H. Koch, J. Noga, *J. Chem. Phys.* **1997**, *106*, 9639.
- [94] D. G. Truhlar, *Chem. Phys. Lett.* **1998**, *294*, 45.
- [95] A. Halkier, W. Klopper, T. Helgaker, P. Jørgensen, *J. Chem. Phys.* **1999**, *111*, 4424.
- [96] F. Neese, E. F. Valeev, *J. Chem. Theory Comput.* **2011**, *7*, 33.
- [97] D. Datta, S. Kossmann, F. Neese, *J. Chem. Phys.* **2016**, *145*, 114101.
- [98] J. C. Rienstra-Kiracofe, G. S. Tschumper, H. F. Schaefer, S. Nandi, G. B. Ellison, *Chem. Rev.* **2002**, *102*, 231.
- [99] M. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation*, OUP Oxford, England **2010**; Ch.4, p. 135.
- [100] T. J. Lee, P. R. Taylor, *Int. J. Quantum Chem.* **1989**, *36*, 199.
- [101] A. Nunes-Alves, G. M. Arantes, *J. Chem. Inf. Model.* **2014**, *54*, 2309.
- [102] M. Drosou, C. A. Mitsopoulou, D. A. Pantazis, *J. Chem. Theory Comput.* **2022**, *18*, 6.
- [103] R. Norrestam, B. Stensland, *Acta Crystallogr. B* **1972**, *28*, 440.

- [104] C. T. W. Moonen, J. Vervoort, F. Mueller, *Biochemistry* **1984**, 23, 4868.
- [105] J. J. P. Stewart, *J. Comput. Chem.* **1989**, 10, 209.
- [106] M. H. Teixeira, G. M. Arantes, *Biochim. Biophys. Acta Bioenerg.* **2019**, 1860, 541.
- [107] M. H. Teixeira, G. M. Arantes, *RSC Adv.* **2019**, 9, 16892.
- [108] M. Kılıç, B. Ensing, *J. Phys. Chem. B* **2019**, 123, 9751.
- [109] G. M. Arantes, *Biochem. J.* **2006**, 399, 343.
- [110] G. M. Arantes, M. C. C. Ribeiro, *J. Chem. Phys.* **2008**, 128, 114503.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: F. Curtolo, G. M. Arantes, *J. Comput. Chem.* **2022**, 43(23), 1561. <https://doi.org/10.1002/jcc.26957>

Supporting Information

Molecular Properties and Tautomeric Equilibria of Isolated Flavins

Felipe Curtolo* 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

E-mail: felipe.curtolo@usp.br, garantes@iq.usp.br

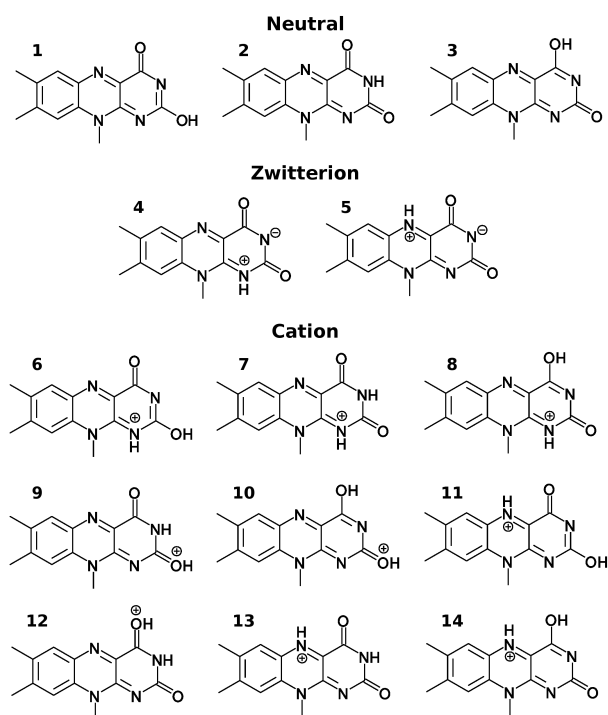


Figure S1: Tautomeric forms of lumiflavin quinones with neutral and cationic charge.

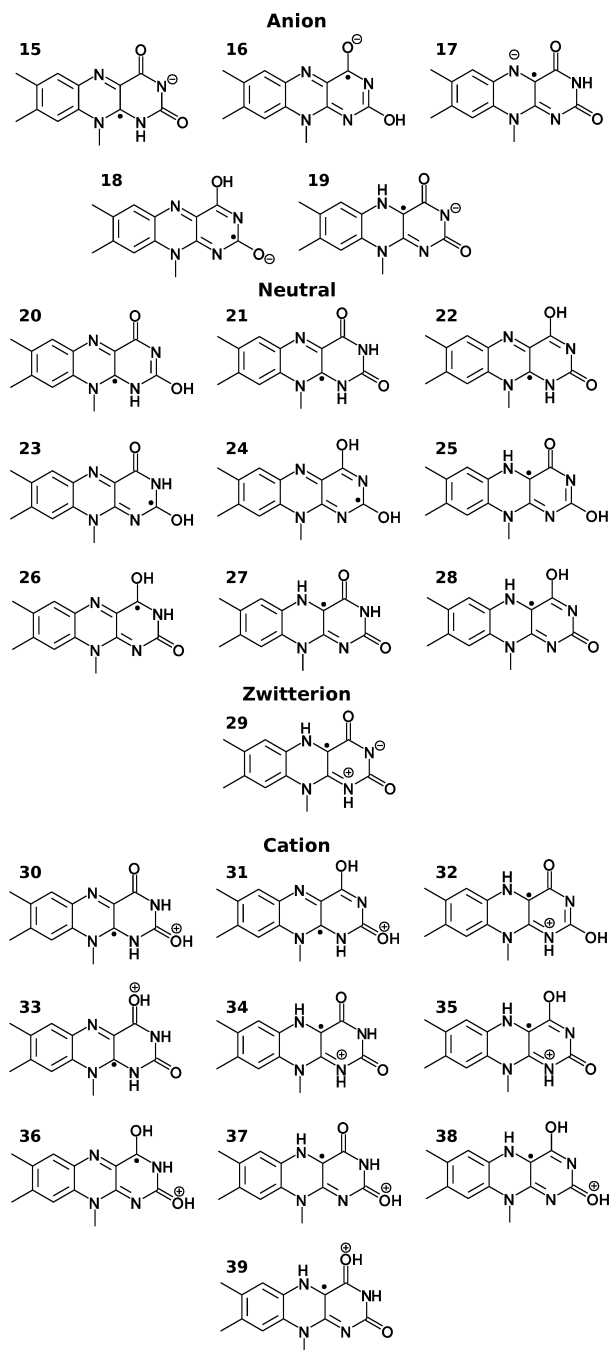


Figure S2: Tautomeric forms of lumiflavin semiquinones with anionic, neutral and cationic charge.

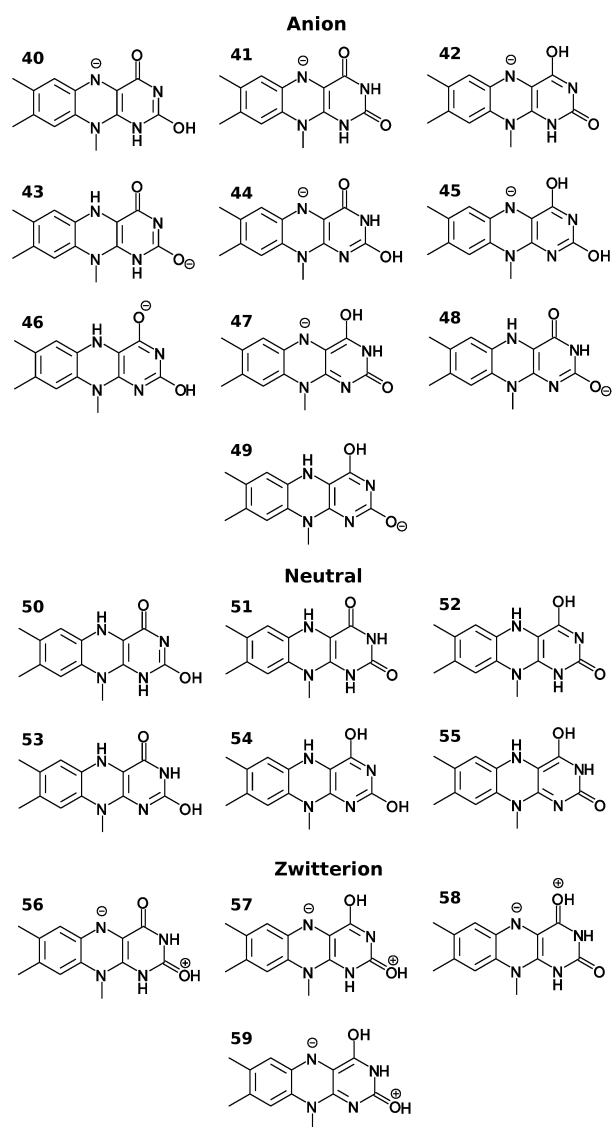


Figure S3: Tautomeric forms of lumiflavin quinols with anionic and neutral charge.

Table S1: *Ab initio* electronic-structure calculations of flavins previously reported.

Methods	Properties	Molecule	Oxidation state	Titration protons	Reference
HF, MP2	tautomerism, conformations	lumiflavin	quinone, semiquinone, quinol	0, 1, 2, 3, 4	¹²
HF	tautomerism, conformations	lumiflavin	quinone, semiquinone, quinol	0, 1, 2, 3, 4	¹³
DFT	tautomerism, conformations	isoalloxazine	quinone, semiquinone, quinol	0, 1, 2, 3, 4	¹⁴
HF	tautomerism, excitation energies	lumiflavin	quinone, semiquinone, quinol	1, 2, 3	¹⁵
DFT, MP2	conformations	lumiflavin	quinone, quinol	1, 3	²⁴
DFT	tautomerism, vibrational frequencies	lumiflavin	quinone, semiquinone	1, 2	¹⁶
DFT	conformations, ionization potential	lumiflavin, 7,8-substituted lumiflavin	quinone, semiquinone, quinol	1	²⁵
CIS, TD-DFT, DFT/MRCI	excitation energies	lumiflavin	quinone	1	³⁷
TD-DFT	excitation energies	isoalloxazine	quinone	1	³⁸
TD-DFT	excitation energies	lumiflavin, lumichrome	quinone	1	³⁹
TD-DFT	excitation energies	lumiflavin, methyl-isoalloxazines	quinone	1	⁴⁰
TD-DFT	excitation energies	methyl-isoalloxazines	quinone	1	⁴¹
DFT	vibrational frequencies	lumiflavin	quinol	2, 3	²⁶
TD-DFT	excitation energies	3-ethyl-lumiflavin	quinone	1	⁴²
TD-DFT	excitation energies	riboflavin, iso-(6,7)-riboflavin	quinone	1	⁴³
TD-DFT	excitation energies	methyl-isoalloxazines	quinone	0	⁴⁴
TD-DFT	excitation energies	5-deaza-riboflavin	quinone	1	⁴⁵
DFT	vibrational frequencies	lumiflavin	quinone	1	²⁷
CASPT2	excitation energies	isoalloxazine	quinone	1	⁵¹

TD-DFT	excitation energies	lumiflavin, 8-substituted lumiflavin	quinone, quinol	1, 3	⁴⁶
SAC-CI	excitation energies	lumiflavin	quinone	1	⁵²
DFT	electron affinity, proton affinity	lumiflavin	quinone, semiquinone, quinol	1, 2, 3	²⁸
DFT	conformations, ionization potential	lumiflavin, 7,8-substituted lumiflavin	quinone, semiquinone	1	²⁹
TD-DFT, DFT/MRCI	excitation energies	isoalloxazine, lumiflavin	quinone	1	⁴⁷
HF, CIS, DFT	vibrational frequencies	riboflavin	quinone	1	³⁰
DFT, DFT/MRCI	tautomerism, proton affinity, excitation energies	lumiflavin	quinone	0, 1, 2	¹⁷
DFT/MRCI	excitation energies	riboflavin, riboflavin derivatives	quinone	1	⁵³
TD-DFT	excitation energies	riboflavin	quinone	1	⁴⁸
DFT	electron affinity, ionization potential	lumiflavin, 7,8-substituted lumiflavin	quinone, semiquinone	1	³¹
DFT	vibrational frequencies	lumiflavin	quinol	2, 3	³²
TD-DFT, CASSCF	excitation energies, vibrational frequencies	lumiflavin	quinone	1	⁴⁹
DFT	electron affinity	lumiflavin	quinone, semiquinone, quinol	1, 2	³³
DFT	electron affinity, ionization potential, proton affinity	lumiflavin	quinone, semiquinone	1, 2	³⁴
HF, MP2, DFT	tautomerism, vibrational frequencies	lumiflavin, lumichrome, riboflavin, FMN	quinone	1, 2	¹⁸
TD-DFT, DFT/MRCI	excitation energies	roseoflavin	quinone	1	⁵⁰
Relativistic 4C-TD-DFT	spin-forbidden electronic transitions	lumiflavin	quinone	1	⁵⁴
TD-DFT, ADC(2), CC2	excitation energies, vibrational frequencies	lumiflavin	quinone	1	⁵⁵

MP2, DFT	dipole, conformations	lumiflavin	quinone, semiquinone, quinol	1, 2, 3	³⁵
DFT, TD-DFT	excitation energies, electron affinity	isoalloxazine	quinone, semiquinone, quinol	1, 2, 3	³⁶
EOM-CCSD	ionization potential	lumiflavin, riboflavin	quinone	1	⁵⁶

Table S2: τ_1 diagnostic calculated with DLPNO-CCSD(T)/aug-cc-pVTZ for flavoquinone tautomers

Tautomer	τ_1	Tautomer	τ_1	Tautomer	τ_1
1	0.014	7	0.014	13	0.014
2	0.014	8	0.015	14	0.014
3	0.014	9	0.014	2 bent	0.015
4	0.015	10	0.015	7 bent	0.015
5	0.015	11	0.014	9 bent	0.015
6	0.014	12	0.015	10 bent	0.015

Table S3: τ_1 diagnostic calculated with DLPNO-CCSD(T)/aug-cc-pVTZ for flavosemiquinone tautomers

Tautomer	τ_1	Tautomer	τ_1	Tautomer	τ_1
15	0.019	26	0.018	36	0.018
16	0.019	27	0.016	37	0.018
17	0.017	28	0.016	38	0.019
18	0.019	29	0.016	39	0.018
19	0.016	30	0.019	17 bent	0.020
20	0.018	31	0.018	24 bent	0.018
21	0.018	32	0.018	27 bent	0.015
22	0.019	33	0.018	34 bent	0.018
23	0.019	34	0.018	37 bent	0.019
24	0.019	35	0.017	38 bent	0.019
25	0.016				

Table S4: τ_1 diagnostic calculated with DLPNO-CCSD(T)/aug-cc-pVTZ for flavoquinol tautomers

Tautomer	τ_1	Tautomer	τ_1	Tautomer	τ_1
40	0.014	48	0.014	56	0.014
41	0.014	49	0.014	57	0.014
42	0.014	50	0.013	58	0.014
43	0.014	51	0.014	59	0.014
44	0.014	52	0.014	48 planar	0.014
45	0.014	53	0.014	51 planar	0.014
46	0.014	54	0.013	53 planar	0.014
47	0.014	55	0.013	54 planar	0.013

Table S5: Calculated electron affinity (EA_{therm}) and gas-phase basicity (GB) for **2**. GB values were obtained at 298.15 K

Method	EA_{therm} (eV)	GB (kJ.mol ⁻¹)
MP2/cc-pVDZ	1.23	944
MP2/cc-pVTZ	1.66	933
MP2/aug-cc-pVDZ	1.81	918
MP2/aug-cc-pVTZ	1.88	922
MP2/aug-cc-pVQZ	1.92	923
DLPNO-CCSD/cc-pVDZ	1.24	951
DLPNO-CCSD/cc-pVTZ	1.59	943
DLPNO-CCSD/aug-cc-pVDZ	1.73	928
DLPNO-CCSD/aug-cc-pVTZ	1.78	933
DLPNO-CCSD/aug-cc-pVQZ	1.79	933

Table S6: Calculated adiabatic electron affinity (EA_{adiab}) for **2** for various basis set.

Method	MP2	DLPNO-CCSD	DLPNO-CCSD(T)
cc-pVDZ	1.19	1.20	1.17
cc-pVTZ	1.62	1.56	1.55
cc-pV[D/T]Z	1.84	1.74	1.75
aug-cc-pVDZ	1.77	1.70	1.71
aug-cc-pVTZ	1.84	1.74	1.75
aug-cc-pVQZ	1.88	1.75	1.76
aug-cc-pV[D/T]Z	1.89	1.77	1.78
aug-cc-pV[T/Q]Z	1.91	1.76	1.77

Table S7: Energetic contributions to the Gibbs free energy G of flavoquinone tautomers at 298.15 K, in E_h .

Molecule	E_{el}	$ZPVE$	H	$-TS$
<i>Neutral and zwitterion</i>				
1	-871.054453	0.238905	-870.798647	-0.058578
2	-871.075697	0.239063	-870.819675	-0.058760
3	-871.050869	0.238485	-870.795390	-0.058746
4	-871.027683	0.237637	-870.772724	-0.059683
5	-871.012071	0.237658	-870.757296	-0.059107
<i>Cation</i>				
6	-871.420880	0.251122	-871.152390	-0.059465
7	-871.448112	0.251827	-871.178980	-0.059407
8	-871.432070	0.251490	-871.163347	-0.059186
9	-871.452954	0.251958	-871.183950	-0.058709
10	-871.446807	0.252042	-871.177856	-0.058342
11	-871.419002	0.251929	-871.150077	-0.058658
12	-871.437656	0.251502	-871.169096	-0.058730
13	-871.431723	0.251877	-871.162711	-0.059023
14	-871.401841	0.250967	-871.133579	-0.059288

Table S8: Energetic contributions to the Gibbs free energy G of flavosemiquinone tautomers at 298.15 K, in E_h .

Molecule	E_{el}	$ZPVE$	H	$-TS$
<i>Anion</i>				
15	-871.101536	0.234920	-870.849091	-0.060554
16	-871.118713	0.235696	-870.865885	-0.059599
17	-871.138262	0.236047	-870.885097	-0.059640
18	-871.121223	0.235867	-870.868277	-0.059474
19	-871.103222	0.235367	-870.850501	-0.060119
<i>Neutral and zwitterion</i>				
20	-871.629599	0.248598	-871.363294	-0.060708
21	-871.660387	0.249503	-871.393345	-0.060455
22	-871.650300	0.249525	-871.383405	-0.060061
23	-871.656626	0.249342	-871.389921	-0.059954
24	-871.666446	0.250141	-871.399251	-0.059249
25	-871.656741	0.250248	-871.389287	-0.059631
26	-871.660511	0.249725	-871.393568	-0.059624
27	-871.671171	0.250252	-871.403587	-0.059954
28	-871.640749	0.249258	-871.373974	-0.060222
29	-871.633570	0.249084	-871.366869	-0.060591
<i>Cation</i>				
30	-871.998347	0.260920	-871.719335	-0.061360
31	-872.031451	0.262657	-871.751333	-0.060070
32	-872.035215	0.262820	-871.754737	-0.060492
33	-872.039920	0.262822	-871.759572	-0.060259
34	-872.057256	0.263379	-871.776246	-0.060522
35	-872.035061	0.262753	-871.754650	-0.060497
36	-872.042577	0.262828	-871.762377	-0.059754
37	-872.061185	0.263466	-871.780311	-0.059930
38	-872.053205	0.263241	-871.772572	-0.059796
39	-872.028801	0.261990	-871.748965	-0.060731

Table S9: Energetic contributions to the Gibbs free energy G of flavoquinol tautomers at 298.15 K, in E_h .

Molecule	E_{el}	$ZPVE$	H	$-TS$
<i>Anion</i>				
40	-871.688511	0.246230	-871.424581	-0.059954
41	-871.721651	0.247104	-871.456969	-0.059840
42	-871.715350	0.246658	-871.450953	-0.060181
43	-871.731174	0.247319	-871.466020	-0.060347
44	-871.699656	0.245586	-871.435996	-0.061007
45	-871.727312	0.247033	-871.462711	-0.059748
46	-871.744043	0.247960	-871.478574	-0.059633
47	-871.726630	0.246906	-871.462021	-0.059990
48	-871.759648	0.248125	-871.493958	-0.059809
49	-871.728751	0.247264	-871.463754	-0.060048
<i>Neutral and zwitterion</i>				
50	-872.265193	0.261311	-871.985919	-0.060382
51	-872.291731	0.262064	-872.011802	-0.060291
52	-872.272457	0.261663	-871.992994	-0.060066
53	-872.285623	0.261869	-872.005995	-0.060017
54	-872.285746	0.261775	-872.006203	-0.059955
55	-872.275674	0.262195	-871.995904	-0.059667
56	-872.192797	0.259069	-871.915488	-0.061014
57	-872.231252	0.260364	-871.953001	-0.060188
58	-872.241593	0.260705	-871.962993	-0.060292
59	-872.238925	0.260381	-871.960600	-0.060272

Table S10: EA_{adiab} (eV) of neutral and zwitterionic flavoquinone tautomers calculated with various DFT functionals. Reference is DLPNO-CCSD(T)/aug-cc-pV[D/T]Z.

Tautomer	1	2	3	4	5
BLYP	1.78	1.83	1.97	2.00	2.45
OLYP	1.71	1.77	1.89	1.95	2.39
PBE	1.92	1.98	2.13	2.15	2.58
BPBE	1.88	1.93	2.09	2.10	2.54
OPBE	1.82	1.88	2.02	2.06	2.48
TPSS	1.85	1.91	2.06	2.07	2.53
M06-L	1.73	1.79	1.94	1.97	2.45
B3LYP	1.84	1.90	2.03	2.11	2.61
B3LYP+D3	1.84	1.89	2.04	2.11	2.60
PBE0	1.92	1.98	2.13	2.20	2.70
O3LYP	1.85	1.91	2.06	2.11	2.58
M06	1.90	1.95	2.09	2.17	2.69
M06-2X	1.95	1.98	2.13	2.22	2.77
CAM-B3LYP	1.97	2.01	2.16	2.26	2.81
LC-BLYP	2.01	2.05	2.23	2.30	2.86
ω B97X	1.93	1.98	2.14	2.24	2.81
B2PLYP	1.74	1.78	1.93	1.99	2.51
Reference	1.84	1.78	1.99	2.08	2.54

Table S11: EA_{adiab} (eV) of cationic flavoquinone tautomers calculated with various DFT functionals.

Tautomer	6	7	8	9	10	11	12	13	14
BLYP	5.60	5.69	5.82	5.48	5.80	6.28	5.89	6.35	6.33
OLYP	5.56	5.65	5.76	5.43	5.74	6.23	5.85	6.31	6.29
PBE	5.74	5.83	5.98	5.62	5.96	6.42	6.04	6.49	6.47
BPBE	5.71	5.80	5.95	5.58	5.92	6.38	6.01	6.46	6.44
OPBE	5.68	5.77	5.90	5.54	5.87	6.34	5.98	6.42	6.41
TPSS	5.69	5.79	5.93	5.57	5.91	6.38	6.00	6.46	6.44
M06-L	5.61	5.71	5.86	5.47	5.83	6.33	5.92	6.41	6.39
B3LYP	5.72	5.82	5.94	5.58	5.93	6.49	6.06	6.57	6.55
B3LYP+D3	5.72	5.82	5.95	5.58	5.93	6.49	6.06	6.57	6.55
PBE0	5.82	5.92	6.06	5.67	6.04	6.59	6.17	6.68	6.66
O3LYP	5.72	5.82	5.97	5.58	5.94	6.46	6.06	6.54	6.52
M06	5.81	5.91	6.03	5.66	6.02	6.59	6.14	6.67	6.66
M06-2X	5.87	5.96	6.09	5.72	6.11	6.73	6.23	6.79	6.78
CAM-B3LYP	5.89	5.99	6.13	5.74	6.13	6.74	6.28	6.83	6.81
LC-BLYP	5.93	6.03	6.19	5.78	6.21	6.81	6.33	6.88	6.88
ω B97X	5.88	5.98	6.12	5.72	6.13	6.76	6.28	6.84	6.84
B2PLYP	5.61	5.70	5.84	5.47	5.84	6.43	5.95	6.48	6.45
Reference	5.75	5.84	5.99	5.61	6.03	6.51	6.11	6.56	6.55

Table S12: EA_{adiab} (eV) of neutral and zwitterionic flavosemiquinone tautomers calculated with various DFT functionals.

Tautomer	20	21	22	23	24	25	26	27	28	29
BLYP	1.38	1.43	1.61	1.15	1.49	1.96	1.58	2.02	2.00	2.17
OLYP	1.34	1.39	1.54	1.08	1.41	1.87	1.51	1.94	1.94	2.11
PBE	1.47	1.52	1.74	1.26	1.62	2.06	1.71	2.12	2.12	2.27
BPBE	1.42	1.47	1.69	1.21	1.57	2.01	1.66	2.07	2.07	2.23
OPBE	1.40	1.45	1.64	1.15	1.49	1.93	1.59	2.01	2.01	2.17
TPSS	1.37	1.42	1.64	1.14	1.51	1.98	1.61	2.05	2.04	2.20
M06-L	1.14	1.20	1.49	0.98	1.37	1.87	1.46	1.94	1.93	2.10
B3LYP	1.36	1.41	1.60	1.07	1.47	2.03	1.59	2.09	2.08	2.29
B3LYP+D3	1.36	1.41	1.59	1.06	1.47	2.03	1.58	2.09	2.08	2.29
PBE0	1.36	1.42	1.64	1.08	1.51	2.06	1.62	2.13	2.13	2.33
O3LYP	1.37	1.42	1.66	1.14	1.52	2.01	1.64	2.09	2.08	2.26
M06	1.44	1.50	1.70	1.16	1.58	2.14	1.68	2.20	2.20	2.41
M06-2X	1.55	1.61	1.77	1.19	1.66	2.27	1.78	2.32	2.32	2.55
CAM-B3LYP	1.40	1.46	1.66	1.06	1.53	2.17	1.66	2.23	2.23	2.45
LC-BLYP	1.45	1.51	1.73	1.11	1.62	2.24	1.74	2.28	2.31	2.51
ω B97X	1.38	1.44	1.61	0.99	1.49	2.17	1.62	2.22	2.24	2.46
B2PLYP	1.55	1.60	1.73	1.20	1.62	2.14	1.74	2.19	2.20	2.39
Reference	1.67	1.73	1.85	1.27	1.74	2.44	1.88	2.47	2.45	2.70

Table S13: EA_{adiab} (eV) of the cationic flavosemiquinone tautomers calculated with various DFT functionals. Reference is DLPNO-CCSD(T)/aug-cc-pV[D/T]Z.

Tautomer	30	31	32	33	34	35	36	37	38	39
BLYP	5.02	5.22	5.76	5.26	5.89	5.93	5.13	5.62	5.79	5.98
OLYP	4.99	5.16	5.72	5.21	5.85	5.90	5.07	5.57	5.74	5.93
PBE	5.11	5.36	5.88	5.40	6.00	6.05	5.27	5.73	5.91	6.10
BPBE	5.08	5.31	5.84	5.36	5.97	6.02	5.22	5.69	5.87	6.07
OPBE	5.07	5.27	5.81	5.32	5.94	6.00	5.17	5.65	5.83	6.03
TPSS	5.01	5.25	5.81	5.30	5.94	5.99	5.17	5.67	5.85	6.06
M06-L	4.78	5.13	5.74	5.18	5.87	5.93	5.04	5.59	5.79	5.98
B3LYP	5.01	5.20	5.88	5.26	6.02	6.07	5.12	5.73	5.93	6.19
B3LYP+D3	5.01	5.20	5.88	5.25	6.02	6.07	5.12	5.72	5.93	6.20
PBE0	5.02	5.24	5.94	5.30	6.07	6.14	5.16	5.78	5.99	6.27
O3LYP	5.04	5.28	5.87	5.34	6.01	6.07	5.19	5.72	5.91	6.15
M06	5.06	5.30	6.03	5.35	6.17	6.22	5.21	5.87	6.08	6.33
M06-2X	5.18	5.39	6.15	5.44	6.28	6.34	5.31	5.98	6.22	6.55
CAM-B3LYP	5.03	5.24	6.05	5.31	6.19	6.25	5.17	5.89	6.12	6.45
LC-BLYP	5.06	5.31	6.10	5.37	6.23	6.31	5.24	5.94	6.19	6.53
ω B97X	4.97	5.18	6.06	5.25	6.19	6.27	5.11	5.89	6.14	6.50
B2PLYP	5.22	5.37	5.98	5.42	6.11	6.17	5.28	5.83	6.04	6.33
Reference	5.34	5.50	6.30	5.55	6.42	6.49	5.41	6.15	6.37	6.71

Table S14: GB (kJ.mol⁻¹) at 298.15 K of neutral and zwitterionic flavoquinone tautomers calculated with various DFT functionals. Reference is DLPNO-CCSD(T)/aug-cc-pV[D/T]Z.

Tautomer	1	2	3	4	5
BLYP	986	923	970	1029	1036
OLYP	999	937	987	1044	1052
PBE	985	923	973	1030	1036
BPBE	993	931	980	1038	1044
OPBE	1006	946	998	1052	1059
TPSS	996	934	983	1041	1047
M06-L	1005	939	990	1053	1059
B3LYP	991	930	980	1041	1044
B3LYP+D3	995	932	982	1044	1047
PBE0	993	933	985	1045	1046
O3LYP	990	929	982	1039	1043
M06	989	927	978	1040	1042
M06-2X	979	923	971	1031	1032
CAM-B3LYP	984	923	973	1039	1037
LC-BLYP	964	902	952	1019	1016
ω B97X	991	928	979	1048	1045
B2PLYP	984	924	972	1036	1038
Reference	986	930	977	1040	1038

Table S15: GB (kJ.mol⁻¹) at 298.15 K of anionic flavosemiquinone tautomers calculated with various DFT functionals.

Tautomer	15	16	17	18	19
BLYP	1385	1350	1329	1344	1413
OLYP	1401	1366	1343	1360	1430
PBE	1385	1354	1329	1344	1413
BPBE	1394	1362	1338	1353	1422
OPBE	1410	1378	1351	1369	1439
TPSS	1399	1366	1340	1357	1426
M06-L	1413	1376	1349	1369	1441
B3LYP	1399	1364	1336	1356	1426
B3LYP+D3	1402	1368	1341	1359	1430
PBE0	1403	1371	1339	1362	1431
O3LYP	1396	1365	1337	1355	1426
M06	1400	1365	1333	1357	1427
M06-2X	1391	1362	1327	1355	1420
CAM-B3LYP	1399	1363	1333	1356	1425
LC-BLYP	1378	1344	1317	1335	1404
ω B97X	1409	1369	1341	1363	1433
B2PLYP	1393	1358	1334	1349	1421
Reference	1402	1373	1336	1367	1425

Table S16: GB (kJ.mol⁻¹) at 298.15 K of neutral and zwitterionic flavosemiquinone tautomers calculated with various DFT functionals.

Tautomer	20	21	22	23	24	25	26	27	28	29
BLYP	1005	984	962	1003	966	998	934	950	1004	1035
OLYP	1018	996	975	1017	980	1011	948	965	1021	1050
PBE	1006	984	961	1004	963	996	933	950	1006	1034
BPBE	1013	991	970	1012	972	1005	941	958	1014	1042
OPBE	1025	1002	981	1024	985	1017	954	973	1031	1057
TPSS	1016	993	972	1014	973	1007	943	960	1016	1045
M06-L	1021	998	980	1020	978	1016	951	965	1023	1057
B3LYP	1010	987	968	1009	969	1004	940	958	1015	1046
B3LYP+D3	1015	992	971	1013	972	1008	943	961	1018	1049
PBE0	1012	988	970	1011	968	1005	943	962	1020	1049
O3LYP	1012	988	966	1010	968	1002	938	957	1016	1043
M06	1004	981	965	1004	963	1001	937	955	1013	1044
M06-2X	999	976	953	998	956	993	933	955	1011	1037
CAM-B3LYP	1004	981	962	1004	961	998	934	953	1011	1044
LC-BLYP	989	966	942	988	941	978	914	933	991	1024
ω B97X	1010	987	970	1010	966	1006	941	959	1017	1054
B2PLYP	1008	985	958	1006	965	997	933	954	1011	1041
Reference	1001	979	962	999	955	1001	942	963	1019	1048

Table S17: GB (kJ.mol⁻¹) at 298.15 K of anionic flavoquinol tautomers calculated with various DFT functionals.

Tautomer	40	41	42	43	44	45	46	47	48	49
BLYP	1429	1415	1374	1393	1434	1381	1352	1343	1319	1370
OLYP	1441	1427	1390	1411	1449	1398	1368	1360	1336	1389
PBE	1432	1417	1369	1394	1433	1377	1351	1340	1318	1372
BPBE	1440	1425	1379	1403	1443	1386	1360	1350	1327	1381
OPBE	1452	1437	1395	1420	1457	1403	1376	1366	1343	1399
TPSS	1445	1430	1384	1407	1449	1391	1363	1356	1330	1385
M06-L	1466	1450	1396	1421	1463	1404	1375	1366	1342	1397
B3LYP	1447	1432	1391	1406	1457	1398	1361	1363	1330	1387
B3LYP+D3	1452	1436	1395	1409	1462	1402	1365	1368	1334	1389
PBE0	1454	1438	1393	1410	1462	1400	1363	1366	1332	1394
O3LYP	1447	1431	1383	1405	1451	1391	1359	1354	1327	1386
M06	1448	1432	1390	1407	1456	1397	1360	1362	1330	1388
M06-2X	1443	1427	1389	1398	1458	1395	1351	1365	1323	1388
CAM-B3LYP	1453	1437	1396	1404	1468	1403	1356	1370	1329	1387
LC-BLYP	1439	1422	1376	1383	1452	1382	1334	1351	1309	1366
ω B97X	1463	1446	1408	1414	1481	1415	1365	1384	1339	1395
B2PLYP	1437	1421	1385	1400	1451	1391	1353	1358	1325	1382
Reference	1449	1432	1397	1407	1468	1401	1360	1375	1335	1398

Table S18: Electric dipole moment $||\mu||$ (in D) of selected flavin tautomers calculated with various DFT functionals. Reference is obtained with MP2/aug-cc-pV[D/T]Z relaxed electron density.

Tautomer	2	7	9	10	17	24	27	34	37	38	48	51	53	54
BLYP	9.45	8.30	6.48	1.90	9.79	4.48	9.34	6.19	3.56	2.80	10.50	5.84	3.95	2.83
OLYP	9.36	8.22	6.45	1.83	9.64	4.32	9.26	6.13	3.54	2.71	10.37	5.77	3.94	2.71
PBE	9.39	8.26	6.46	1.86	9.69	4.40	9.26	6.14	3.53	2.76	10.39	5.79	3.92	2.77
BPBE	9.37	8.24	6.45	1.84	9.70	4.39	9.24	6.13	3.53	2.76	10.40	5.79	3.93	2.75
OPBE	9.28	8.16	6.42	1.76	9.55	4.22	9.16	6.06	3.51	2.67	10.27	5.71	3.92	2.62
TPSS	9.38	8.22	6.46	1.80	9.79	4.40	9.25	6.09	3.52	2.75	10.48	5.80	3.95	2.76
M06-L	9.16	8.02	6.30	1.51	9.64	4.09	9.04	5.85	3.34	2.72	10.40	5.55	3.87	2.49
B3LYP	9.61	8.40	6.55	1.90	10.29	4.61	9.57	6.27	3.58	2.93	11.01	5.95	4.08	2.92
PBE0	9.57	8.37	6.53	1.84	10.30	4.55	9.52	6.21	3.55	2.91	11.02	5.93	4.08	2.88
O3LYP	9.59	8.41	6.56	1.93	10.07	4.57	9.54	6.29	3.61	2.88	10.79	5.94	4.04	2.90
M06	9.32	8.21	6.40	1.63	9.94	4.30	9.22	6.07	3.43	2.92	10.76	5.66	3.95	2.66
M06-2X	9.51	8.30	6.50	1.81	10.47	4.67	9.67	6.16	3.53	2.93	11.35	5.97	4.12	3.03
CAM-B3LYP	9.58	8.37	6.54	1.78	10.67	4.70	9.61	6.18	3.52	2.99	11.41	5.99	4.14	3.01
LC-BLYP	9.46	8.28	6.52	1.69	10.66	4.75	9.56	6.05	3.46	2.98	11.49	5.97	4.12	3.08
ω B97X	9.36	8.17	6.44	1.55	10.63	4.53	9.43	5.95	3.39	2.94	11.44	5.87	4.11	2.90
B2PLYP	9.42	8.27	6.50	1.87	10.12	4.65	9.52	6.11	3.54	2.82	11.07	5.87	4.02	2.97
Reference	9.03	8.05	6.42	2.01	9.44	4.68	10.11	7.20	4.73	4.24	11.19	5.73	3.85	3.08

Table S19: Torsion energy (E_{tor} , in kJ.mol^{-1}) of selected flavin tautomers calculated with various DFT functionals. Reference is DLPNO-CCSD(T)/aug-cc-pV[D/T]Z.

Tautomer	2	7	9	10	17	24	27	34	37	38	48	51	53	54
BLYP	39.0	43.4	48.7	48.6	25.8	30.0	21.1	20.4	26.1	26.6	-1.0	-7.2	-1.8	-1.4
OLYP	37.9	41.5	47.3	47.1	25.0	28.8	20.3	18.7	24.8	25.3	-1.3	-8.9	-2.8	-2.1
PBE	39.6	43.1	48.8	48.5	26.6	30.4	22.1	20.6	26.7	27.2	-0.5	-7.0	-1.3	-1.0
BPBE	39.5	43.0	48.7	48.5	26.6	30.4	22.0	20.4	26.6	27.0	-0.4	-7.1	-1.3	-1.0
OPBE	38.3	41.0	47.1	46.8	25.6	29.0	21.0	18.6	25.1	25.6	-0.8	-8.9	-2.3	-1.8
TPSS	40.0	44.2	49.9	49.6	26.4	30.6	21.8	21.0	27.0	27.5	-1.3	-7.4	-2.1	-1.5
M06-L	41.5	46.5	51.9	51.5	28.1	32.1	23.1	22.9	28.8	29.1	-0.8	-6.8	-1.7	-1.2
B3LYP	41.4	47.0	52.7	52.5	26.8	31.2	21.9	22.2	28.0	28.5	-1.8	-8.1	-2.7	-1.9
B3LYP+D3	40.9	46.5	52.3	52.1	26.3	30.7	21.3	21.7	27.7	28.1	-2.1	-8.3	-2.9	-2.1
PBE0	42.2	47.5	53.5	53.2	27.5	31.5	22.7	22.6	28.8	29.2	-1.6	-8.2	-2.6	-1.8
O3LYP	41.6	45.7	51.5	51.3	27.7	31.6	23.2	21.8	28.1	28.6	-0.5	-7.4	-1.4	-1.1
M06	40.5	46.1	51.9	51.8	26.9	31.0	21.8	22.2	28.1	28.7	-1.1	-7.5	-2.0	-1.3
M06-2X	40.5	47.5	54.0	54.1	25.3	30.4	20.5	22.1	28.2	28.6	-3.4	-9.8	-4.6	-3.1
CAM-B3LYP	41.8	48.9	54.9	54.9	26.6	31.2	21.4	23.1	29.1	29.5	-2.7	-9.0	-3.6	-2.4
LC-BLYP	41.1	48.6	54.7	54.9	26.3	31.1	21.3	23.1	29.3	29.7	-2.7	-8.9	-3.4	-2.2
ω B97X	40.7	48.7	54.8	55.0	25.7	30.4	20.1	22.9	28.9	29.3	-3.7	-9.9	-4.8	-3.1
B2PLYP	40.2	46.8	52.4	52.2	24.6	30.1	19.9	21.1	26.7	27.2	-3.7	-9.9	-4.8	-3.3
Reference	38.5	47.7	53.8	53.6	19.6	30.4	18.4	22.8	27.3	29.4	-5.2	-10.9	-6.5	-4.2

3 Mechanisms for flavin-mediated oxidation: hydride or hydrogen-atom transfer?

Felipe Curtolo and Guilherme M. Arantes

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

Reproduced with permission from Curtolo, F.; Arantes, G. M. *J. Chem. Inf. Model.* **2020**, 60(12), 6282–6287. DOI: [10.1021/acs.jcim.0c00945](https://doi.org/10.1021/acs.jcim.0c00945). Copyright 2020 American Chemical Society.

Mechanisms for Flavin-Mediated Oxidation: Hydride or Hydrogen-Atom Transfer?

Felipe Curtolo and Guilherme M. Arantes*



Cite This: *J. Chem. Inf. Model.* 2020, 60, 6282–6287



Read Online

ACCESS |



Metrics & More

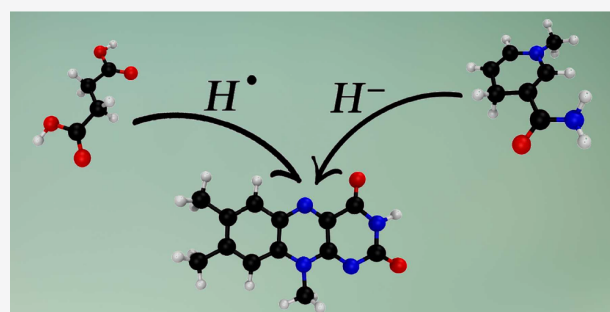


Article Recommendations



Supporting Information

ABSTRACT: Flavins are versatile biological cofactors which catalyze proton-coupled electron transfers (PCET) with varying number and coupling of electrons. Flavin-mediated oxidations of nicotinamide adenine dinucleotide (NADH) and of succinate, initial redox reactions in cellular respiration, were examined here with multiconfigurational quantum chemical calculations and a simple analysis of the wave function proposed to quantify electron transfer along the proton reaction coordinate. The mechanism of NADH oxidation is a prototypical hydride transfer, with two electrons moving concerted with the proton to the same acceptor group. However, succinate oxidation depends on the elimination step and can proceed through the transfer of a hydride or hydrogen atom, with proton and electrons moving to different groups in both cases. These results help to determine the mechanism of fundamental but still debated biochemical reactions and illustrate a new diagnostic tool for electron transfer that can be useful to characterize a broad class of PCET processes.



INTRODUCTION

Enzymes equipped with flavin cofactors comprise the most abundant class of natural catalysts for combined proton and electron transfer.^{1,2} The redox center in all natural flavins is formed by the heteronuclear tricyclic isoalloxazine ring (Figure 1), primarily attached to the protein by noncovalent hydrogen bonds, stacking, and cation- π contacts.^{3–5} These interactions also modulate the flavin redox potential from –400 to 60 mV, allowing oxidation of a range of aliphatic and aromatic substrates.^{6–8}

Flavin redox reactions are an example of proton-coupled electron transfers or PCET, a broad family of reactions and energy conversion processes in chemistry.^{9–11} PCET mechanisms are characterized by the number of electrons involved, such as in hydride ($2e^-/1H^+$) versus hydrogen-atom ($1e^-/1H^+$, or HAT) transfer;¹² the order of steps or their concurrency, such as electron transfer first, proton second (ETPT), or concerted proton–electron transfer (CPET);^{9,13} whether the transfers proceed from(to) the same or different chemical groups in the donor(acceptor), as in multiple-site PCET;^{10,14} and the tunneling behavior and the adiabaticity or participation of excited-states in the transfer processes.^{13,15}

Thermochemical and redox potential measurements determined that flavins may undergo hydride and single-electron ($1e^-$) transfers,^{10,16} and molecular simulations explored their tunneling effects.^{17–20} However, a description of possible HAT, particularly in enzymatic mechanisms, has received less attention.²¹ A remarkable example is the BLUF flavoprotein

which has been shown by ultrafast spectroscopy and simulations to sustain HAT by light activation.^{22,23}

Quantitative diagnostics of PCET mechanisms can be obtained from quantum-chemical calculations of molecular properties along the reaction progress (usually the proton transfer coordinate).¹³ Nonadiabatic couplings¹⁵ and dipole moments²⁴ are well-behaved properties. However, the calculation of partial charges and spins,²⁵ the simplest and most common diagnostic used,^{13,26–28} can be problematic because these are not physical observables and may be calculated with different recipes.^{29–31} Realizing these difficulties, intrinsic bond orbitals³² were recently proposed and successfully applied to identify proton and electron donors in PCET.^{33,34}

Here, we address the PCET mechanism in flavin-mediated oxidation of NADH (reduced nicotinamide adenine dinucleotide) and of succinate catalyzed respectively by respiratory complex I (or type I NADH dehydrogenase)^{35–37} and by complex II (or succinate dehydrogenase, see Figure 1).^{38,39} These fundamental steps for cellular respiration have been traditionally assigned to hydride transfers,^{36,39} but recent

Received: August 13, 2020

Published: November 2, 2020



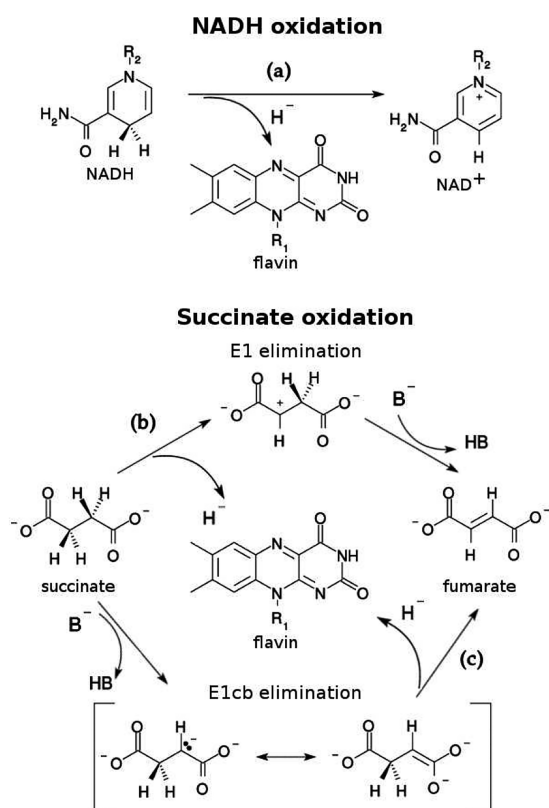


Figure 1. Flavins mediate oxidation of NADH and succinate. NADH reacts by elimination from the nicotinamide ring, step (a). Two reaction sequences are possible for succinate oxidation: unimolecular elimination (E1) is shown in step (b) and conjugate base elimination (E1cb) in step (c). Biologically relevant flavins differ only in the R_1 substituent. R_2 is the ribosyl-ADP moiety, and B^- is a general base.

studies disputed this view and proposed HAT mechanisms.^{26,27}

Computational Methods. The three reactions studied here are models of flavin-mediated oxidations catalyzed by respiratory complexes I and II (Figure 1). The flavin isoalloxazine ring was represented by lumiflavin (LF, IUPAC name 7,8,10-trimethylbenzo[*g*]pteridine-2,4-dione) instead of flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD) natural cofactors. Ribityl-ADP (adenosine diphosphate) and ribityl-phosphate moieties were replaced by methyl groups. The ribosyl-ADP moiety in NADH was replaced by a methyl group, generating 1-methylnicotinamide in the reduced form (MNAH or 1-methyl-4H-pyridine-3-carboxamide, Figure 2). Succinate had its carboxylates protonated into succinic acid for both E1 elimination from H_2Succ (butanedioic acid) and E1cb elimination from H_2Succ^- [(E)-4-hydroxy-4-oxidobut-3-enoic acid]. This is consistent with the active site in respiratory complex II³⁸ that shields the negative charges in succinate carboxylates by hydrogen bonding with positively charged residues and with previous calculations on similar reactions.⁴⁰ These models had 45, 44, and 51 atoms, respectively.

Molecular geometries were optimized at the MP2 level⁴¹ with the def2-SVP basis set.⁴² Resolution of identity with the def2/J⁴³ and def2-SVP/C⁴⁴ auxiliary basis was used. Optimizations with methods neglecting dispersion interactions resulted in disruption of the experimental^{35,38} stacked orientation between reactants. Transition structures were

confirmed by analysis of the Hessian eigenvector with the negative eigenvalue and of the intrinsic reaction coordinates (Table S1, Supporting Information).⁴⁵

A reaction coordinate $RC = r_1 - r_2$ was defined by the broken bond distance (r_1 , between the transferred H and C2 atom in H_2Succ/H_2Succ^- or C4 atom in MNAH) minus the formed bond distance (r_2 , between H and N5 atom in LF, Figure 2). The reaction profile was obtained with constrained geometry optimizations fixing r_1 while relaxing all others degrees of freedom using the ORCA 4.1.1 program.⁴⁶ Intrinsic bond orbitals (IBOs)³² were obtained from localization³³ of the broken-symmetry solution³⁴ of unrestricted M06⁴⁷ single-point calculations with ORCA.

The complete active space self-consistent field (CASSCF)^{28,41} method was used with the PySCF package version 1.5⁴⁸ for singlet spin states in C_1 point-group symmetry. All single-point calculations used the def2-TZVP basis set. Configuration and population analysis on CASSCF wave functions were done with localized orbitals by the Pipek–Mezey scheme, while Foster–Boys localization led to equivalent results.⁴⁹ This analysis was also performed on CASSCF/def2-SVP wave functions, and similar results were observed. Active spaces contained six electrons in five MOs (6e,5o) for H_2Succ , (8e,6o) for H_2Succ^- , and (12e,10o) for MNAH reactions as determined from analysis of occupation numbers and selection of reactive MOs. As MOs may change shape and ordering along the reaction path, CASSCF calculations started from the geometry with highest multi-configurational character (product for H_2Succ and reactant for H_2Succ^- and MNAH reactions), and guess MOs were taken from the CASSCF calculation of an adjacent geometry along the path.

The short-range influence of the enzymatic environment on the mechanism of flavin-mediated oxidation was emulated with truncated active site models (Figure S1, Supporting Information). For succinate oxidation catalyzed by respiratory complex II, the model contained 192 atoms including FAD (for which the ribityl-ADP moiety was replaced by 2-hydroxyethyl), succinate (deprotonated, $Succ$), side chains of H365, T377, E378, R402, H504, and R544, backbones of G170 and G546, and eight water molecules, which compose the first coordination shell around the flavin and substrate groups. The initial model structure was based on the PDB ID 1YOP³⁸ from *F. frigidimarina* fumarate reductase flavoprotein, homologous to the respiratory complex II. This complete enzyme bound to succinate was solvated, neutralized, and equilibrated for 22 ns of classical molecular dynamics simulation, using the CHARMM36 force field.⁵⁰ The listed groups were extracted from this equilibrated structure, and dangling covalent bonds were saturated by hydrogen atoms. Reactive geometries were obtained by relaxed potential energy surface scans along the same reaction coordinate defined above (RC) with position restraints applied to C_α or C_β of capped amino acids to preserve the active site architecture. This corresponds to the H_2Succ reaction in Figure 2B.

For NADH oxidation catalyzed by respiratory complex I, the active site model contained 167 atoms including FMN, NADH (for which phosphate group and ADP moiety were replaced by methyl groups), side chains of residues N92, E97, Y180, N220, S295, and T325, and backbone of residues G183 and E184 from the Nqo1 subunit of *T. thermophilus* complex I as found in the PDB ID 3IAM³⁵ structure. This model corresponds to the MNAH reaction in Figure 2A. Geometries of stationary

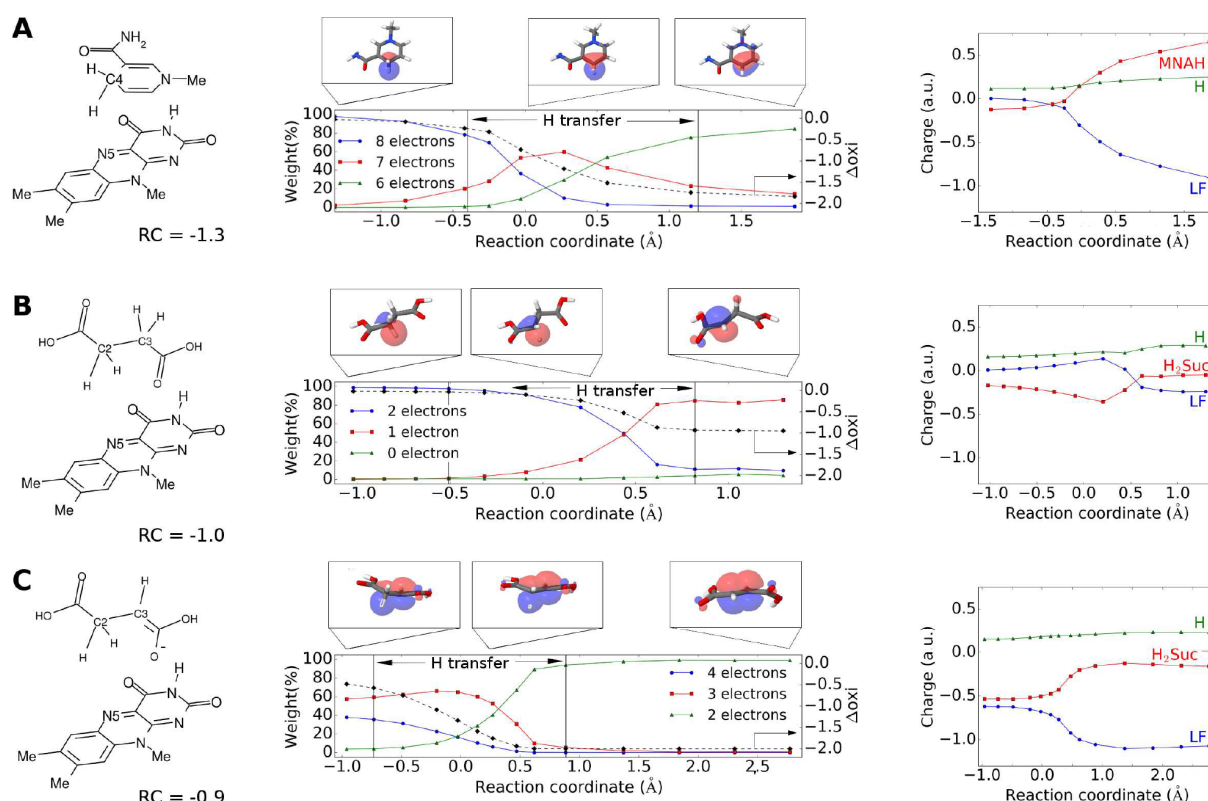


Figure 2. Mechanisms of flavin mediated oxidation from analysis of the wave function along the reaction. Panels A, B, and C show the reactions studied here, corresponding respectively to steps (a), (b), and (c) in Figure 1. The left column presents the reactant states. The middle column shows the contribution or weight of wave function configurations with n -electrons in molecular orbitals (MOs) localized in the donor along the proton-transfer reaction coordinate. Δoxi is calculated from eq 1 and shown in the dashed line. In the H transfer range, the transferred hydrogen is not fully bound to either donor or acceptor. The broken C–H bond σ MO is shown on top, and the lone-pair MO is also shown for H₂Suc⁻. The right column shows Mulliken partial charges for the transferred H, flavin, and donor groups.

points were optimized in the B3LYP-D3/def2-SVP level⁵¹ and kindly provided by Prof. Ville Kaila from their original study.²⁷ Geometries for the complex II model were optimized at the M06-L/def2-SVP level.⁵² Geometries of the enzyme models showed little dependency on the optimization level if dispersion interactions were accounted for (as in B3LYP-D3 or M06-L functionals). Wave functions for stationary points were obtained at the CASSCF/def2-SVP level with the same active spaces used for the isolated reactions.

RESULTS AND DISCUSSION

Three model reactions corresponding to steps (a)–(c) in Figure 1 were studied with a lumiflavin (LF, R_1 = methyl) acceptor, NADH modeled as 1-methylnicotinamide (MNAH), and succinate protonated to H₂Suc (succinic acid) for E1 elimination and H₂Suc⁻ (carbanion ↔ enolate electromer) for E1cb elimination. See panels A, B, and C in Figure 2 for the reactant structures, consistent with enzymatic active sites.

In order to quantify the extent of electron transfer along the reaction and avoid the pitfalls of using only partial charges to analyze PCET mechanisms,^{29,33} we propose a simple quantitative diagnostic of electron transfer employing weights of wave function configurations constructed with orbitals localized in the donor. This electronic charge transferred, Δoxi , is defined along the reaction coordinate (RC) as

$$\Delta\text{oxi}(\text{RC}) = \sum_n n w_n(\text{RC}) - \max\{n\} \quad (1)$$

where n is the number of electrons occupying active MOs localized in the donor, $\max\{n\}$ is their maximum occupation at the reactant state, and w_n is the combined weight of wave function configurations with n -electrons. Only configurations with $w_n > 1\%$ in at least one geometry along the reaction were considered. This analysis may be used with any kind of multiconfigurational wave function obtained with localized MOs.⁴¹

We start with the H₂Suc reaction (model for succinate E1 elimination, Figure 2B) in which the broken C–H bond σ MO becomes a nonbonding MO in the product. This is the only active MO localized in the H₂Suc donor, thus $\{n\} = (0, 1, 2)$ electrons and $\max\{n\} = 2$. Figure 2B middle column shows that the contribution (combined weight) to the wave function of this localized MO changes from almost 100% for two-electron occupation in the reactant to 80% for one-electron occupation in the product. The average electronic charge transferred is $\Delta\text{oxi} = -1$, showing that the H₂Suc reaction is a hydrogen-atom transfer (HAT).

In the biradical product, an unpaired electron occupies the LF π^* MO and spin-couples to a singlet state with the unpaired electron in the carboxylate (Figure 2B, top of middle column), resulting in a short distance ($r_1 = 2.3$ Å) between the two molecules. Interestingly, this is in line with experimental

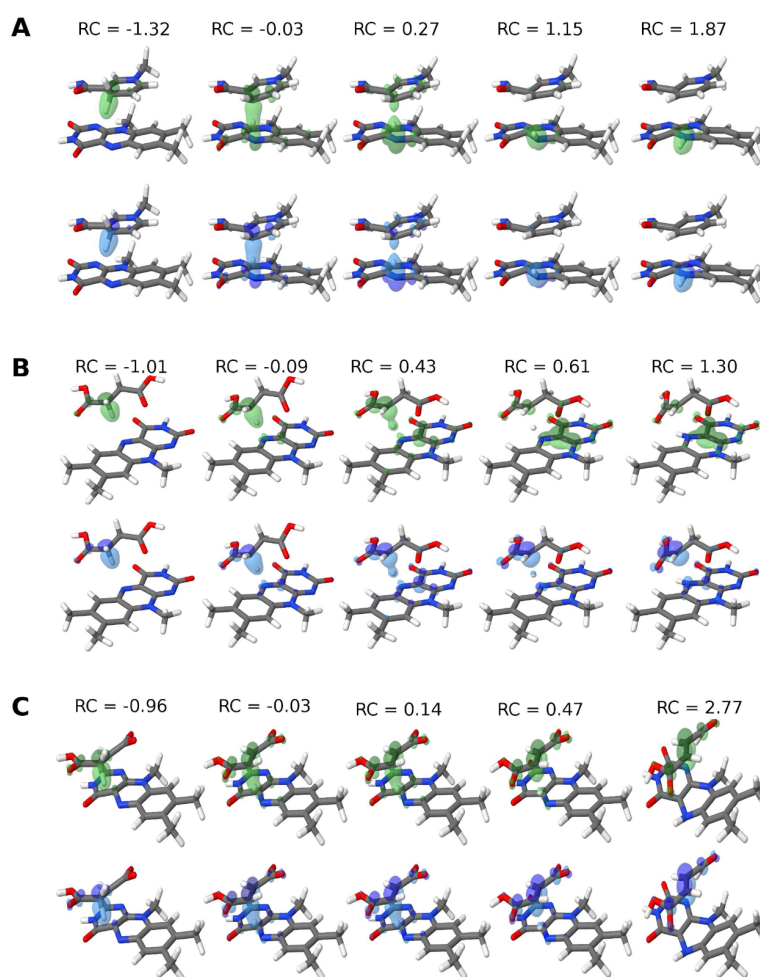


Figure 3. Intrinsic bond orbitals of the broken C–H bond for geometries with a given reaction coordinate (RC) along the pathway for MNAH (A), H₂Suc (B), and H₂Suc[−] (C) oxidation. Green α and blue β spin-orbitals are shown.

observation of a FAD flavosemiquinone radical signal during the catalytic cycle of respiratory complex II.⁵³

Partial charges condensed on the LF and H₂Suc groups are similar (within 0.2e, Figure 2B, right column) and nearly neutral when comparing reactant and product (end) states, suggesting there is no net charge separation or redistribution, in accordance to a HAT.^{13,15} The partial charge on the transferred H atom is close to 0.2e along the complete pathway for all reactions studied here (Figure 2, right column). A naive interpretation of this partial charge would suggest a HAT is observed in all reactions, but this is not the case as clearly shown for the other two reactions below.

For H₂Suc[−] (E1cb elimination, Figure 2C), the nonbonding MO with the extra electron lone-pair and the σ MO of the broken C–H bond localize in the donor. Only configurations with $\{n\} = (2, 3, 4)$ electrons in these orbitals contribute significantly (more than 1%) to the wave function, thus $\max\{n\} = 4$. Remarkably, the anionic reactant already shows a substantial charge transfer ($\Delta\text{oxi} = -0.5$ at RC = -1.0 Å). When fumaric acid is formed, only configurations with two electrons in the forming double bond become relevant, resulting in a much longer distance ($r_1 = 3.8$ Å) between LF[−] and fumaric acid. The average $\Delta\text{oxi} = -2$ in the product, and this reaction is clearly a hydride transfer. Partial charges

change considerably between end states, and LF has a $-1e$ charge in the product, as expected from receiving a hydride.

For the MNAH reaction (Figure 2A), the MO of the broken C–H bond becomes a π MO and conjugates with the other nicotinamide MOs. A total of six active MOs are localized in the donor. Only configurations with $\{n\} = (6, 7, 8)$ electrons in these orbitals contribute significantly to the wave function, thus $\max\{n\} = 8$. Weights change from almost 100% for configurations with eight electrons in the reactant state to 85% for six electrons in the product, with a final $\Delta\text{oxi} \simeq -2$. We conclude this reaction is also a hydride transfer.

It is noteworthy that weights for configurations with an odd number of electrons localized in the donor (one electron for H₂Suc, three for H₂Suc[−], and seven for MNAH reactions) peak near the transition state (TS, RC = -0.03 Å for MNAH, 0.61 Å for H₂Suc, and 0.14 Å for H₂Suc[−], Table S1). This is in line with $\Delta\text{oxi} \simeq -1$ at the TS for the three reactions (Figure 2, middle column) and suggests that the activation energies correspond to the first (and only in H₂Suc reaction) electron transfer. Partial charge or other population analysis^{30–32} cannot usually provide this level of detail.

Electron and proton transfer occur concerted along the same range of reaction coordinates for the three studied reactions (H transfer indicated in Figure 2). However, analysis of intrinsic bond orbitals (IBO,^{32,33} Figure 3A) suggests that the

two transfers take place from (or to) the same donor (acceptor) site only for the MNAH reaction as orbitals (and the corresponding electron density) from the C–H bond in the reactant turn into the N5–H bond in the product after H^- transfer. On the other hand, electron and proton transfer in the two succinic acid reactions proceed to different sites of the acceptor flavin. These two reactions should be classified as concerted PCET^{11,13} or multiple-site PCET¹⁴ as orbitals of the broken C–H bond in the reactant do not turn into the N5–H bond in the product. For the H_2 Suc reaction, the one electron transferred from the broken C–H bond to the flavin delocalizes over the pyrimidinedione ring (Figure 3B). For the H_2 Suc[−] reaction, the broken C–H bond rearranges to form the double bond in fumaric acid (Figure 3C), and the two electrons transferred to flavin come from the nonbonding MO with the extra electron lone-pair (Figure S2).

Finally, we tested whether the mechanisms observed above are conserved in the enzymatic environment. Truncated active site models with 167 and 192 atoms, respectively, coordinating NADH in respiratory complex I and succinate in respiratory complex II were built from these protein crystallographic structures (Figure S1).^{35,37,38} Analysis of Δoxi (eq 1) calculated for these active site models using the same methods depicted above shows NADH oxidation proceeds via hydride transfer, and succinate oxidation in E1 elimination occurs through HAT (Table S2), showing that the enzymatic mechanisms are equivalent to those of the isolated models.

CONCLUSIONS

We confirm the traditional view³⁶ that NADH oxidation by flavin proceeds via a prototypical hydride transfer, with the two electrons moving concerted with the proton from the donor to the same acceptor group. A previous proposal of HAT for this reaction²⁷ is incorrect and illustrates the pitfalls of assigning PCET mechanisms with diagnostics based only on partial charges. For succinate oxidation by flavin, two reaction sequences are possible. E1 elimination may occur by a HAT, opposed to usual proposals,³⁹ but E1cb elimination will also proceed via hydride with an advanced charge transfer in the reactant state.

PCET mechanisms of flavin-mediated oxidation depend on the donor molecule. This should have implications for the mechanisms of several flavoproteins besides those of the respiratory chain studied here. It is also expected that the simple diagnostic of electron transfer along the reaction pathway proposed here will be useful to characterize a broad class of PCET processes.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jcim.0c00945>.

Two figures with the structure of truncated enzyme active site models and additional IBO analysis and two tables with quantification of electron transfer (Δoxi) for enzyme models and TS analysis. (PDF)

AUTHOR INFORMATION

Corresponding Author

Guilherme M. Arantes — Department of Biochemistry, Instituto de Química, Universidade de São Paulo, 05508-900 São Paulo,

SP, Brazil; orcid.org/0000-0001-5356-7703;
Email: garantes@iq.usp.br

Author

Felipe Curtolo — Department of Biochemistry, Instituto de Química, Universidade de São Paulo, 05508-900 São Paulo, SP, Brazil; orcid.org/0000-0002-4459-0968

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.jcim.0c00945>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Fruitful discussions with Prof. Peter R. Taylor and funding from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, scholarships 2016/23525-0 and 2017/26109-0 to F.C. and Grants 2016/24096-5 and 2019/21856-7 to G.M.A.) are gratefully acknowledged.

REFERENCES

- (1) Joosten, V.; van Berkel, W. J. Flavoenzymes. *Curr. Opin. Chem. Biol.* **2007**, *11*, 195–202.
- (2) Sobrado, P.; Gadda, G. Introduction to Flavoproteins: Beyond the Classical Paradigms. *Arch. Biochem. Biophys.* **2017**, *632*, 1–3.
- (3) Gray, M.; Goodman, A. J.; Carroll, J. B.; Bardon, K.; Markey, M.; Cooke, G.; Rotello, V. M. Model Systems for Flavoenzyme Activity: Interplay of Hydrogen Bonding and Aromatic Stacking in Cofactor Redox Modulation. *Org. Lett.* **2004**, *6*, 385–388.
- (4) Nandwana, V.; Samuel, I.; Cooke, G.; Rotello, V. M. Aromatic Stacking Interactions in Flavin Model Systems. *Acc. Chem. Res.* **2013**, *46*, 1000–1009.
- (5) Reis, A. A. O.; Sayegh, R. S. R.; Marana, S. R.; Arantes, G. M. Combining Free Energy Simulations and NMR Chemical-Shift Perturbation To Identify Transient Cation- π Contacts in Proteins. *J. Chem. Inf. Model.* **2020**, *60*, 890–897.
- (6) Chaiyen, P.; Scrutton, N. S. Flavins and Flavoproteins. *FEBS J.* **2015**, *282*, 3001–3002.
- (7) Bresnahan, C. G.; Reinhardt, C. R.; Bartholow, T. G.; Rumpel, J. P.; North, M.; Bhattacharyya, S. Effect of Stacking Interactions on the Thermodynamics and Kinetics of Lumiflavin: A Study with Improved Density Functionals and Density Functional Tight-Binding Protocol. *J. Phys. Chem. A* **2015**, *119*, 172–182.
- (8) Buckel, W.; Thauer, R. K. Flavin-Based Electron Bifurcation, A New Mechanism of Biological Energy Coupling. *Chem. Rev.* **2018**, *118*, 3862–3886.
- (9) Cukier, R. I.; Nocera, D. G. Proton-Coupled Electron Transfer. *Annu. Rev. Phys. Chem.* **1998**, *49*, 337–369.
- (10) Huynh, M. H. V.; Meyer, T. J. Proton-Coupled Electron Transfer. *Chem. Rev.* **2007**, *107*, 5004–5064.
- (11) Warren, J. J.; Tronic, T. A.; Mayer, J. M. Thermochemistry of Proton-Coupled Electron Transfer Reagents and its Implications. *Chem. Rev.* **2010**, *110*, 6961–7001.
- (12) Hammes-Schiffer, S. Comparison of Hydride, Hydrogen Atom, and Proton-Coupled Electron Transfer Reactions. *ChemPhysChem* **2002**, *3*, 33–42.
- (13) Hammes-Schiffer, S. Proton-Coupled Electron Transfer: Classification Scheme and Guide to Theoretical Methods. *Energy Environ. Sci.* **2012**, *5*, 7696–7703.
- (14) Darcy, J. W.; Koronkiewicz, B.; Parada, G. A.; Mayer, J. M. A Continuum of Proton-Coupled Electron Transfer Reactivity. *Acc. Chem. Res.* **2018**, *51*, 2391–2399.
- (15) Hammes-Schiffer, S.; Stuchebrukhov, A. A. Theory of Coupled Electron and Proton Transfer Reactions. *Chem. Rev.* **2010**, *110*, 6939–6960.

- (16) Brenner, S.; Hay, S.; Heyes, D. J.; Scrutton, N. S. *Proton-Coupled Electron Transfer: A Carrefour of Chemical Reactivity Traditions*; The Royal Society of Chemistry, 2012; pp 57–88.
- (17) Poulsen, T. D.; Garcia-Viloca, M.; Gao, J.; Truhlar, D. G. Free Energy Surface, Reaction Paths, and Kinetic Isotope Effect of Short-Chain Acyl-CoA Dehydrogenase. *J. Phys. Chem. B* **2003**, *107*, 9567–9578.
- (18) Bhattacharyya, S.; Stankovich, M. T.; Truhlar, D. G.; Gao, J. Combined Quantum Mechanical and Molecular Mechanical Simulations of One- and Two-Electron Reduction Potentials of Flavin Cofactor in Water, Medium-Chain Acyl-CoA Dehydrogenase, and Cholesterol Oxidase. *J. Phys. Chem. A* **2007**, *111*, 5729–5742.
- (19) Pang, J.; Hay, S.; Scrutton, N. S.; Sutcliffe, M. J. Deep Tunneling Dominates the Biologically Important Hydride Transfer Reaction from NADH to FMN in Morphinone Reductase. *J. Am. Chem. Soc.* **2008**, *130*, 7092–7097.
- (20) Delgado, M.; Görlisch, S.; Longbotham, J. E.; Scrutton, N. S.; Hay, S.; Moliner, V.; Tuñón, I. Convergence of Theory and Experiment on the Role of Preorganization, Quantum Tunneling, and Enzyme Motions into Flavoenzyme-Catalyzed Hydride Transfer. *ACS Catal.* **2017**, *7*, 3190–3198.
- (21) Tan, S. L. J.; Novianti, M. L.; Webster, R. D. Effects of Low to Intermediate Water Concentrations on Proton-Coupled Electron Transfer (PCET) Reactions of Flavins in Aprotic Solvents and a Comparison with the PCET Reactions of Quinones. *J. Phys. Chem. B* **2015**, *119*, 14053–14064.
- (22) Mathes, T.; Zhu, J.; van Stokkum, I. H. M.; Groot, M. L.; Hegemann, P.; Kennis, J. T. M. Hydrogen Bond Switching among Flavin and Amino Acids Determines the Nature of Proton-Coupled Electron Transfer in BLUF Photoreceptors. *J. Phys. Chem. Lett.* **2012**, *3*, 203–208.
- (23) Goings, J. J.; Hammes-Schiffer, S. Early Photocycle of Slr1694 Blue-Light Using Flavin Photoreceptor Unraveled through Adiabatic Excited-State Quantum Mechanical/Molecular Mechanical Dynamics. *J. Am. Chem. Soc.* **2019**, *141*, 20470–20479.
- (24) Soudackov, A. V.; Hammes-Schiffer, S. Probing Non-adiabaticity in the Proton-Coupled Electron Transfer Reaction Catalyzed by Soybean Lipoygenase. *J. Phys. Chem. Lett.* **2014**, *5*, 3274–3278.
- (25) Sirjoosingh, A.; Hammes-Schiffer, S. Proton-Coupled Electron Transfer versus Hydrogen Atom Transfer: Generation of Charge-Localized Diabatic States. *J. Phys. Chem. A* **2011**, *115*, 2367–2377.
- (26) Abe, Y.; Shoji, M.; Nishiya, Y.; Aiba, H.; Kishimoto, T.; Kitaura, K. The Reaction Mechanism of Sarcosine Oxidase Elucidated Using FMO and QM/MM Methods. *Phys. Chem. Chem. Phys.* **2017**, *19*, 9811–9822.
- (27) Saura, P.; Kaila, V. R. I. Energetics and Dynamics of Proton-Coupled Electron Transfer in the NADH/FMN Site of Respiratory Complex I. *J. Am. Chem. Soc.* **2019**, *141*, 5710–5719.
- (28) 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.
- (29) Politzer, P.; Harris, R. R. Properties of Atoms in Molecules. I. Proposed Definition of the Charge on an Atom in a Molecule. *J. Am. Chem. Soc.* **1970**, *92*, 6451–6454.
- (30) Reed, A. E.; Weinstock, R. B.; Weinhold, F. Natural Population Analysis. *J. Chem. Phys.* **1985**, *83*, 735–746.
- (31) Martin, F.; Zipse, H. Charge Distribution in the Water Molecule - A Comparison of Methods. *J. Comput. Chem.* **2005**, *26*, 97–105.
- (32) Knizia, G. Intrinsic Atomic Orbitals: An Unbiased Bridge between Quantum Theory and Chemical Concepts. *J. Chem. Theory Comput.* **2013**, *9*, 4834–4843.
- (33) Klein, J. E. M. N.; Knizia, G. cPCET versus HAT: A Direct Theoretical Method for Distinguishing X-H Bond-Activation Mechanisms. *Angew. Chem., Int. Ed.* **2018**, *57*, 11913–11917.
- (34) Mandal, M.; Elwell, C. E.; Bouchev, C. J.; Zerk, T. J.; Tolman, W. B.; Cramer, C. J. Mechanisms for Hydrogen-Atom Abstraction by Mononuclear Copper(III) Cores: Hydrogen-Atom Transfer or Concerted Proton-Coupled Electron Transfer? *J. Am. Chem. Soc.* **2019**, *141*, 17236–17244.
- (35) Berrisford, J. M.; Sazanov, L. A. Structural Basis for the Mechanism of Respiratory Complex I. *J. Biol. Chem.* **2009**, *284*, 29773–29783.
- (36) Hirst, J. Mitochondrial Complex I. *Annu. Rev. Biochem.* **2013**, *82*, 551–575.
- (37) Teixeira, M. H.; Arantes, G. M. Balanced Internal Hydration Discriminates Substrate Binding to Respiratory Complex I. *Biochim. Biophys. Acta, Bioenerg.* **2019**, *1860*, 541–548.
- (38) Wardrope, C.; Mowat, C. G.; Walkinshaw, M. D.; Reid, G. A.; Chapman, S. K. Fumarate Reductase: Structural and Mechanistic Insights from the Catalytic Reduction of 2-methylfumarate. *FEBS Lett.* **2006**, *580*, 1677–1680.
- (39) Maklashina, E.; Cecchini, G.; Dikanov, S. A. Defining a Direction: Electron Transfer and Catalysis in Escherichia coli Complex II Enzymes. *Biochim. Biophys. Acta, Bioenerg.* **2013**, *1827*, 668–678.
- (40) Lucas, M. F.; Ramos, M. J. Mechanism of a Soluble Fumarate Reductase from *Shewanella frigidimarina*: A Theoretical Study. *J. Phys. Chem. B* **2006**, *110*, 10550–10556.
- (41) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*, 1st ed.; Wiley: New York, 2000.
- (42) 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.
- (43) Weigend, F. Accurate Coulomb-Fitting Basis Sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.
- (44) Hellweg, A.; Hättig, C.; Höfener, S.; Klopper, W. Optimized Accurate Auxiliary Basis Sets for RI-MP2 and RI-CC2 Calculations for the Atoms Rb to Rn. *Theor. Chem. Acc.* **2007**, *117*, 587–597.
- (45) Fukui, K. The Path of Chemical Reactions - The IRC Approach. *Acc. Chem. Res.* **1981**, *14*, 363–368.
- (46) Neese, F. Software Update: The ORCA Program System, Version 4.0. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, No. e1327.
- (47) 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.
- (48) Sun, Q.; Berkelbach, T. C.; Blunt, N. S.; Booth, G. H.; Guo, S.; Li, Z.; Liu, J.; McClain, J. D.; Sayfutyarova, E. R.; Sharma, S.; Wouters, S.; Chan, G. K.-L. PySCF: The Python-based Simulations of Chemistry Framework. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, No. e1340.
- (49) Pipek, J.; Mezey, P. A Fast Intrinsic Localization Procedure Applicable for ab Initio and Semiempirical Linear Combination of Atomic Orbital Wave Functions. *J. Chem. Phys.* **1989**, *90*, 4916–4926.
- (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) Becke, A. D. Density Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648.
- (52) 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.
- (53) Maklashina, E.; Iverson, T. M.; Sher, Y.; Kotlyar, V.; Andréll, J.; Mirza, O.; Hudson, J. M.; Armstrong, F. A.; Rothery, R. A.; Weiner, J. H.; Cecchini, G. 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.* **2006**, *281*, 11357–11365.

Supporting Information

Mechanisms for Flavin Mediated Oxidation: Hydride or Hydrogen-Atom Transfer?

Felipe Curtolo 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

E-mail: garantes@iq.usp.br

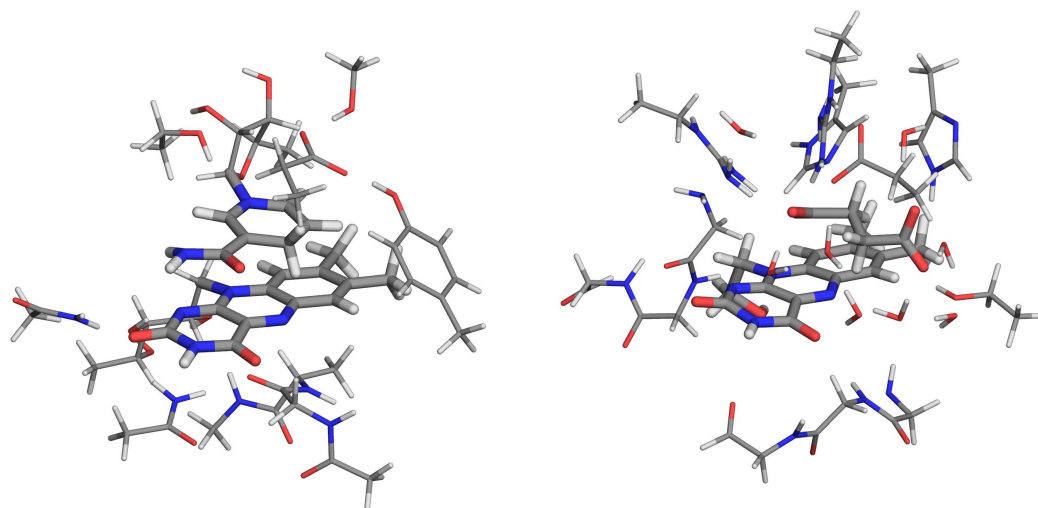


Figure S1: Structure of truncated active site models for respiratory complexes I (left) and II (right) used for calculations in Table S2. Flavin, the nicotinamide ring of NADH (left) and succinate (right) are shown in thick sticks.

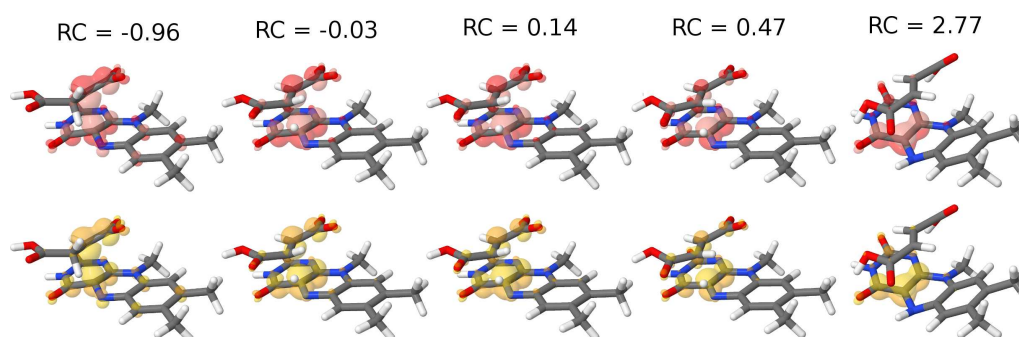


Figure S2: Intrinsic bond orbitals of the 2-electrons transferred to flavin for geometries with a given reaction coordinate (RC) along the pathway for H_2Suc^- oxidation. Pink α and yellow β spin-orbitals are shown.

Table S1: Reaction coordinates and mode component analysis for transition states (TS) of flavin mediated oxidation.

Donor	r_1 (Å)	r_2 (Å)	RC (Å)	Frequency (cm^{-1})
MNAH	1.30	1.33	-0.03	-859.2
H ₂ Suc	1.68	1.07	0.61	-827.1
H ₂ Suc ⁻	1.40	1.26	0.14	-947.0

Table S2: Average number of electrons transferred (Δoxi , equation 1 in the main text) and reaction coordinate (RC in Å) for stationary points of flavin mediated oxidation in active site models of the respiratory complex I reaction with NADH and complex II reaction with succinate via E1 elimination (Fig. S1).

	Complex I + NADH		Complex II + succinate	
	Δoxi	RC	Δoxi	RC
Reactant	-0.06	-1.00	-0.02	-1.67
TS	-0.93	0.00	-1.00	1.36
Product	-1.92	1.86	-1.00	1.70

4 Mechanism of flavin oxidation catalyzed by fumarate reductases

Felipe Curtolo and Guilherme M. Arantes

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

Mechanism of flavin oxidation catalyzed by fumarate reductases

Felipe Curtolo 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

E-mail: garantes@iq.usp.br

Abstract

The enzymatic reduction of fumarate to succinate is central to energy metabolism. Flavoenzymes that catalyze this reaction are ubiquitous and employ a mechanism with sequential hydride and proton transfers. However, it is still unknown if intermediates are involved. Here, electronic structure calculations were used to probe the reaction mechanism of fumarate reduction mediated by flavin in a large cluster model of the *S. frigidimarina* soluble fumarate reductase. A benchmark study was performed to show that static correlation effects are small and density functional methods can be safely used. In particular, the ω B97X-D3 functional has average errors of 1.4 kcal.mol⁻¹ in comparison to coupled-cluster reference calculations. This density functional method was applied to reduction mechanisms with carbanions, adducts and radical intermediates, and details of the reaction were revealed: the lowest activation energies were obtained for carbanion intermediates, discarding the other possible mechanisms; fumarate twisting and protonation are not critical for catalysis, in opposition to previous proposals; hydride transfer is vibronically adiabatic and tunneling contributions are small. Thus catalysis is driven by electrostatic interactions. The comprehensive mechanistic description of fumarate reduction provided here clarifies contradictory views and gives new insights into catalysis by essential flavoenzyme reductases and dehydrogenases involved in cellular respiration.

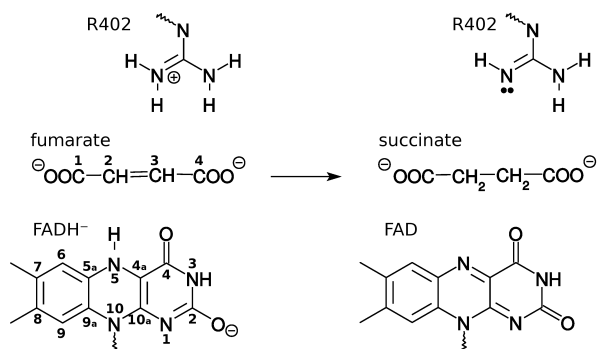
Introduction

Fumarate reductases and succinate dehydrogenases are homologous flavoenzymes that catalyze the interconversion of fumarate to succinate in anaerobic and aerobic respiration, respectively.¹ Even though some reductases are soluble,²⁻⁴ all succinate dehydrogenases⁵⁻⁸ and most fumarate reductases⁹⁻¹¹ are associated to large protein complexes embedded in membranes. The main difference between them is that soluble fumarate reductases cannot oxidize succinate,¹² whereas the membrane-bound enzymes, which have a covalently bound FAD,¹³ can work in both directions oxidizing succinate and reducing fumarate.^{14,15}

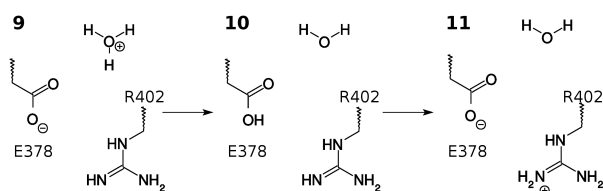
The residues that form the active site of these enzymes are completely conserved,^{1,2} indicating an ubiquitous reaction mechanism. Because soluble fumarate reductases are easier to manipulate than the membrane-associated complexes, most studies that address the enzymatic mechanism of fumarate reduction were conducted in the soluble fumarate reductase flavocytochrome *c*₃ (Fcc₃) of *Shewanella frigidimarina*.^{2,12,16-22} In fact, the investigation of the mechanisms of fumarate reduction in Fcc₃ was intended not only to characterize this enzyme, but also to provide an understanding of the mechanism of succinate oxidation in respiratory complex II.¹

Fumarate reduction is an addition reaction, where two protons and two electrons are added to the C=C double bond of fumarate

to generate succinate (Scheme 1). In the enzyme, a hydride equivalent is transferred from FADH^- nitrogen 5 to fumarate carbon 2,^{2,22} with the other proton coming from a catalytic arginine residue^{17,18,23} (Arg402 in the numbering of Fcc3 from *S. frigidimarina*, that will be used for residue numbering hereafter). Because arginines are poor proton donors,²⁴ a Grotthus-like proton transfer mechanism involving conserved residues and water is suggested to allow fast proton exchange between active site and solvent²¹ (Scheme 2).



Scheme 1: Fumarate reduction, with respective atom numbering. Only the flavin ring of FADH^- and the side-chain of Arg402 are shown.



Scheme 2: Mechanism proposed to deliver protons from solvent into the enzyme active site. The numbers **9**, **10** and **11** label reactive intermediates.

The sequence in which the hydride equivalent from FADH^- and the proton from Arg402 are transferred to fumarate is not understood.²⁵ Kinetic isotope effects show disputable views^{26–28} between a concerted mechanism, where proton and hydride are transferred in a single step, and a stepwise mechanism, where proton and hydride transfers occur separately. The observation of an “intermediate-like” molecule in the active

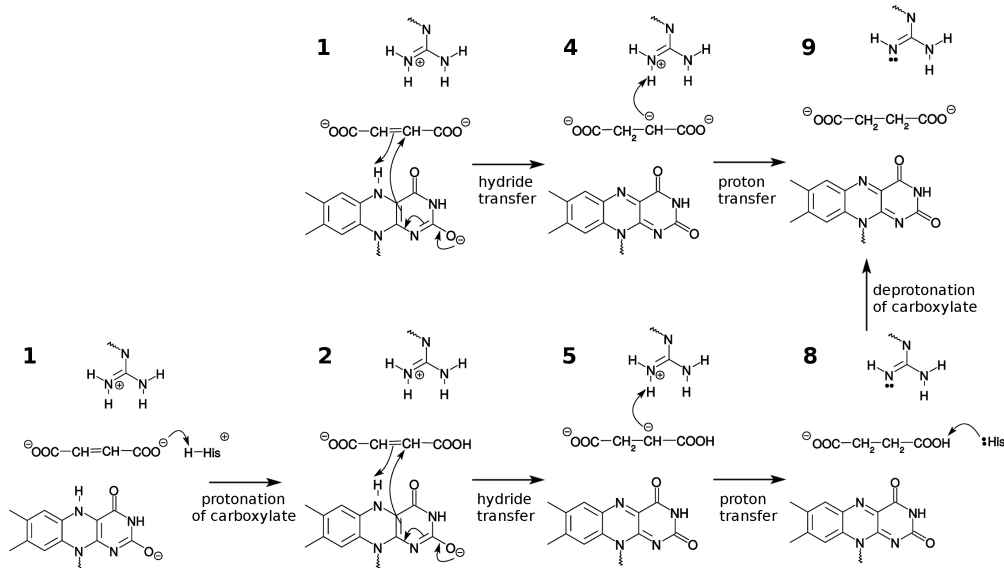
site of some crystallographic structures of fumarate reductases, was interpreted as evidence of a stepwise mechanism.^{2,7} The mechanism most accepted is the reduction of fumarate in two steps, with hydride transfer from FADH^- first, generating a carbanion as reaction intermediate¹ (Scheme 3, reaction sequence **1** \rightarrow **4** \rightarrow **9**).

This mechanism was probed in the computational study performed by Lucas and Ramos,²⁹ employing minimal model of the active site. The barrier found to form a carbanion intermediate was compatible to experimental kinetic data.¹⁷ The authors also argued that hydride transfer is rate-limiting and that the enzyme employs general acid catalysis, with His504 protonating carboxylate 4 of fumarate prior to carbanion formation³⁰ (Scheme 3, reaction sequence **1** \rightarrow **2** \rightarrow **5** \rightarrow **8** \rightarrow **9**). To our knowledge, this is the only computational study investigating the reaction mechanisms of fumarate reductases or succinate dehydrogenases.

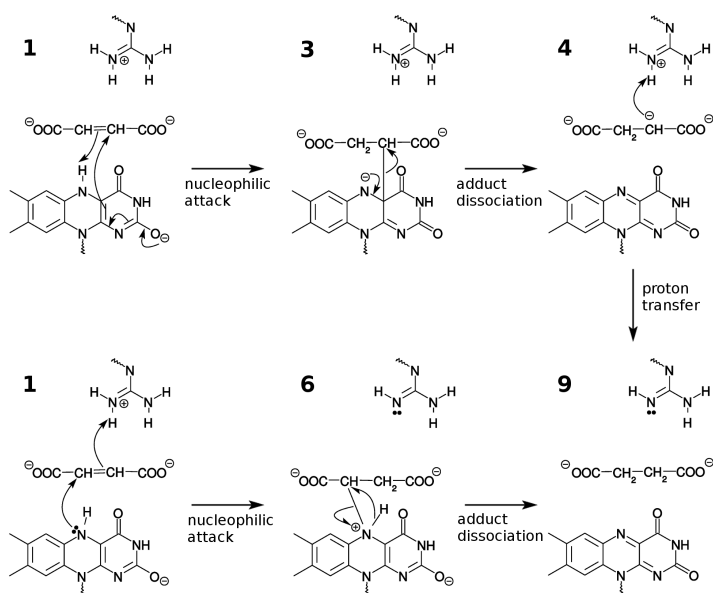
During transfer of a hydride equivalent from FADH^- to fumarate, proton and electrons may not be transferred together as a prototypical hydride transfer, but asynchronously and to different parts of the molecule.²⁵ Specifically, when the carbanion intermediate is formed, its electron density can be shared with FAD and lead to adduct formation (Scheme 3, reaction **1** \rightarrow **3** \rightarrow **4** \rightarrow **9**). In fact, this mechanism was observed in similar reactions and proposed to take place in fumarate reduction.³¹

Proton transfer from Arg402 before complete fumarate reduction is also possible. This proton addition to fumarate $\text{C}=\text{C}$ bond generates a radical intermediate instead of a carbocation, with an electron from FADH^- accompanying in a proton-coupled electron transfer reaction²⁵ (Scheme 3, reaction **1** \rightarrow **7** \rightarrow **9**). Alternatively, proton transfer from Arg402 might occur simultaneously to a nucleophilic attack of FADH^- forming an adduct intermediate instead of a radical (Scheme 3 – Reaction: **1** \rightarrow **6** \rightarrow **9**). A similar intermediate was recently observed in a flavoenzyme of the fumarate reductase/succinate dehydrogenase family³²

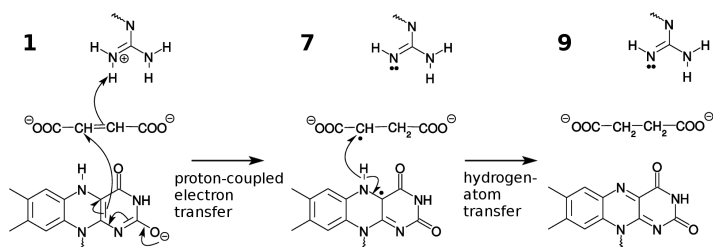
(A) Nucleophilic addition by direct hydride transfer



(B) Nucleophilic addition with adduct formation



(C) Radical addition



Scheme 3: Proposed mechanisms for enzymatic reduction of fumarate. Reaction intermediates are labeled from **1** to **9**.

Here, these mechanistic proposals are tested using electronic structure calculations in order to answer three questions: What is

the mechanism of enzymatic reduction of fumarate? What enzymatic features drive catalysis? Do tunneling facilitate hydride

transfer? The last question is relevant because hydride tunneling was already shown to be non negligible in other flavoenzymes.^{33,34}

Computational methods

Model construction: Two cluster models, named small cluster and large cluster, were built based on the crystal structure of Fcc₃ from *S. frigidimarina* bound to a malate-like intermediate with 1.8 Å resolution.² The small cluster consisted of fumarate, anionic flavoquinol FADH⁻, the catalytic Arg402 and the first solvation shell around fumarate, that is, side-chains of His365, Thr377, His504, Arg544 and backbones of Gly170, Glu378 and Gly547. The amino acid residues and FADH⁻ were truncated as shown in Figure 1a and 1b. The large cluster (Fig. 1c) is an expansion of the small cluster with larger truncation region and a second solvation shell formed by Ala168, Ala169, Gly171, Asn173, Met236, Thr336, Gln363, Pro366, Thr367, Met375, Val376, Ala379, Val380, Arg381, Tyr435, Leu545, Gly546, Asn548, Ala549, Ile550 and nine structural waters. To maintain the active site architecture, some atoms were kept fixed (indicated by * in Fig. 1). In total the small cluster model comprised 121 atoms and the large cluster comprised 475 atoms, both with zero net charge.

There are eight ionizable groups in the models: His365, Glu378, Arg381, Arg402, Tyr435, His504, His505, Arg544 and the prosthetic group FADH⁻. Among them, His365, Arg402, His504, Arg544 and FADH⁻ are directly involved in catalysis or substrate binding,¹ and have already been studied.^{12,17,35,36} Fully reduced FAD was shown to be in the anionic form FADH⁻ under neutral pH for both soluble¹² and membrane-bound³⁶ members of the succinate dehydrogenase/fumarate reductase family, with protonation most likely at nitrogens N3 and N5.³⁷ A pK_a of 7.4 was assigned to His504,¹⁷ suggesting it is positively charged during enzyme turnover. His365 is believed to be neutral, because it can be mutated to a non-ionizable serine residue without loss of fumarate reductase activity.³⁵ Chapman et al.¹⁷ even argued based on the

active site topology that His365 is protonated at N ϵ because N δ receives a hydrogen bond from a nearby peptide bond. Arginine residues were all modeled in their charged state because deprotonated arginines are very unlikely to occur inside proteins under neutral pH.³⁸ Glu378 is salt-bridged by the side-chains of Arg381 and Arg402,²¹ which might lower its pK_a stabilizing the anionic form. Finally, Tyr435 and His505, that are part of the second solvation shell, were modeled in their neutral forms with His505 protonated at N δ due to a favorable hydrogen bond contact. Fumarate was treated fully deprotonated with charge -2.

Calculation settings: Nearly all single-point energy calculations and geometry optimizations were performed with ORCA version 4.1.1.^{39,40} In all calculations, the protein environment beyond the cluster model was treated as an implicit solvent with the Conductor-like Continuum Polarization Model (C-PCM)⁴¹ and a dielectric constant $\epsilon = 4.0$,⁴² unless stated otherwise.

Geometries were optimized with density functional theory (DFT) using the composite method PBEh-3c⁴³⁻⁴⁶ and increased integration grids (keyword "Grid5", in ORCA). For the small cluster, default geometry convergence settings were employed, but for the large cluster, looser convergence criteria ("LooseOpt" keyword) had to be employed because of its size. Additionally, as the large cluster has 465 atoms and more flexibility, for each configuration that was optimized, we visually inspected if the environment remained in the same configuration after each optimization by checking all hydrogen bonds. Transition state (TS) optimizations were performed only for the small cluster as the Hessian matrix calculation is prohibitively costly for the large cluster. Approximate transition state geometries for the large cluster were generated by constrained geometry optimizations, restraining the Cartesian coordinates of the reactive atoms in the same position obtained in the small cluster. To check that the TS obtained for the small clusters are correct, intrinsic reaction coordinate analysis⁴⁷ were performed.

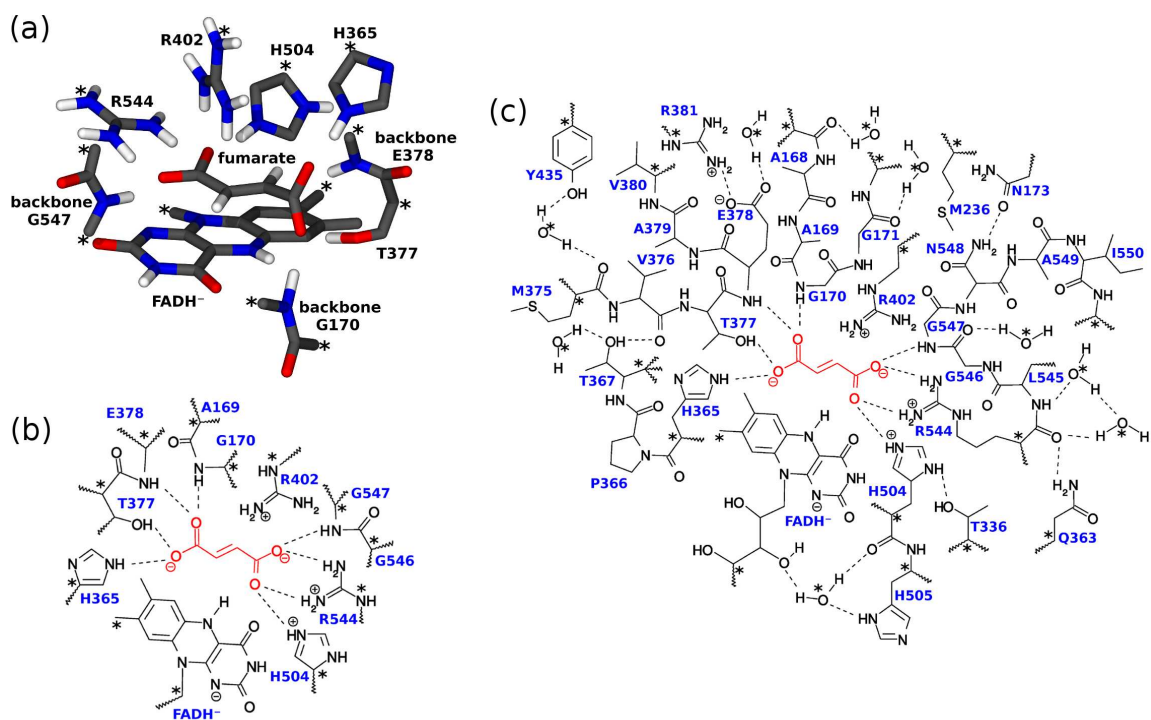


Figure 1: Cluster models of Fcc₃ from *S. frigidimarina* employed here. All atoms pointed with * were kept fixed in space to maintain the active-site architecture. (a) 3D orientation of the small cluster, showing the residues and cofactors that hydrogen-bond to fumarate. Only polar hydrogens are shown. (b) 2D representation of the small cluster with fumarate shown in red. (c) 2D representation of the large cluster with fumarate in red.

To benchmark the reaction, reference energies were obtained with coupled-cluster DLPNO-CCSD(T)⁴⁸ and a complete basis-set (CBS) extrapolation^{49–51} from cc-pVDZ to cc-pVTZ.⁵² For domain-based local-pair approximation (DLPNO), auxiliary cc-pVTZ/C⁵³ and normal thresholds (“NormalPNO” keyword) were used. Additionally, the resolution of identity approximation was employed to fit both Coulomb and exchange integrals.⁵⁴

Multireference calculations were performed with complete active space self-consistent field (CASSCF)⁵⁵ followed by perturbation treatment with partially contracted *n*-electron valence state perturbation theory NEVPT2.^{56,57} Energies at the CBS limit were calculated with two-point extrapolations⁵⁸ from cc-pVDZ and cc-pVTZ. We also tested the single-reference Møller–Plesset perturbation theory MP2.⁵⁵ Both perturbational treatments used the frozen-core approximation and the integrals were fitted with cc-pVTZ/C.⁵³

For DFT single-point energy calculations, basis sets def2-TZVPP and def2-QZVPP⁵⁹ were used with increased integration grids (“Grid6”) and the resolution of identity approximation for Coulomb integrals.⁶⁰ Thirteen density functional approximations were tested: BLYP,^{61,62} PBE,⁶³ M06-L,⁶⁴ TPSS,⁶⁵ B3LYP,^{66,67} PBE0,⁶⁸ M06,⁵⁰ M06-2X,⁵⁰ TPSSH,⁶⁹ CAM-B3LYP,⁷⁰ ωB97X-D3,^{44,71} ωB97X-V^{72–74} and B2PLYP.⁷⁵ For functionals that do not account for dispersion, empirical D3^{44,45} correction was added. In hybrid functionals, the chain-of-spheres approximation (COSX)^{76,77} was used for the calculation of Hartree–Fock exchange with increased grid resolution (“GridX9”). For B2PLYP we also used the resolution of identity for the MP2 integrals with auxiliary basis def2-TZVPP/C.⁷⁸

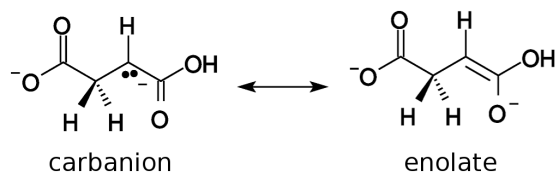
Zero-point vibrational energies (ZPVE) and thermal contributions to free energies were calculated using the quasi rigid-rotor-harmonic-oscillator approximation⁷⁹ with the semiempirical tight-binding method GFN2-

xTB.⁸⁰ Hessian calculations with DFT are unfeasible for the large cluster, so we had to use a semiempirical method. To overcome the problem of having geometries with nonzero gradients at GFN2-xTB level, we employed the single-point Hessian formalism of Spicher and Grimme⁸¹ that uses a biased potential to avoid imaginary frequencies. For transition states however, regular Hessian calculations had to be used because the biased potential was unable to restrain the broken bond geometry. These calculations were performed with xtb program version 6.3.3.⁸²

Results & Discussion

Strong correlation effects

The multiconfigurational character of **5** and **7** were evaluated here with CASSCF calculations. **5** is typically described as a carbanion intermediate,^{1,35,83} but it can also adopt an enolate form (Scheme 4), suggesting the existence of competing valence structures that could require multiconfigurational treatments. Whereas **7** is a singlet diradical species with nearly degenerate orbitals.²⁵



Scheme 4: Possible electromers of **5**.

Active spaces employed are shown in Figures S1 and S2. They were composed of flavin π and π^* molecular orbitals (MOs) that forms the N5=C4a double bond, the σ and σ^* orbitals of the transferred hydrogen and the carbanion/radical non-bonding MO that is doubly/singly occupied respectively in **5** and **7**. For **5** we also included the C=O π and π^* MOs that might become a C-O⁻ bond in the enolate electromer. The resulting active space contained 8 electrons in 7 MOs (8e,7o) for **5** and (6e,5o) for **7**.

In **5**, no CASSCF optimized MOs have natural occupation numbers that differ much

from 2.00 or 0.00, indicating negligible multi-configurational character (Fig. S1). Its C=O π MO has occupation 2.00 and resembles the non-bonding orbital of C-O⁻, demonstrating the “carbanion” intermediate is actually an enolate intermediate. Contrastingly, **7** shows two strongly correlated MOs with occupation 1.09 and 0.91 with essentially the same shape and a change in phase (Fig. S2).

The weight of the leading configurations in the natural orbital CASSCF expansion for **7** is 53.9% and 45.2%, suggesting non-negligible multiconfigurational character. These weights are in line with only a pair of strongly-correlated MOs found from analysis of occupation numbers.

To check the multiconfigurational character in transition states and nonequilibrium geometries, we also performed CASSCF calculations in two representative reactions: the hydride transfer reaction **2** \rightarrow **5** and the adduct dissociation **6** \rightarrow **9**. For the hydride transfer reaction, a single configuration with weight 93% in the expansion is present in all geometries along the reaction path, showing it is essentially monoconfigurational. For the hydrogen-atom transfer **6** \rightarrow **9**, the leading configurations have weights varying from 94% to 98% for all geometries, except the transition state geometry that shows two dominant configurations with weights 84% and 15% indicating multiconfigurational character.

Figure 2 shows the energy profiles calculated with single-reference MP2 and multireference NEVPT2 for these two reactions in comparison to DLPNO-CCSD(T), all at CBS limit. MP2 agrees with DLPNO-CCSD(T), within 1.5 kcal.mol⁻¹ in reaction energies, but underestimates the reaction barriers in 15 and 6 kcal.mol⁻¹ for **2** \rightarrow **5** and **6** \rightarrow **9** respectively. This difference is expected, because the lack of triples excitations in MP2 is known to affect its accuracy in calculating reaction barriers.^{84,85}

Because **2** \rightarrow **5** do not show multiconfigurational character, the difference between MP2 and multireference NEVPT2, that reaches at maximum 6 kcal.mol⁻¹, is probably due to uneven dynamical correlation being recovered by these methods (Fig. 2a). Con-

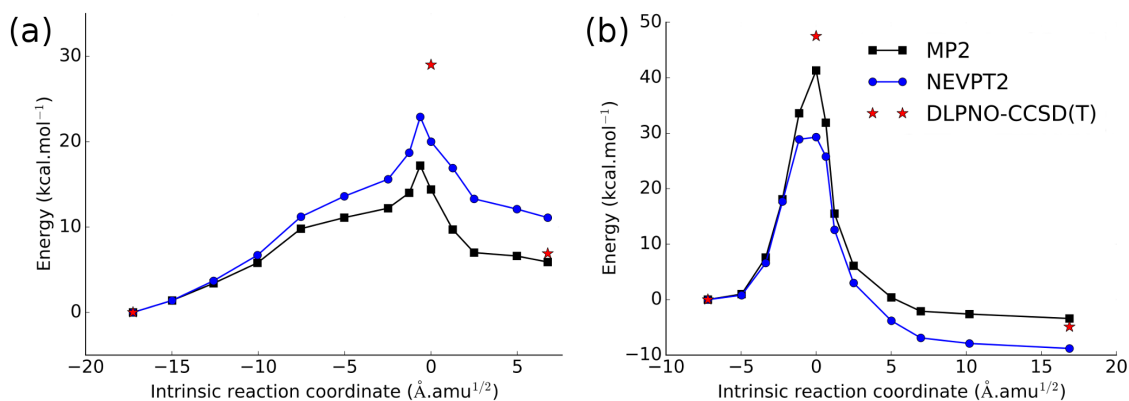


Figure 2: Energies for hydride transfer **2** → **5** (a) and hydrogen-atom transfer **6** → **9** (b). MP2 and DLPNO-CCSD(T) calculations used Hartree-Fock MOs and NEVPT2 used CASSCF(8e,7o) and CASSCF(6e,5o) optimized MOs in (a) and (b) respectively. The energies reported are at the complete basis set limit (CBS) and the reaction path was generated with IRC using PBEh-3c method.

trastingly, for **6** → **9**, in addition to the dynamical correlation error, the transition state is probably overestimated by MP2 and DLPNO-CCSD(T) because they lack static correlation, yielding the larger difference of 12 kcal.mol⁻¹ (Figure. 2b).

Accuracy of the cluster model and DFT methods

Because one electronic configuration dominates the mean-field description for all but radical addition mechanism **1** → **7** → **9** and the transition state of **6** → **9**, we expect single-reference CCSD(T) will successfully recover electron correlation effects. However, even for the small cluster, with 121 atoms, canonical CCSD(T)/CBS calculations cannot be performed, so the approximate DLPNO-CCSD(T) had to be used to generate reference reaction energies and barriers. Recently, we have shown that DLPNO-CCSD(T) provides accurate energies for flavins under different redox states,³⁷ so it is safe to assume it will yield accurate energies for reactions studied here. In fact, the use of DLPNO-CCSD(T) to benchmark other enzyme-catalyzed reactions showed errors of less than 1 kcal.mol⁻¹ compared to canonical CCSD(T).^{86,87}

Despite the accuracy of DLPNO-CCSD(T)

in calculating reaction energies and barriers, the small cluster model might limit mechanistic conclusions because it lacks long-range electrostatic effects and restrains the flexibility of the active site, so we decided to check the effect of truncating the system by comparing the small cluster with the large cluster.

To model the reactions in the large cluster, more approximate calculation methods such as DFT are required. Thus, the reaction energies and barriers for the small cluster were tested here with several density functional approximations in comparison to reference DLPNO-CCSD(T)/CBS (Table 1). The best performance, with mean unsigned (MUE) and maximum absolute errors (MAX) of 1.4 kcal.mol⁻¹ and 3.3 kcal.mol⁻¹, respectively, was obtained with the range-separated hybrid ω B97X-D3/def2-TZVPP. We checked for possible basis set incompleteness error by increasing one cardinal number from def2-TZVPP to def2-QZVPP and the errors remained equal, showing energies are already converged at triple-zeta level. The other range-separated functionals tested here, CAM-B3LYP and ω B97X-V, are only slightly less accurate than ω B97X-D3, indicating the inclusion of balanced Hartree-Fock exchange is crucial. M06-2X, which contains the largest amount of Hartree-Fock exchange (54%) among the hybrid functionals tested here, is the only one with accuracy compa-

rable to the range-separated functionals, suggesting self-interaction error, which is minimized by inclusion of Hartree–Fock exchange, is the main source of errors for predicting accurate reaction energies and barriers.

Table 1: Mean unsigned (MUE) and maximum absolute (MAX) errors calculated with DFT methods for the reaction energies and barriers of the mechanisms studied here in the small cluster, with the exception of the multiconfigurational radical addition mechanism ($N = 20$ relative energy values). All quantities in kcal.mol⁻¹. TZ and QZ stand for def2-TZVPP and def2-QZVPP respectively.

Functional	MUE	MAX
GGA and meta-GGA		
BLYP+D3/TZ	6.8	25.3
PBE+D3/TZ	6.3	22.7
TPSS+D3/TZ	5.7	20.4
M06-L/TZ	5.2	20.6
Hybrid		
B3LYP+D3/TZ	4.2	13.7
PBE0+D3/TZ	3.0	8.5
TPSSH+D3/TZ	4.5	14.9
M06/TZ	2.6	9.9
M06-2X/TZ	2.0	5.4
Range-separated		
CAM-B3LYP+D3/TZ	1.6	3.3
ω B97X-D3/TZ	1.4	3.3
ω B97X-D3/QZ	1.4	3.3
ω B97X-V/TZ	2.0	5.7
Double-hybrid		
B2PLYP+D3/TZ	3.6	12.5

Note: Reference energies were calculated with DLPNO-CCSD(T)/CBS

All GGA and meta-GGA functionals gave poor results with MAX errors higher than 20 kcal.mol⁻¹, the best was M06-L, which is highly parametrized and trained to reproduce reaction energies and barriers.⁶⁴ Because GGA approximation do not include Hartree–Fock exchange, self-interaction errors are particularly critical for them, explaining their low accuracy. For double-hybrid B2PLYP, self-interaction errors usually are not a problem (53% of Hartree–Fock exchange),⁷⁵ yet its accuracy was similar to simple hybrid functionals B3LYP, PBE0

and TPSSH. We checked the HOMO-LUMO gaps calculated with B2PLYP and noted that species with the lowest gaps ended up with the largest errors, indicating the problem with B2PLYP was the perturbative correction it employs, which can be problematic when the energy gaps become small.⁵⁵

B3LYP, which was used in the early study of Lucas and Ramos²⁹ has MUE of 4.2 kcal.mol⁻¹ and MAX error of 13.7 kcal.mol⁻¹ which is similar to the error estimated by them of 5 kcal.mol⁻¹. The biggest problem with their calculations is that their model was too small.

As ω B97X-D3/def2-TZVPP is the most accurate density functional tested here, it was employed to model the reaction in the large cluster. Its maximum absolute error (MAX) of 3.3 kcal.mol⁻¹ is considered the uncertainty associated with our energy calculations. In addition, the calculation of free energies may increase our uncertainty due to the approximate Hessian used to compute ZPVE and thermal contributions. Spicher and Grimme⁸¹ estimated a MUE of 0.6 kcal.mol⁻¹ for this approximation, so in the worst case scenario the error in energy of 3.3 kcal.mol⁻¹ will add to the 0.6 kcal.mol⁻¹, yielding a final uncertainty of approximately 4 kcal.mol⁻¹ for free energies.

Reaction energies and barriers for the large cluster were calculated for all possible reaction steps. By comparing the energies calculated in the small cluster with the large cluster we estimated how the inclusion of a second solvation shell around fumarate can influence catalysis. In average, relative energies varied 11 kcal.mol⁻¹ between the models, with the largest difference being 22 kcal.mol⁻¹. This shows the small model (with 121 atoms!) is unable to reproduce the active site environment and that implicit solvent alone cannot polarize the system appropriately. Thus, the model of Lucas and Ramos²⁹ with only 53 atoms is probably too small to describe the enzymatic reduction of fumarate.

The benchmark study and error assessment reported here is important because it allowed us to set a theoretical framework that minimizes possible sources of errors and legiti-

mate our calculations. More importantly, an uncertainty was estimated, making it possible to compare the different reaction mechanisms with experimental data. The fact that we know the limitations of our model allow us to make fair comparisons without resorting to fortuitous error cancellations.

The mechanism of fumarate reduction

The results reported here were computed at the large cluster. Figure 3 shows the free energies calculated for the reaction mechanisms proposed in Scheme 3, with the exception of the radical addition mechanism.

The mechanism of nucleophilic addition by direct hydride transfer is shown in black and blue in Fig. 3. The difference is that the reaction profile in black does not involve the protonation of fumarate by His504, while the diagram in blue does. In the first case, both the hydride transfer from FADH^- to fumarate ($1 \rightarrow 4$) and the proton transfer from Arg402 to the carbanion intermediate ($4 \rightarrow 9$) have transition states with 23 kcal.mol $^{-1}$ of energy. Because **1** is the rate-determining intermediate (TDI⁸⁸), the effective activation energy of the reaction is 23 kcal.mol $^{-1}$, with both transition states $\text{TS}_{1,4}$ and $\text{TS}_{4,9}$ contributing to reaction kinetics. This contrasts the previous work of Lucas and Ramos²⁹ that indicated hydride transfer as the only rate-limiting step.

Despite its high pK_a in water, Arg402 is easily deprotonated in the enzyme, with a barrier of only 5 kcal.mol $^{-1}$ for the reaction $4 \rightarrow 9$, suggesting the carbanion intermediate is a stronger base than Arg402.

For the second possibility of nucleophilic addition by direct hydride transfer, the initial step is protonation of fumarate by His504 ($1 \rightarrow 2$) with barrier of 15 kcal.mol $^{-1}$. The hydride transfer from FADH^- ($2 \rightarrow 5$) and the proton transfer from Arg402 ($5 \rightarrow 8$) have barriers of 13 kcal.mol $^{-1}$ and 14 kcal.mol $^{-1}$, respectively. The proton transfer to His504 restoring its initial protonation state ($8 \rightarrow 9$) has a barrier of only 2 kcal.mol $^{-1}$. Despite, the low barriers of the individual

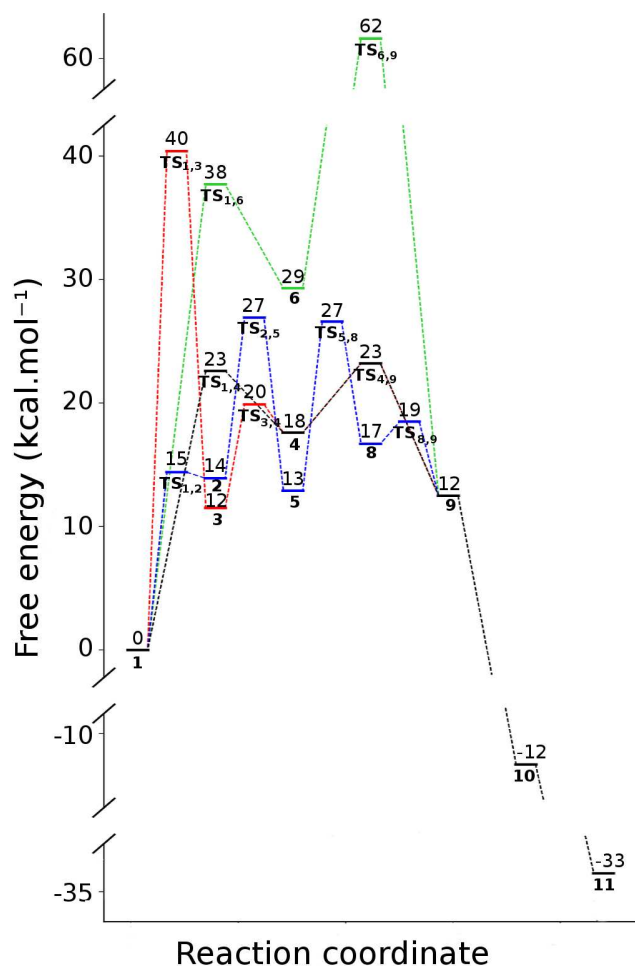


Figure 3: Diagrams of free energies for possible reaction mechanisms of fumarate reduction. The black diagram represents the nucleophilic addition mechanism by direct hydride transfer. Blue is also the nucleophilic addition by direct hydride transfer but with an initial protonation of fumarate (general acid catalysis). In red and green are the nucleophilic addition mechanisms with adduct formation. Free energy values are shown over the bars and the indices below are the ones used in Scheme 3 to identify the molecule.

reaction steps, the overall reaction rate will have an effective activation energy of 27 kcal.mol $^{-1}$, that is the difference between the rate-determining transition states (TDTS⁸⁸) $\text{TS}_{2,5}$ and $\text{TS}_{5,8}$ in relation to the TDI **1**. Therefore, the mechanism of direct hydride transfer without protonation of fumarate by His504 (black diagram in Fig. 3) with an activation energy of 23 kcal.mol $^{-1}$ is favored over the mechanism with protonation by His504 (blue diagram in Fig. 3).

The nucleophilic addition mechanisms with adduct formation³¹ (red and green diagrams in Fig. 3) have **TS**_{1,3} and **TS**_{6,9} determining the effective activation energies of 40 kcal.mol⁻¹ and 62 kcal.mol⁻¹ respectively. Both are much higher than the activation energies calculated for the direct hydride transfer mechanisms, indicating they will not be employed by the enzyme. Even if we consider, that our neglecting of static correlation is overestimating **TS**_{6,9}, the mechanism would still be too energetic since intermediate **6**, which is monoconfigurational has 29 kcal.mol⁻¹ of energy, which is higher than the activation energies calculated for nucleophilic addition mechanisms of 23 kcal.mol⁻¹ and 27 kcal.mol⁻¹. Similar arguments can be used to discard the radical addition mechanism **1** → **7** → **9** (not shown in Fig. 3), whose diradical intermediate **7** has an energy of 71 kcal.mol⁻¹.

Finally, a comment must be made about the energies calculated for the reprotonation of Arg402: steps **9** → **10** and **10** → **11**. Because Arg402 is believed to be reprotonated by an excess proton from the solvent via a Grotthuss mechanism,^{21,89} we changed the system composition in **9** by making one of the structural waters a hydronium and then calculated the energies of **10** and **11** relative to it. For these steps, we abstained from calculating the transition states, as Grotthuss mechanism have naturally low barriers. In fact the overall reaction was very exergonic with $\Delta G = -33$ kcal.mol⁻¹. The reverse reaction of succinate oxidation, proceeding via the black diagram of Fig. 3 has an activation energy of 56 kcal.mol⁻¹, which is much higher than fumarate reduction, corroborating the experimental observations that Fcc₃ is unable to oxidize succinate.^{1,12}

To complete the catalytic cycle, the binding and unbinding free energies of fumarate and succinate to the enzyme must be accounted. It has been shown experimentally that the Michaelis constant K_M is similar to the dissociation constant K_D ^{17,90,91} for Fcc₃, thus the binding and unbinding free energies were calculated using the experimental values¹² of $K_M \approx K_D = -RT \ln(\Delta G)$. For fumarate

complexation forming the enzyme-substrate complex, a free energy of -6.2 kcal.mol⁻¹ is released, whereas for the dissociation of the enzyme-product complex 3.6 kcal.mol⁻¹ is required. Therefore, the chemical reaction and not the binding/unbinding process is responsible for the turnover rate observed in fumarate reduction by Fcc₃.

Considering the uncertainty of 4 kcal.mol⁻¹ of our calculations, the activation energies of 23 kcal.mol⁻¹ and 27 kcal.mol⁻¹ are indistinguishable and we cannot discard the participation of His504 as a general-acid in the catalysis. However, we can safely assume the reaction proceeds via nucleophilic addition by direct hydride transfer, as the other mechanisms involving adduct or radical formation have significant higher activation energies.

Note that the enzyme is not fully regenerated here because we started with FADH⁻ in **1** and end up with FAD in **11**. Because the regeneration of FADH⁻ involves electron transfers from heme groups that are not part of the clusters models (Fig. 1), these steps could not be modeled. Nevertheless, experimental data from electrochemical measurements⁹² have shown that enzyme turnover is not determined by FADH⁻ regeneration, hence the conclusions presented here remain valid.

The experimental k_{cat} for this reaction corresponds to an effective barrier of 14 kcal.mol⁻¹.¹⁷ Considering the activation energy of 23 ± 4 kcal.mol⁻¹ calculated here, the error of our model is 6 kcal.mol⁻¹. The lack of dynamical effects, long-range electrostatic interaction beyond the cluster model and possible hydride tunneling might be contributing to the difference between experiment and the calculations reported here. In the next sections we investigate some of these effects.

How is the reaction catalyzed?

Several strategies can be employed by enzymes to catalyze reactions. We have shown in the previous section that one strategy that might be employed in fumarate reduction is the protonation of fumarate by His504 in a

general acid catalysis. We also discarded covalent catalysis, because the reactions that form adduct intermediates with FAD are too energetic.

A third common strategy employed by enzymes is electrostatic catalysis. We checked the extent to which the active site environment modulates the reaction by mapping the electrostatic potential of the large cluster into the electron density obtained with the ω B97X-D3 energy calculation (Fig. 4). Regions colored in blue represent negative charged regions (δ^-), while red regions represent positively charged (δ^+) regions. The hydride (H^-) is transferred in an axis perpendicular to the plane of Figure 4 towards the positively charged region (Region colored in red in Fig. 4), i.e the active site has its charges naturally distributed in a way that facilitates the hydride transfer.

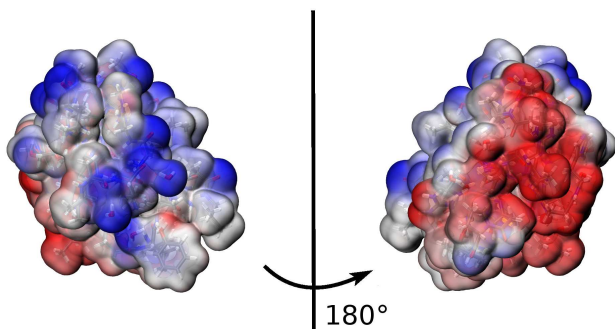


Figure 4: Electrostatic mapping of the large cluster model bound to fumarate **1** into the electron density calculated with ω B97X-D3/def2-TZVPP. The color scale goes from blue to red representing negative charged regions (δ^-) to positively charged regions (δ^+). Hydride transfer takes place in an axis perpendicular to plane of the representation towards the positively charged regions in the right figure.

Dipoles moments were also calculated for the large and small cluster complexed to fumarate (**1** in Scheme 3). In the small cluster dipole was 50.4 D and in the large cluster it was 77.8 D, demonstrating long-range electrostatic interactions increase the dipole moment in the large cluster to facilitate the hydride transfer from $FADH^-$ to dianionic fumarate. Therefore, electrostatic catalysis is

employed by Fcc₃ to facilitate oxidoreduction of fumarate and succinate.

As we used an implicit solvent beyond the cluster region to polarize our model, we decided to check how much would the energetics of the reaction change using implicit solvents with different dielectric constants ϵ . With $\epsilon = 20$, all energies increased compared to default $\epsilon = 4$, with an average deviation of 1 kcal.mol⁻¹. Without implicit solvent, the average difference is -3 kcal.mol⁻¹. Therefore, even implicit solvents far from the reaction center, can still affect the reaction energies in few kcal.mol⁻¹, indicating electrostatic catalysis is a major effect in Fcc₃. Note that the mechanistic conclusions of the previous section remain valid with different dielectrics, because the energy differences due to ϵ are smaller than the energy differences between mechanisms.

Other hypothesis usually discussed for enzymatic reduction of fumarate is that there is a geometric constraint applied by the enzyme to fumarate that twists its carboxylate 1 to a less optimal geometry in order to facilitate hydride transfer.^{1,2,89,93} In fact, fumarate was found twisted in the optimized geometry calculated here agreeing with the conformation observed in crystal structures complexed to fumarate.^{2,10,94} To check the stability of planar fumarate, we performed a constrained geometry optimization to keep it in a planar conformation, which yield an energy difference of only 6 kcal.mol⁻¹ higher than the twisted conformation. Therefore both conformations could be visited in a thermalized enzyme-substrate complex, suggesting there is not actually a geometric constraint to the substrate and the contribution of this feature to catalysis must be small.

The role of Thr377 in coordinating the twist of fumarate was also tested by replacement of this threonine to an alanine in our calculations. In the mutant T377A, the fumarate remained twisted, indicating Thr377 hydrogen-bond with carboxylate is not critical for this conformation. The other hydrogen-bonds with His365 and the backbone of Glu378 are sufficient to stabilize fumarate in the twisted conformation. The loss of activity observed

experimentally upon T377A mutation might be related to larger changes in the enzyme environment that could increase for instance active site hydration.

Characterization of the hydride transfer step

The hydride transfer from FADH^- to fumarate (**1** \rightarrow **4** in Fig. 3) is one of the rate-limiting steps of the reaction. We have shown recently that this step is not a prototypical hydride transfer, with electrons and protons being transferred asynchronously to different parts of the molecule.²⁵ Because proton-coupled electron transfer reactions might involve excited states and nonadiabatic regimes,^{95–99} we decided to further investigate the role of nonadiabaticity and nuclear quantum effects in this reaction step. Yet, our aim here is to provide a qualitative or semiquantitative picture of these effects, therefore the small cluster model was employed in all calculations reported in this section.

Electronic nonadiabatic effects due to state crossings with excited states were calculated for first-excited state singlets (S_1) with time-dependent DFT (TD-DFT) at $\omega\text{B97X-D3/def2-TZVPP}$ level for the species **1**, **4** and **TS_{1,4}** along the reaction path. For fumarate **1**, S_1 is 70 kcal.mol⁻¹ more energetic than ground state singlet (S_0), for the carbanion intermediate **4**, this energy difference is 64 kcal.mol⁻¹ and for the transition state **TS_{1,4}**, 72 kcal.mol⁻¹. Thus, electronic excited states are significantly higher than ground state along the hydride transfer pathway and do not approach or cross with it, indicating the reaction is electronically adiabatic when thermally activated.

Consequently, the nuclei move in the electronic ground state potential energy surface (Born–Oppenheimer approximation). Figure 5a shows the minimum energy path connecting **1** to **4** obtained with an intrinsic reaction coordinate (IRC) calculation. In the first third of the reaction, until IRC of approximately -3.0, the distances between the transferred hydrogen and the donor or ac-

ceptor atoms remain equal, while the other atoms not involved directly in the reaction, i.e the environment, are being adjusted to facilitate de hydride transfer. From IRC = -3.0 to IRC = -0.2, the hydride donor(nitrogen N5 of FADH^-) and hydride acceptor(carbon C2 of fumarate) approach each other, with the proton transfer taking place in the narrow interval between IRC = -0.7 to 0.6. After the hydride transfer, the donor-acceptor distance increases and the environment adjusts to accommodate the product **4**. At the transition state, i.e IRC = 0.0 (Fig. 5b), the environment is optimized to facilitate the hydride transfer, therefore we employed the **TS_{1,4}** geometry to check for nuclear quantum effects of the transferring proton.

Figure 5c shows the two-dimensional proton potential at the transition state. The transferring proton was placed in different positions in the plane formed by atoms N5 of FADH^- , C2 of fumarate and the proton at the **TS_{1,4}** geometry (see Fig. 5b) to calculate the proton potential. The points indicated by stars are stationary points at the proton potential: one is the optimized transition state, and the others are degenerate minima at the reactant and product wells. The energy difference ΔV_p^\ddagger between **TS_{1,4}** and the minima is only 2 kcal.mol⁻¹, indicating an almost barrierless proton transfer potential, usually found in vibronically adiabatic PCET.⁹⁸

The transferring proton vibrational wave function amplitudes calculated at the two-dimensional proton potential are shown in Fig. 5d with their respective energy splitting Δ_{vib} . In the ground vibrational state, the proton is delocalized between the reactant and product wells, with the largest amplitudes found around the transition state geometry, whereas for the first- and second-excited vibrational states, the proton vibrational wave function have larger amplitudes around the reactant and product wells respectively. The energy splitting between the ground and first-excited vibrational states $\Delta_{\text{vib}} = 3$ kcal.mol⁻¹ is higher than ΔV_p^\ddagger , suggesting the reaction is not only electronically adiabatic but also vibrationally adiabatic. That means electrons will respond instantaneously to the motion of

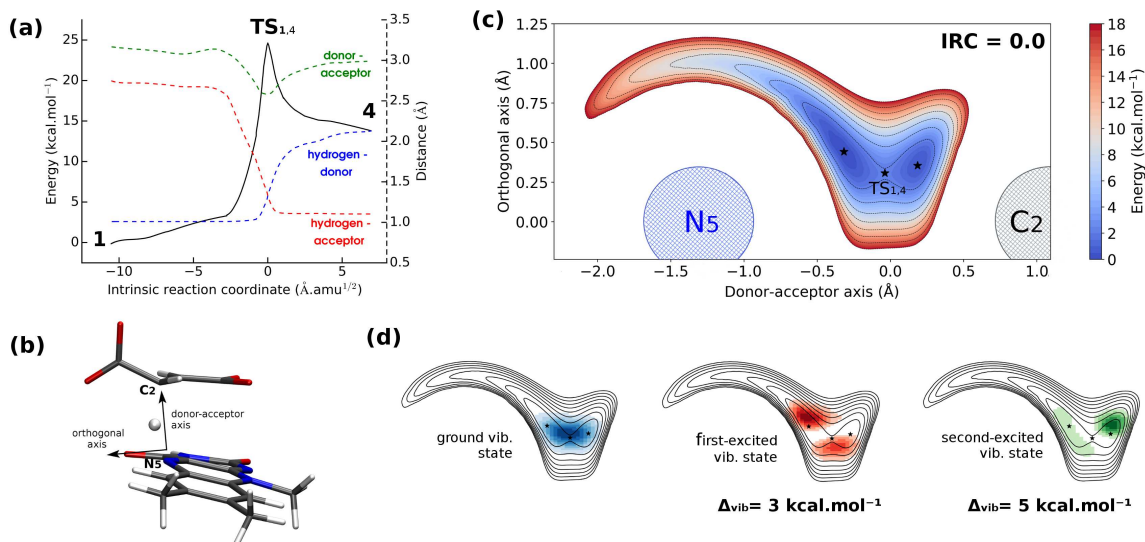


Figure 5: Proton vibrational analysis of the hydride transfer step. Minimum energy path for the reaction $1 \rightarrow 4$ calculated with IRC at PBEh-3c level (a), where dashed lines correspond to distances (right axis) and the continuous line correspond to energy (left axis). Representation of the fumarate and flavin geometries at the transition state, showing the donor-acceptor axis and the orthogonal axis (b). Two-dimensional proton potential calculated at the $TS_{1,4}$ geometry (IRC = 0.0) with $\omega B97X-D3/def2-TZVPP$ (c). Proton vibrational wave function amplitudes calculated with Fourier grid Hamiltonian method (d). All data were obtained from the small cluster model.

the transferring proton, and the proton itself will respond instantaneously to the motion of the other nuclei.^{96,98}

In this regime, excited states will not contribute to the reaction rate, therefore the classical “over-the-barrier” view for the reaction mechanism will be a good approximation. Nevertheless, as shown in Fig. 5d, the proton is delocalized over the reactant and product wells in the ground state, challenging the classical view of an adiabatic transfer. In fact, due to the large charge reorganization upon hydride transfer, the environment adjustment prior to the proton transfer seems to be the biggest driving force for the reaction, in a way that when the environment and the donor-acceptor distance reaches that of the transition state geometry, the proton transfer will be essentially barrierless. This view is analogous to Marcus theory,^{100,101} where electrons are transferred upon appropriate solvent reorganization.

Even though, the hydride transfer is vibronically adiabatic, deep tunneling of the proton directly from the reactant to the proton wells is still possible, as was already observed

in other flavoproteins.^{33,34} The simplest diagnostic for this phenomenon is Wigner’s first order correction to ground state tunneling^{102–104} that calculates the transmission coefficient κ as a function of the modulus of the transition state imaginary frequency ν^\ddagger and the temperature T :

$$\kappa \sim 1 + \frac{1}{24} \left(\frac{h c \nu^\ddagger}{k_B T} \right)^2 \quad (1)$$

where h is Planck’s constant, c is the speed of light and k_B is the Boltzmann constant.

Proton quantum parameters discussed here are shown in Table 2. The calculated κ of 2.6 is consistent with small to moderate tunneling of the hydride,¹⁰² which is in accordance with the low deuterium kinetic isotope effects (KIE) observed experimentally.^{26–28,92} Therefore, we believe tunneling is present but not critical for enzyme kinetics.

Table 2: Nuclear quantum effects parameters of the transferred proton

Proton vibration adiabaticity ^a	
ΔV_p^\ddagger (kcal.mol ⁻¹)	2
Δ_{vib} (kcal.mol ⁻¹)	3
Hydride tunneling ^b	
ν^\ddagger (cm ⁻¹)	1268.5
κ	2.6

^a Proton potential and vibrational wave function calculated in a two-dimensional grid using Fourier grid Hamiltonian method.^{105–107}

^b κ was calculated with Wigner’s first order correction for ground state tunneling.^{102–104}

Conclusions

Accurate predictions and new explanations of how the Fcc₃ enzyme works are presented here. Mechanisms involving adducts or radical intermediates were unequivocally discarded, confirming the current experimental view that fumarate reduction occurs in two steps via a carbanion intermediate. Fumarate protonation prior to hydride transfer appears to be unnecessary.

The electrostatic potential and dipole moment of the active site is responsible to facilitate hydride transfer towards fumarate, giving directionality to the reaction. The twisted conformation of fumarate, considered to be determining for catalysis,^{2,89,93} was shown to be less relevant.

Both hydride transfer from FADH⁻ and proton transfer from Arg402 have equal energy barriers and contribute to the effective activation energy, contrasting the previous work of Ramos and Lucas²⁹ that suggested the hydride transfer as rate-limiting.

Hydride transfer was shown to be both electronically and vibrationally adiabatic with low tunneling contribution. The DFT benchmark and static correlation analysis performed will help to build models for future hybrid QM/MM calculations. Usage of the ω B97X-D3 functional with at least a triple zeta basis set and a large quantum region is recommended.

Our findings can be extrapolated to other fumarate reductases and potentially to succi-

nate dehydrogenases. The effect of covalently bound FAD in membrane-associated enzymes and the possibility of a concerted mechanism were not addressed here and will be explored in a future study.

Acknowledgement Funding from FAPESP (scholarships 2016/23525-0 and 2017/26109-0 to F.C., and grants 2016/24096-5 and 2019/21856-7 to G.M.A.) are gratefully acknowledged.

References

- (1) Iverson, T. M. Catalytic mechanisms of complex II enzymes: A structural perspective. *Biochim. Biophys. Acta Bioenerg.* **2013**, 1827, 648–657.
- (2) Taylor, P.; Pealing, S. L.; Reid, G. A.; Chapman, S. K.; Walkinshaw, M. D. Structural and mechanistic mapping of a unique fumarate reductase. *Nat. Struct. Biol.* **1999**, 6, 1108–1112.
- (3) Besteiro, S.; Biran, M.; Biteau, N.; Couston, V.; Baltz, T.; Canioni, P.; Bringaud, F. Succinate Secreted by *Trypanosoma brucei* Is Produced by a Novel and Unique Glycosomal Enzyme, NADH-dependent Fumarate Reductase. *J. Biol. Chem.* **2002**, 277, 38001–38012.
- (4) Kim, S.; Kim, C. M.; Son, Y.-J.; Choi, J. Y.; Siegenthaler, R. K.; Lee, Y.; Jang, T.-H.; Song, J.; Kang, H.; Kaiser, C. A.; Park, H. H. Molecular basis of maintaining an oxidizing environment under anaerobiosis by soluble fumarate reductase. *Nat. Commun.* **2018**, 9, 4867.
- (5) Yankovskaya, V.; Horsefield, R.; Törnroth, S.; Luna-Chavez, C.; Miyoshi, H.; Léger, C.; Byrne, B.; Cecchini, G.; Iwata, S. Architecture of Succinate Dehydrogenase and Reactive Oxygen Species Generation. *Science* **2003**, 299, 700–704.
- (6) Sun, F.; Huo, X.; Zhai, Y.; Wang, A.; Xu, J.; Su, D.; Bartlam, M.; Rao, Z. Crystal Structure of Mitochondrial Respiratory Membrane Protein Complex II. *Cell* **2005**, 121, 1043–1057.
- (7) Huang, L.-S.; Shen, J. T.; Wang, A. C.; Berry, E. A. 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.* **2006**, 1757, 1073–1083.

- (8) Gong, H. et al. Cryo-EM structure of trimeric *Mycobacterium smegmatis* succinate dehydrogenase with a membrane-anchor SdhF. *Nat. Commun.* **2020**, *11*, 4245.
- (9) Iverson, T. M.; Luna-Chavez, C.; Cecchini, G.; Rees, D. C. Structure of the *Escherichia coli* Fumarate Reductase Respiratory Complex. *Science* **1999**, *284*, 1961–1966.
- (10) Lancaster, C. R. D.; Kröger, A.; Auer, M.; Michel, H. Structure of fumarate reductase from *Wolinella succinogenes* at 2.2 Å resolution. *Nature* **1999**, *402*, 377–385.
- (11) Shimizu, H.; Osanai, A.; Sakamoto, K.; Inaoka, D. K.; Shiba, T.; Harada, S.; Kita, K. Crystal structure of mitochondrial quinol-fumarate reductase from the parasitic nematode *Ascaris suum*. *J. Biochem.* **2012**, *151*, 589–592.
- (12) Turner, K. L.; Doherty, M. K.; Heering, H. A.; Armstrong, F. A.; Reid, G. A.; Chapman, S. K. Redox Properties of Flavocytochrome c_3 from *Shewanella frigidimarina* NCIMB400. *Biochemistry* **1999**, *38*, 3302–3309.
- (13) Blaut, M.; Whittaker, K.; Valdovinos, A.; Ackrell, B. A. C.; Gunsalus, R. P.; Cecchini, G. Fumarate Reductase Mutants of *Escherichia coli* That Lack Covalently Bound Flavin. *J. Biol. Chem.* **1989**, *264*, 13599–13604.
- (14) Guest, J. R. Partial Replacement of Succinate Dehydrogenase Function by Phage- and Plasmid-specified Fumarate Reductase in *Escherichia coli*. *Microbiology* **1981**, *122*, 171–179.
- (15) Maklashina, E.; Berthold, D. A.; Cecchini, G. Anaerobic Expression of *Escherichia coli* Succinate Dehydrogenase: Functional Replacement of Fumarate Reductase in the Respiratory Chain during Anaerobic Growth. *J. Bacteriol.* **1998**, *180*, 5989–5996.
- (16) Bamford, V.; Dobbin, P. S.; Richardson, D. J.; Hemmings, A. M. Open conformation of a flavocytochrome c_3 fumarate reductase. *Nat. Struct. Biol.* **1999**, *6*, 1104–1107.
- (17) Doherty, M. K.; Pealing, S. L.; Miles, C. S.; Moysey, R.; Taylor, P.; Walkinshaw, M. D.; Reid, G. A.; Chapman, S. K. Identification of the Active Site Acid/Base Catalyst in a Bacterial Fumarate Reductase: A Kinetic and Crystallographic Study. *Biochemistry* **2000**, *39*, 10695–10701.
- (18) Mowat, C. G.; Moysey, R.; Miles, C. S.; Leys, D.; Doherty, M. K.; Taylor, P.; Walkinshaw, M. D.; Reid, G. A.; Chapman, S. K. Kinetic and Crystallographic Analysis of the Key Active Site Acid/Base Arginine in a Soluble Fumarate Reductase. *Biochemistry* **2001**, *40*, 12292–12298.
- (19) Pankhurst, K. L.; Mowat, C. G.; Miles, C. S.; Leys, D.; Walkinshaw, M. D.; Reid, G. A.; Chapman, S. K. Role of His505 in the Soluble Fumarate Reductase from *Shewanella frigidimarina*. *Biochemistry* **2002**, *41*, 8551–8556.
- (20) Rothery, E. L.; Mowat, C. G.; Miles, C. S.; Mott, S.; Walkinshaw, M. D.; Reid, G. A.; Chapman, S. K. Probing Domain Mobility in a Flavocytochrome. *Biochemistry* **2004**, *43*, 4983–4989.
- (21) Pankhurst, K. L.; Mowat, C. G.; Rothery, E. L.; Hudson, J. M.; Jones, A. K.; Miles, C. S.; Walkinshaw, M. D.; Armstrong, F. A.; Reid, G. A.; Chapman, S. K. A Proton Delivery Pathway in the Soluble Fumarate Reductase from *Shewanella frigidimarina*. *J. Biol. Chem.* **2006**, *281*, 20589–20597.
- (22) Wardrope, C.; Mowat, C. G.; Walkinshaw, M. D.; Reid, G. A.; Chapman, S. K. Fumarate reductase: Structural and mechanistic insights from the catalytic reduction of 2-methylfumarate. *FEBS Lett.* **2006**, *580*, 1677–1680.
- (23) Tedeschi, G.; Ronchi, S.; Simonic, T.; Treu, C.; Mattevi, A.; Negri, A. Probing the Active Site of L-Aspartate Oxidase by Site-Directed Mutagenesis: Role of Basic Residues in Fumarate Reduction. *Biochemistry* **2001**, *40*, 4738–4744.
- (24) Schlippe, Y. V. G.; Hedstrom, L. A twisted base? The role of arginine in enzyme-catalyzed proton abstractions. *Arch. Biochem. Biophys.* **2005**, *433*, 266–278.
- (25) Curtolo, F.; Arantes, G. M. Mechanisms for Flavin-Mediated Oxidation: Hydride or Hydrogen-Atom Transfer? *J. Chem. Inf. Model.* **2020**, *60*, 6282–6287.
- (26) Vitale, L.; Rittenberg, D. The Rates of Oxidation of Some Deuterio Isomers of Succinate by Succinic Dehydrogenase. *Biochemistry* **1967**, *6*, 690–699.
- (27) Hollocher, T. C.; You, K.-s.; Conjalka, M. Rate-Determining Steps in the Oxidation of Succinate Catalyzed by Succinic Dehydrogenase. *J. Am. Chem. Soc.* **1970**, *92*, 1032–1035.
- (28) Kaczorowski, G. J.; Cheung, Y.-F.; Walsh, C. Substrate Kinetic Isotope Effects in Dehydrogenase Coupled Active Transport in Membrane Vesicles of *Escherichia coli*. *Biochemistry* **1977**, *16*, 2619–2628.

- (29) Lucas, M. F.; Ramos, M. J. Mechanism of a Soluble Fumarate Reductase from *Shewanella frigidimarina*: A Theoretical Study. *J. Phys. Chem. B* **2006**, *110*, 10550–10556.
- (30) Vik, S. B.; Hatefi, Y. Possible occurrence and role of an essential histidyl residue in succinate dehydrogenase. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 6749–6753.
- (31) Venkataram, U. V.; Bruice, T. C. On the Mechanism of Flavin-Catalyzed Dehydrogenation α,β to an Acyl Function. The Mechanism of 1,5-Dihydroflavin Reduction of Maleimides. *J. Am. Chem. Soc.* **1984**, *106*, 5703–5709.
- (32) Dai, Y.; Kizjakina, K.; Campbell, A. C.; Korasick, D. A.; Tanner, J. J.; Sobrado, P. Flavin-N5 Covalent Intermediate in a Nonredox Dehalogenation Reaction Catalyzed by an Atypical Flavoenzyme. *ChemBioChem* **2018**, *19*, 53–57.
- (33) Pang, J.; Hay, S.; Scrutton, N. S.; Sutcliffe, M. J. Deep Tunneling Dominates the Biologically Important Hydride Transfer Reaction from NADH to FMN in Morphinone Reductase. *J. Am. Chem. Soc.* **2008**, *130*, 7092–7097.
- (34) Delgado, M.; Görlich, S.; Longbotham, J. E.; Scrutton, N. S.; Hay, S.; Moliner, V.; Tuñón, I. Convergence of Theory and Experiment on the Role of Preorganization, Quantum Tunneling, and Enzyme Motions into Flavoenzyme-Catalyzed Hydride Transfer. *ACS Catal.* **2017**, *7*, 3190–3198.
- (35) Schröder, I.; Gunsalus, R. P.; Ackrell, B. A. C.; Cochran, B.; Cecchini, G. Identification of Active Site Residues of *Escherichia coli* Fumarate Reductase by Site-directed Mutagenesis. *J. Biol. Chem.* **1991**, *266*, 13572–13579.
- (36) Léger, C.; Heffron, K.; Pershad, H. R.; Maklashina, E.; Luna-Chavez, C.; Cecchini, G.; Ackrell, B. A. C.; Armstrong, F. A. Enzyme Electrokinetics: Energetics of Succinate Oxidation by Fumarate Reductase and Succinate Dehydrogenase. *Biochemistry* **2001**, *40*, 11234–11245.
- (37) Curtolo, F.; Arantes, G. M. Molecular properties and tautomeric equilibria of isolated flavins. *J. Comput. Chem.* **2022**, *43*, 1561–1572.
- (38) Harms, M. J.; Schlessman, J. L.; Sue, G. R.; García-Moreno E., B. Arginine residues at internal positions in a protein are always charged. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 18954–18959.
- (39) Neese, F. The ORCA program system. *WIREs Comput. Mol. Sci.* **2012**, *2*, 73–78.
- (40) Neese, F. Software update: the ORCA program system, version 4.0. *WIREs Comput. Mol. Sci.* **2018**, *8*, e1327.
- (41) Barone, V.; Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A* **1998**, *102*, 1995–2001.
- (42) Siegbahn, P. E.; Himo, F. The quantum chemical cluster approach for modeling enzyme reactions. *WIREs Comput. Mol. Sci.* **2011**, *1*, 323–336.
- (43) Grimme, S.; Brandenburg, J. G.; Banwarth, C.; Hansen, A. Consistent structures and interactions by density functional theory with small atomic orbital basis sets. *J. Chem. Phys.* **2015**, *143*, 054107.
- (44) 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.
- (45) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- (46) Kruse, H.; Grimme, S. A geometrical correction for the inter- and intra-molecular basis set superposition error in Hartree-Fock and density functional theory calculations for large systems. *J. Chem. Phys.* **2012**, *136*, 154101.
- (47) Ishida, K.; Morokuma, K.; Komornicki, A. The intrinsic reaction coordinate. An *ab initio* calculation for $\text{HNC} \rightarrow \text{HCN}$ and $\text{H}^- + \text{CH}_4 \rightarrow \text{CH}_4 + \text{H}^-$. *J. Chem. Phys.* **1977**, *66*, 2153–2156.
- (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*, 024109.
- (49) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-set convergence of correlated calculations on water. *J. Chem. Phys.* **1997**, *106*, 9639–9646.
- (50) 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

- p>testing of four M06-class functionals and 12 other functionals.
- Theor. Chem. Acc.*
- 2008**
- ,
- 120*
- , 215–241.
- (51) Karton, A.; Martin, J. M. L. Comment on: “Estimating the Hartree–Fock limit from finite basis set calculations” [Jensen F (2005) *Theor Chem Acc* 113:267]. *Theor. Chem. Acc.* **2006**, *115*, 330–333.
 - (52) Dunning, T. H., Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
 - (53) Weigend, F.; Köhn, A.; Hättig, C. Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. *J. Chem. Phys.* **2002**, *116*, 3175–3183.
 - (54) Weigend, F. A fully direct RI-HF algorithm: Implementation, optimised auxiliary basis sets, demonstration of accuracy and efficiency. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285–4291.
 - (55) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*, 1st ed.; John Wiley & Sons, Ltd: Chichester, 2000.
 - (56) Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. Introduction of n -electron valence states for multireference perturbation theory. *J. Chem. Phys.* **2001**, *114*, 10252–10264.
 - (57) Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. n -electron valence state perturbation theory: A spinless formulation and an efficient implementation of the strongly contracted and of the partially contracted variants. *J. Chem. Phys.* **2002**, *117*, 9138–9153.
 - (58) Pansini, F. N. N.; Neto, A. C.; Varandas, A. J. C. Extrapolation of Hartree–Fock and multiconfiguration self-consistent-field energies to the complete basis set limit. *Theor. Chem. Acc.* **2016**, *135*, 261.
 - (59) 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.
 - (60) Weigend, F. Accurate Coulomb-fitting basis sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.
 - (61) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.
 - (62) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789.
 - (63) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
 - (64) 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.
 - (65) Tao, J.; 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.
 - (66) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
 - (67) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *Ab Initio* Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
 - (68) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.
 - (69) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes. *J. Chem. Phys.* **2003**, *119*, 12129–12137.
 - (70) Yanai, T.; Tew, D. P.; Handy, N. C. A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57.
 - (71) 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*, 263–272.
 - (72) Mardirossian, N.; Head-Gordon, M. ω B97X-V: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904–9924.

- (73) Vydrov, O. A.; Voorhis, T. V. Nonlocal van der Waals density functional: The simpler the better. *J. Chem. Phys.* **2010**, *133*, 244103.
- (74) Hujo, W.; Grimme, S. Performance of the van der Waals Density Functional VV10 and (hybrid)GGA Variants for Thermochemistry and Noncovalent Interactions. *J. Chem. Theory Comput.* **2011**, *7*, 3866–3871.
- (75) Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **2006**, *124*, 034108.
- (76) Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. Efficient, approximate and parallel Hartree–Fock and hybrid DFT calculations. A ‘chain-of-spheres’ algorithm for the Hartree–Fock exchange. *Chem. Phys.* **2009**, *356*, 98–109.
- (77) Izsák, R.; Neese, F. An overlap fitted chain of spheres exchange method. *J. Chem. Phys.* **2011**, *135*, 144105.
- (78) Hellweg, A.; Hättig, C.; Höfener, S.; Kloppe, W. Optimized accurate auxiliary basis sets for RI-MP2 and RI-CC2 calculations for the atoms Rb to Rn. *Theor. Chem. Acc.* **2007**, *117*, 587–597.
- (79) Grimme, S. Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory. *Chem. Eur. J.* **2012**, *18*, 9955–9964.
- (80) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.
- (81) Spicher, S.; Grimme, S. Single-Point Hessian Calculations for Improved Vibrational Frequencies and Rigid-Rotor-Harmonic-Oscillator Thermodynamics. *J. Chem. Theory Comput.* **2021**, *17*, 1701–1714.
- (82) Bannwarth, C.; Caldeweyher, E.; Ehlert, S.; Hansen, A.; Pracht, P.; Seibert, J.; Spicher, S.; Grimme, S. Extended tight-binding quantum chemistry methods. *WIREs Comput. Mol. Sci.* **2021**, *11*, e1493.
- (83) Farng, O. L.; Bruice, T. C. Carbon–Carbon Double Bond Formation Accompanying Hydride Transfer from a Carbanion to 5-Carbalumiflavin. *J. Chem. Soc., Chem. Commun.* **1984**, *3*, 185–186.
- (84) Ess, D. H.; Houk, K. N. Activation Energies of Pericyclic Reactions: Performance of DFT, MP2, and CBS-QB3 Methods for the Prediction of Activation Barriers and Reaction Energetics of 1,3-Dipolar Cycloadditions, and Revised Activation Enthalpies for a Standard Set of Hydrocarbon Pericyclic Reactions. *J. Phys. Chem. A* **2005**, *109*, 9542–9553.
- (85) Cremer, D. Møller–Plesset perturbation theory: from small molecule methods to methods for thousands of atoms. *WIREs Comput. Mol. Sci.* **2011**, *1*, 509–530.
- (86) Wappett, D. A.; Goerigk, L. Toward a Quantum-Chemical Benchmark Set for Enzymatically Catalyzed Reactions: Important Steps and Insights. *J. Phys. Chem. A* **2019**, *123*, 7057–7074.
- (87) Paiva, P.; Ramos, M. J.; Fernandes, P. A. Assessing the validity of DLPNO-CCSD(T) in the calculation of activation and reaction energies of ubiquitous enzymatic reactions. *J. Comput. Chem.* **2020**, *41*, 2459–2468.
- (88) Kozuch, S.; Shaik, S. How to Conceptualize Catalytic Cycles? The Energetic Span Model. *Acc. Chem. Res.* **2011**, *44*, 101–110.
- (89) Tomasiak, T. M.; Maklashina, E.; Cecchini, G.; Iverson, T. M. A Threonine on the Active Site Loop Controls Transition State Formation in *Escherichia coli* Respiratory Complex II. *J. Biol. Chem.* **2008**, *283*, 15460–15468.
- (90) Morris, C. J.; Black, A. C.; Pealing, S. L.; Manson, F. D. C.; Chapman, S. K.; Reid, G. A.; Gibson, D. M.; Ward, F. B. Purification and properties of a novel cytochrome: flavocytochrome c from *Shewanella putrefaciens*. *Biochem. J.* **1994**, *302*, 587–593.
- (91) Pealing, S. L.; Cheesman, M. R.; Reid, G. A.; Thomson, A. J.; Ward, F. B.; Chapman, S. K. Spectroscopic and Kinetic Studies of the Tetraheme Flavocytochrome c From *Shewanella putrefaciens* NCIMB400. *Biochemistry* **1995**, *34*, 6153–6158.
- (92) Hirst, J.; Ackrell, B. A. C.; Armstrong, F. A. Global Observation of Hydrogen/Deuterium Isotope Effects on Bidirectional Catalytic Electron Transport in an Enzyme: Direct Measurement by Protein-Film Voltammetry. *J. Am. Chem. Soc.* **1997**, *119*, 7434–7439.
- (93) Tomasiak, T. M.; Archuleta, T. L.; Andréll, J.; Luna-Chávez, C.; Davis, T. A.; Sarwar, M.; Ham, A. J.; McDonald, W. H.; Yankovskaya, V.; Stern, H. A.; Johnston, J. N.; Maklashina, E.; Cecchini, G.; Iverson, T. M.

- Geometric Restraint Drives On- and Off-pathway Catalysis by the *Escherichia coli* Menaquinol:Fumarate Reductase. *J. Biol. Chem.* **2011**, *286*, 3047–3056.
- (94) Tomasiak, T. M.; Archuleta, T. L.; Andr  ll, J.; Luna-Ch  vez, C.; Davis, T. A.; Sarwar, M.; Ham, A. J.; McDonald, W. H.; Yankovskaya, V.; Stern, H. A.; Johnston, J. N.; Maklashina, E.; Cecchini, G.; Iverson, T. M. Geometric Restraint Drives On- and Off-pathway Catalysis by the *Escherichia coli* Menaquinol:Fumarate Reductase. *J. Biol. Chem.* **2011**, *286*, 3047–3056.
- (95) Hammes-Schiffer, S. Comparison of Hydride, Hydrogen Atom, and Proton-Coupled Electron Transfer Reactions. *ChemPhysChem* **2002**, *3*, 33–42.
- (96) Skone, J. H.; Soudackov, A. V.; Hammes-Schiffer, S. Calculation of Vibronic Couplings for Phenoxy/Phenol and Benzyl/Toluene Self-Exchange Reactions: Implications for Proton-Coupled Electron Transfer Mechanisms. *J. Am. Chem. Soc.* **2006**, *128*, 16655–16663.
- (97) Hammes-Schiffer, S.; Stuchebrukhov, A. A. Theory of Coupled Electron and Proton Transfer Reactions. *Chem. Rev.* **2010**, *110*, 6939–6960.
- (98) Hammes-Schiffer, S. Proton-coupled electron transfer: classification scheme and guide to theoretical methods. *Energy Environ. Sci.* **2012**, *5*, 7696–7703.
- (99) Camilo, S. R. G.; Curtolo, F.; Galassi, V. V.; Arantes, G. M. Tunneling and Nonadiabatic Effects on a Proton-Coupled Electron Transfer Model for the Q_o Site in Cytochrome *bc*₁. *J. Chem. Inf. Model.* **2021**, *61*, 1840–1849.
- (100) Marcus, R. A. On the Theory of Oxidation-Reduction Reactions Involving Electron Transfer. I. *J. Chem. Phys.* **1956**, *24*, 966–978.
- (101) Marcus, R. A.; Sutin, N. Electron transfers in chemistry and biology. *Biochim. Biophys. Acta, Rev. Bioenerg.* **1985**, *811*, 265–322.
- (102) Nandi, A.; Molpeceres, G.; Gupta, P. K.; Major, D. T.; K  stner, J.; Martin, J. M.; Kozuch, S. Quantum Tunneling in Computational Catalysis and Kinetics: Is it Really Important? In *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*; Elsevier, 2022.
- (103) Bell, R. P. The application of tunnel corrections in chemical kinetics. In *The Tunnel Effect in Chemistry*; Springer US: Boston, 1980; Chapter 3, pp 51–76.
- (104) Wigner, E. P.   ber das   berschreiten von Potentialschwelen bei chemischen Reaktionen. *Z. Phys. Chem.* **1932**, *19B*, 203–216.
- (105) Marston, C. C.; Balint-Kurti, G. G. The Fourier grid Hamiltonian method for bound state eigenvalues and eigenfunctions. *J. Chem. Phys.* **1989**, *91*, 3571–3576.
- (106) Webb, S. P.; Hammes-Schiffer, S. Fourier grid Hamiltonian multiconfigurational self-consistent-field: A method to calculate multidimensional hydrogen vibrational wavefunctions. *J. Chem. Phys.* **2000**, *113*, 5214–5227.
- (107) Soudackov, A. V.; Hammes-Schiffer, S. *fgh_casci_3d*. <http://webpcet.chem.yale.edu/downloads.html>, (accessed 2022-08-31).

Supporting Information

Mechanism of flavin oxidation catalyzed by fumarate reductases

Felipe Curtolo 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

E-mail: garantes@iq.usp.br

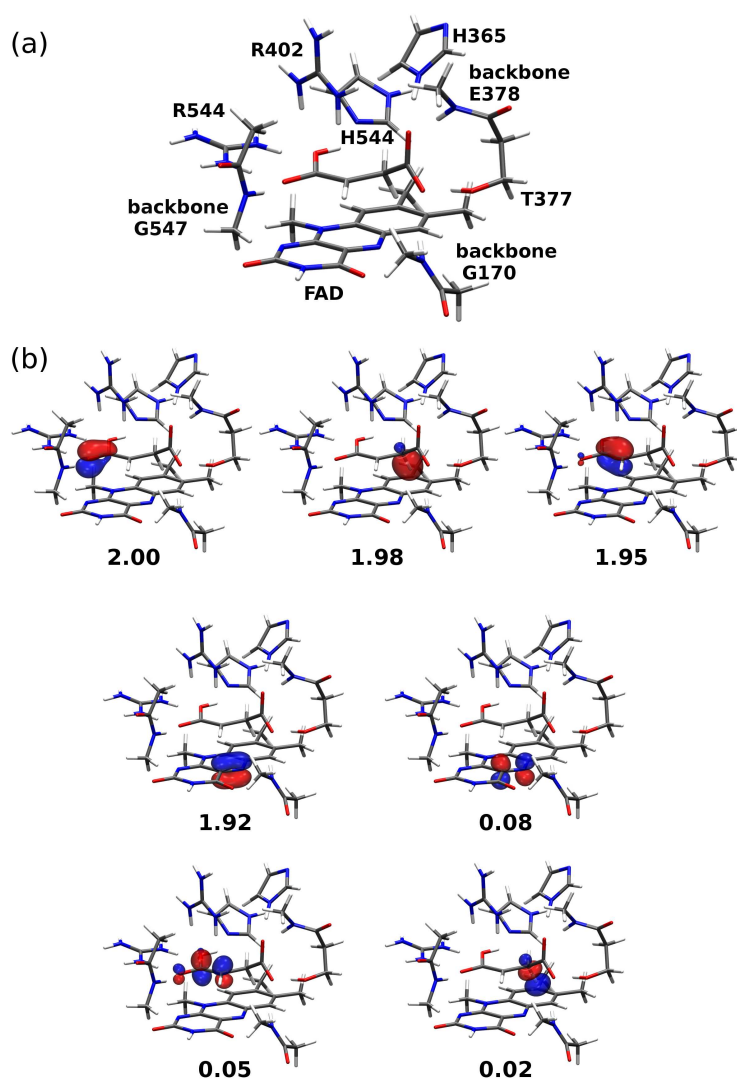


Figure S1: Small cluster model of **5** (a) and CASSCF(8e,7o) final MOs, with their natural orbital occupation numbers (b).

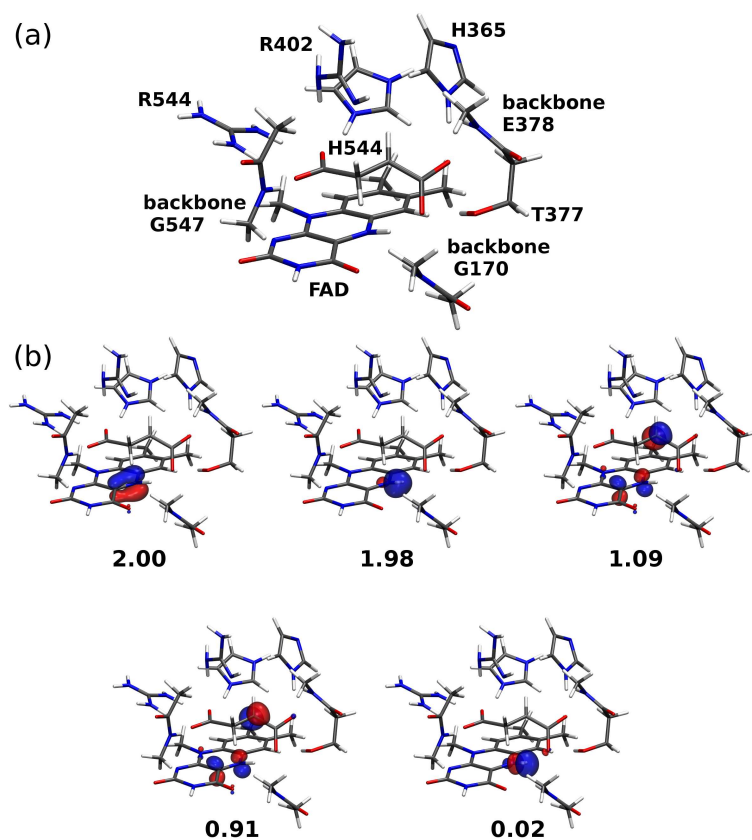


Figure S2: Small cluster model of **7** (a) and CASSCF(6e,5o) final MOs, with their natural orbital occupation numbers (b).

5 Conclusions

This chapter summarizes the main findings of the thesis returning to the original objectives of section 1.3. In the end, the research is put into a more general perspective with some guidelines for future computational works.

The benchmark study of isolated flavin properties of chapter 2 provided guidelines of how chemical accuracy can be achieved with electronic-structure calculations. Because static correlation is small for ground-state properties of flavins, single-reference calculations can be safely used. Particularly the coupled-cluster DLPNO-CCSD(T) at the complete basis set limit provides accurate results. For more approximate calculations, hybrid DFT with triple-zeta basis showed excellent results for all properties, with the exception of dipole moments of flavosemiquinones. Therefore, the use of DFT to calculate any response properties that involve polarization in flavosemiquinones should be evaluated with caution. Overall, the functional M06-2X is the most accurate and is recommended. For gas-phase basicity however, M06 and CAM-B3LYP are best suited.

Among flavin tautomers, the ones with nitrogens protonated have the lowest energies in both anionic and neutral forms, whereas in cationic flavins, tautomers with protonated oxygens are the most stable. In equilibrium conditions at normal temperature, more than one tautomeric form is present, indicating tautomerism cannot be neglected.

Flavin mediated dehydrogenation reactions are better characterized by the analysis of Δ_{oxi} parameter we have proposed in chapter 3. The use of partial charges can lead to wrong interpretations, such as assigning a hydride transfer as a hydrogen-atom transfer. Because Δ_{oxi} uses the weights from MCSCF configurations generated with *localized orbitals* they can easily distinguish between hydride and hydrogen-atom transfers without relying on nonphysical approximations. More importantly, Δ_{oxi} is not restricted to flavin characterization, but to any proton coupled-electron transfer that involves varying numbers of protons and electrons.

NADH and succinate oxidation were characterized, with the first assigned to involve a prototypical hydride transfer with proton and electrons moving concertedly from NADH to flavin and the second depending if E1 or E1cb eliminations take place. In E1 elimination, a hydrogen-atom is transferred from succinate to flavin, while in E1cb mechanism, a hydride is transferred with electrons and proton transferred asynchronously and to different parts of flavin (multi-center proton-coupled electron transfer).

In chapter 4, the mechanism of fumarate reduction by the Fcc₃ enzyme was assigned as a nucleophilic addition via carbanion intermediate, in agreement with experimental views. Both hydride and proton transfer steps contribute to activation energy and Fcc₃ accelerate the reaction mainly due to electrostatic catalysis. Other proposals like fumarate protonation and twisting are less relevant.

Hydride transfer is vibronically adiabatic in Fcc₃, with low tunneling contribution, indicating the “over the barrier” formalism of transition state theory is adequate to investigate fumarate reduction. This information and the benchmark of chapter 4, that recommended ω B97X-D3 functional, can be used in future works to investigate additional mechanistic questions of Fcc₃ and related enzymes.

In summary, the studies that compose this thesis were useful not only to answer the objectives discussed above, but also to provide hints and tools for modeling flavoenzyme reactions in general. Because a large part of this thesis is dedicated to method evaluation and error minimization, the results reported here also give hints in how to build better models of flavins and flavoenzymes. Therefore, the following recommendations are proposed:

- DFT calculations are suitable for modeling flavins, however it is recommended to test them first. This is particularly important when flavosemiquinone radicals or excited states are studied.
- The protonation state and tautomeric form of a flavin should be checked before starting a computational study. In different environments they may assume alternative forms and affect the final predictions.
- In redox processes, conclusions should not be taken simply by analyzing partial

charges and spins. The Δoxi parameter analysis presented here is an excellent alternative to differentiate between hydride and hydrogen-atom transfers in flavoenzymes.

Bibliography

- 1 WYNTER, B. A. LVI.—The composition of cows' milk in health and disease. *J. Chem. Soc., Trans.*, v. 35, p. 530–539, 1879.
- 2 CIBULKA, R. Artificial flavin systems for chemoselective and stereoselective oxidations. *Eur. J. Org. Chem.*, v. 2015, n. 5, p. 915–932, 2015.
- 3 SRIVASTAVA, V. et al. Synthetic applications of flavin photocatalysis: a review. *RSC Adv.*, v. 11, n. 23, p. 14251–14259, 2021.
- 4 TAYLOR, P. et al. Structural and mechanistic mapping of a unique fumarate reductase. *Nat. Struct. Biol.*, v. 6, n. 12, p. 1108–1112, 1999.
- 5 FARIA, E. *Dicionário escolar latino-português*. 3rd. ed. Rio de Janeiro, Brazil: Ministério da Educação e Cultura, 1962.
- 6 GHISLA, S.; EDMONDSON, D. E. Flavin coenzymes. In: *eLS*. [S.l.]: John Wiley & Sons, Ltd, 2014. ISBN 978-0-47-001590-2.
- 7 EDWARDS, A. M. General properties of flavins. In: SILVA, E.; EDWARDS, A. M. (Ed.). *Flavins: Photochemistry and Photobiology*. 1st. ed. Cambridge: The Royal Society of Chemistry, 2006. v. 6, cap. 1, p. 1–11. ISBN 978-0-85-404331-6.
- 8 HEELIS, P. F. The photophysical and photochemical properties of flavins (isoalloxazines). *Chem. Soc. Rev.*, v. 11, n. 1, p. 15–39, 1982.
- 9 WALSH, C. Flavin coenzymes: At the crossroads of biological redox chemistry. *Acc. Chem. Res.*, v. 13, n. 5, p. 148–155, 1980.
- 10 SAID, H. M.; ROSS, A. C. Riboflavin. In: ROSS, A. C. et al. (Ed.). *Modern Nutrition in Health and Disease*. 11th. ed. Baltimore, USA: Lippincott Williams & Wilkins, 2013. cap. 22, p. 325–330. ISBN 978-1-60-547461-8.
- 11 IINUMA, S. Synthesis of riboflavin by intestinal bacteria. *J. Vitaminol.*, v. 1, n. 2, p. 6–13, 1955.
- 12 HEFTI, M. H.; VERVOORT, J.; VAN BERKEL, W. J. H. Deflavination and reconstitution of flavoproteins. *Eur. J. Biochem.*, v. 270, n. 21, p. 4227–4242, 2003.
- 13 SCHNERWITZKI, D.; VABULAS, R. M. Dynamic association of flavin cofactors to regulate flavoprotein function. *IUBMB Life*, v. 74, n. 7, p. 645–654, 2022.
- 14 DYM, O.; EISENBERG, D. Sequence-structure analysis of FAD-containing proteins. *Protein Sci.*, v. 10, p. 1712–1728, 2001.
- 15 SPENCER, R.; FISHER, J.; WALSH, C. Preparation, characterization, and chemical properties of the flavin coenzyme analogues 5-deazariboflavin, 5-deazariboflavin 5'-phosphate, and 5-deazariboflavin 5'-diphosphate, 5' → 5'-adenosine ester. *Biochemistry*, v. 15, n. 5, p. 1043–1053, 1976.

- 16 KARTHIKEYAN, S. et al. Crystal structure of human riboflavin kinase reveals a β barrel fold and a novel active site arch. *Structure*, v. 11, n. 3, p. 265–273, 2003.
- 17 MANSTEIN, D. J.; PAI, E. F. Purification and characterization of FAD synthetase from *Brevibacterium ammoniagenes*. *J. Biol. Chem.*, v. 261, n. 34, p. 16169–16173, 1986.
- 18 GIANCASPERO, T. A. et al. Remaining challenges in cellular flavin cofactor homeostasis and flavoprotein biogenesis. *Front. Chem.*, v. 3, p. 30, 2015.
- 19 MÜLLER, D.; DOPFER, O. Vibronic optical spectroscopy of cryogenic flavin ions: The O2+ and N1 tautomers of protonated lumiflavin. *Phys. Chem. Chem. Phys.*, v. 22, n. 33, p. 18328–18339, 2020.
- 20 LANGER, J. et al. Probing protonation sites of isolated flavins using IR spectroscopy: From lumichrome to the cofactor flavin mononucleotide. *ChemPhysChem*, v. 15, n. 12, p. 2550–2562, 2014.
- 21 ZHANG, T. et al. Stability of flavin semiquinones in the gas phase: The electron affinity, proton affinity, and hydrogen atom affinity of lumiflavin. *J. Phys. Chem. A*, v. 117, n. 44, p. 11136–11141, 2013.
- 22 PORCAL, G. et al. Electron transfer quenching of singlet and triplet excited states of flavins and lumichrome by aromatic and aliphatic electron donors. *Phys. Chem. Chem. Phys.*, v. 5, n. 19, p. 4123–4128, 2003.
- 23 REINHARDT, C. R. et al. Insight into the kinetics and thermodynamics of the hydride transfer reactions between quinones and lumiflavin: A density functional theory study. *J. Mol. Model.*, v. 22, n. 9, p. 199, 2016.
- 24 BRESNAHAN, C. G. et al. Effect of stacking interactions on the thermodynamics and kinetics of lumiflavin: A study with improved density functionals and density functional tight-binding protocol. *J. Phys. Chem. A*, v. 119, n. 1, p. 172–182, 2015.
- 25 KARASULU, B.; THIEL, W. Amine oxidation mediated by N-methyltryptophan oxidase: Computational insights into the mechanism, role of active-site residues, and covalent flavin binding. *ACS Catal.*, v. 5, n. 2, p. 1227–1239, 2015.
- 26 BHATTACHARYYA, S. et al. Combined quantum mechanical and molecular mechanical simulations of one- and two-electron reduction potentials of flavin cofactor in water, medium-chain acyl-CoA dehydrogenase, and cholesterol oxidase. *J. Phys. Chem. A*, v. 111, n. 26, p. 5729–5742, 2007.
- 27 LUCAS, M. F.; RAMOS, M. J. Mechanism of a soluble fumarate reductase from *Shewanella frigidimarina*: A theoretical study. *J. Phys. Chem. B*, v. 110, n. 21, p. 10550–10556, 2006.
- 28 MATEO, P. L.; STURTEVANT, J. M. Thermodynamics of the binding of flavin adenine dinucleotide to D-amino acid oxidase. *Biosystems*, v. 8, n. 4, p. 247–253, 1977.
- 29 LOSTAO, A. et al. Dissecting the energetics of the apoflavodoxin-FMN complex. *J. Biol. Chem.*, v. 275, n. 13, p. 9518–9526, 2000.

- 30 HAZEKAWA, I. et al. A raman study on the C(4)=O stretching mode of flavins in flavoenzymes: Hydrogen bonding at the C(4)=O moiety. *J. Biochem.*, v. 121, n. 6, p. 1147–1154, 1997.
- 31 NELSON, D. L.; COX, M. M. *Lehninger principles of biochemistry*. 6th. ed. New York, USA: W. H. Freeman and Company, 2013. ISBN 978-1-42-923414-6.
- 32 FOX, B. G. Iron cofactors: Nonhaem. In: *eLS*. [S.l.]: John Wiley & Sons, Ltd, 2015. ISBN 978-0-47-001590-2.
- 33 KIM, H. J. Haem structure and function. In: *eLS*. [S.l.]: John Wiley & Sons, Ltd, 2018. ISBN 978-0-47-001590-2.
- 34 KAR, R. K.; MILLER, A.-F.; MROGINSKI, M.-A. Understanding flavin electronic structure and spectra. *WIREs Comput. Mol. Sci.*, p. e1541, 2021.
- 35 KATRITZKY, A. R.; KARELSON, M.; HARRIS, P. A. Prototropic tautomerism of heteroaromatic compounds. *Heterocycles*, v. 32, n. 2, p. 329–369, 1991.
- 36 SANCAR, A. Structure and function of DNA photolyase. *Biochemistry*, v. 33, n. 1, p. 2–9, 1994.
- 37 GHISLA, S. Mechanism of α,β -dehydrogenation of fatty acid CoA derivatives by flavin enzymes. In: BRAY, R. C.; ENGEL, P. C.; MAYHEW, S. G. (Ed.). *Flavins and Flavoproteins: Proceedings of the Eighth International Symposium*. Brighton, England: De Gruyter, 1984. p. 385–402. ISBN 978-3-11-009879-2.
- 38 THIBODEAUX, C. J.; CHANG, W.-c.; LIU, H.-w. Enzymatic chemistry of cyclopropane, epoxide, and aziridine biosynthesis. *Chem. Rev.*, v. 112, n. 3, p. 1681–1709, 2012.
- 39 WALSH, C. T.; WENCEWICZ, T. A. Flavoenzymes: Versatile catalysts in biosynthetic pathways. *Nat. Prod. Rep.*, v. 30, n. 1, p. 175–200, 2013.
- 40 MIURA, R. Versatility and specificity in flavoenzymes: Control mechanisms of flavin reactivity. *Chem. Rec.*, v. 1, n. 3, p. 183–194, 2001.
- 41 OUEDRAOGO, D.; GADDA, G. Flavoprotein oxidases. In: CIBULKA, R.; FRAAJIE, M. (Ed.). *Flavin-Based Catalysis*. Weinheim, Germany: John Wiley & Sons, Ltd, 2021. cap. 9, p. 225–244. ISBN 978-3-52-783013-8.
- 42 PAUL, C. E.; GUARNERI, A.; VAN BERKEL, W. J. Flavoprotein monooxygenases and halogenases. In: CIBULKA, R.; FRAAJIE, M. (Ed.). *Flavin-Based Catalysis*. Weinheim, Germany: John Wiley & Sons, Ltd, 2021. cap. 7, p. 169–199. ISBN 978-3-52-783013-8.
- 43 POLLEGIONI, L. et al. Physiological functions of D-amino acid oxidases: From yeast to humans. *Cell. Mol. Life Sci.*, v. 64, n. 11, p. 1373–1394, 2007.
- 44 FISHER, A. J. et al. The 1.5-Å resolution crystal structure of bacterial luciferase in low salt conditions. *J. Biol. Chem.*, v. 271, n. 36, p. 21956–21968, 1996.
- 45 BERKHOLZ, D. S. et al. Catalytic cycle of human glutathione reductase near 1 Å resolution. *J. Mol. Biol.*, v. 382, n. 2, p. 371–384, 2008.

- 46 SANCHO, J. Flavodoxins: Sequence, folding, binding, function and beyond. *Cell. Mol. Life Sci.*, v. 63, n. 7, p. 855–864, 2006.
- 47 ROBERTS, D. L.; FRERMAN, F. E.; KIM, J.-J. P. Three-dimensional structure of human electron transfer flavoprotein to 2.1-Å resolution. *Proc. Natl. Acad. Sci. U.S.A.*, v. 93, n. 25, p. 14355–14360, 1996.
- 48 SAZANOV, L. A. A giant molecular proton pump: Structure and mechanism of respiratory complex I. *Nat. Rev. Mol. Cell Biol.*, v. 16, n. 6, p. 375–388, 2015.
- 49 IVERSON, T. M. Catalytic mechanisms of complex II enzymes: A structural perspective. *Biochim. Biophys. Acta Bioenerg.*, v. 1827, n. 5, p. 648–657, 2013.
- 50 GHISLA, S.; THORPE, C. Acyl-CoA dehydrogenases a mechanistic overview. *Eur. J. Biochem.*, v. 271, n. 3, p. 494–508, 2004.
- 51 MACHEROUX, P.; KAPPES, B.; EALICK, S. E. Flavogenomics – a genomic and structural view of flavin-dependent proteins. *FEBS J.*, v. 278, n. 15, p. 2625–2634, 2011.
- 52 CASHMORE, A. R. et al. Cryptochromes: Blue light receptors for plants and animals. *Science*, v. 284, n. 5415, p. 760–765, 1999.
- 53 CHRISTIE, J. M. et al. Plant flavoprotein photoreceptors. *Plant Cell Physiol.*, v. 56, n. 3, p. 401–413, 2014.
- 54 UNNO, H. et al. New role of flavin as a general acid-base catalyst with no redox function in type 2 isopentenyl-diphosphate isomerase. *J. Biol. Chem.*, v. 284, n. 14, p. 9160–9167, 2009.
- 55 THIBODEAUX, C. J.; LIU, H. The type II isopentenyl diphosphate:dimethylallyl diphosphate isomerase (IDI-2): A model for acid/base chemistry in flavoenzyme catalysis. *Arch. Biochem. Biophys.*, v. 632, p. 47–58, 2017.
- 56 KRÖGER, A. et al. Bacterial fumarate respiration. *Arch. Microbiol.*, v. 158, n. 5, p. 311–314, 1992.
- 57 IVERSON, T. M. et al. Structure of the *Escherichia coli* fumarate reductase respiratory complex. *Science*, v. 284, n. 5422, p. 1961–1966, 1999.
- 58 LANCASTER, C. R. D. et al. Structure of fumarate reductase from *Wolinella succinogenes* at 2.2 Å resolution. *Nature*, v. 402, n. 6760, p. 377–385, 1999.
- 59 SHIMIZU, H. et al. Crystal structure of mitochondrial quinol–fumarate reductase from the parasitic nematode *Ascaris suum*. *J. Biochem.*, v. 151, n. 6, p. 589–592, 2012.
- 60 BESTEIRO, S. et al. Succinate secreted by *Trypanosoma brucei* is produced by a novel and unique glycosomal enzyme, NADH-dependent fumarate reductase. *J. Biol. Chem.*, v. 277, n. 41, p. 38001–38012, 2002.
- 61 KIM, S. et al. Molecular basis of maintaining an oxidizing environment under anaerobiosis by soluble fumarate reductase. *Nat. Commun.*, v. 9, n. 1, p. 4867, 2018.
- 62 LEYS, D. et al. Structure and mechanism of the flavocytochrome *c* fumarate reductase of *Shewanella putrefaciens* MR-1. *Nat. Struct. Biol.*, v. 6, n. 12, p. 1113–1117, 1999.

- 63 TURNER, K. L. et al. Redox properties of flavocytochrome c_3 from *Shewanella frigidimarina* NCIMB400. *Biochemistry*, v. 38, n. 11, p. 3302–3309, 1999.
- 64 REID, G. A. et al. Catalysis in fumarate reductase. *Biochim. Biophys. Acta Bioenerg.*, v. 1459, n. 2, p. 310–315, 2000.
- 65 PESSANHA, M. et al. Tuning of functional heme reduction potentials in *Shewanella* fumarate reductases. *Biochim. Biophys. Acta Bioenerg.*, v. 1787, n. 2, p. 113–120, 2009.
- 66 SCHWALB, C.; CHAPMAN, S. K.; REID, G. A. The membrane-bound tetrahaem c-type cytochrome CymA interacts directly with the soluble fumarate reductase in *Shewanella*. *Biochem. Soc. Trans.*, v. 30, n. 4, p. 658–662, 2002.
- 67 WARDROPE, C. et al. Fumarate reductase: Structural and mechanistic insights from the catalytic reduction of 2-methylfumarate. *FEBS Lett.*, v. 580, n. 6, p. 1677–1680, 2006.
- 68 DOHERTY, M. K. et al. Identification of the active site acid/base catalyst in a bacterial fumarate reductase: A kinetic and crystallographic study. *Biochemistry*, v. 39, n. 35, p. 10695–10701, 2000.
- 69 MOWAT, C. G. et al. Kinetic and crystallographic analysis of the key active site acid/base arginine in a soluble fumarate reductase. *Biochemistry*, v. 40, n. 41, p. 12292–12298, 2001.
- 70 PANKHURST, K. L. et al. A proton delivery pathway in the soluble fumarate reductase from *Shewanella frigidimarina*. *J. Biol. Chem.*, v. 281, n. 29, p. 20589–20597, 2006.
- 71 YANKOVSKAYA, V. et al. Architecture of succinate dehydrogenase and reactive oxygen species generation. *Science*, v. 299, n. 5607, p. 700–704, 2003.
- 72 SUN, F. et al. Crystal structure of mitochondrial respiratory membrane protein complex II. *Cell*, v. 121, n. 7, p. 1043–1057, 2005.
- 73 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, n. 9, p. 1073–1083, 2006.
- 74 GONG, H. et al. Cryo-EM structure of trimeric *Mycobacterium smegmatis* succinate dehydrogenase with a membrane-anchor SdhF. *Nat. Commun.*, v. 11, n. 1, p. 4245, 2020.
- 75 BLAUT, M. et al. Fumarate reductase mutants of *Escherichia coli* that lack covalently bound flavin. *J. Biol. Chem.*, v. 264, n. 23, p. 13599–13604, 1989.
- 76 TEDESCHI, G. et al. Probing the active site of L -aspartate oxidase by site-directed mutagenesis: Role of basic residues in fumarate reduction. *Biochemistry*, v. 40, n. 15, p. 4738–4744, 2001.
- 77 SCHLIPPE, Y. V. G.; HEDSTROM, L. A twisted base? The role of arginine in enzyme-catalyzed proton abstractions. *Arch. Biochem. Biophys.*, v. 433, n. 1, p. 266–278, 2005.
- 78 VITALE, L.; RITTENBERG, D. The rates of oxidation of some deuterio isomers of succinate by succinic dehydrogenase. *Biochemistry*, v. 6, n. 3, p. 690–699, 1967.

- 79 HOLLOCHER, T. C.; YOU, K.-s.; CONJALKA, M. Rate-determining steps in the oxidation of succinate catalyzed by succinic dehydrogenase. *J. Am. Chem. Soc.*, v. 92, n. 4, p. 1032–1035, 1970.
- 80 TOBER, C. L.; NICHOLLS, P.; BRODIE, J. D. Metabolism and enzymology of fluorosuccinic acids: II. Substrate and inhibitor effects with soluble succinate dehydrogenase. *Arch. Biochem. Biophys.*, v. 138, n. 2, p. 506–514, 1970.
- 81 FARNG, O. L.; BRUICE, T. C. Carbon–carbon double bond formation accompanying hydride transfer from a carbanion to 5-carbalumiflavin. *J. Chem. Soc., Chem. Commun.*, v. 3, p. 185–186, 1984.
- 82 VENKATARAM, U. V.; BRUICE, T. C. On the mechanism of flavin-catalyzed dehydrogenation α,β to an acyl function. The mechanism of 1,5-dihydroflavin reduction of maleimides. *J. Am. Chem. Soc.*, v. 106, n. 19, p. 5703–5709, 1984.
- 83 DAI, Y. et al. Flavin-N5 covalent intermediate in a nonredox dehalogenation reaction catalyzed by an atypical flavoenzyme. *ChemBioChem*, v. 19, n. 1, p. 53–57, 2018.
- 84 PIANO, V.; PALFEY, B. A.; MATTEVI, A. Flavins as covalent catalysts: New mechanisms emerge. *Trends Biochem. Sci.*, v. 42, n. 6, p. 457–469, 2017.
- 85 LANCASTER, C. R. D. et al. Structure of fumarate reductase from *Wolinella succinogenes* at 2.2 Å resolution. *Nature*, v. 402, n. 6760, p. 377–385, 1999.
- 86 TOMASIAK, T. M. et al. A threonine on the active site loop controls transition state formation in *Escherichia coli* respiratory complex II. *J. Biol. Chem.*, v. 283, n. 22, p. 15460–15468, 2008.
- 87 TOMASIAK, T. M. et al. Geometric restraint drives on- and off-pathway catalysis by the *Escherichia coli* menaquinol:fumarate reductase. *J. Biol. Chem.*, v. 286, n. 4, p. 3047–3056, 2011.
- 88 JENSEN, F. *Introduction to Computational Chemistry*. 1st. ed. Chichester, England: John Wiley & Sons, Ltd., 1998. ISBN 978-0-47-198425-2.
- 89 FIELD, M. A. *A Practical Introduction to the Simulation of Molecular Systems*. 2nd. ed. New York, USA: Cambridge University Press, 2007. ISBN 978-0-52-185252-4.
- 90 GRIFFITHS, D. J. *Introduction to Quantum Mechanics*. 2nd. ed. Upper Saddle River, USA: Pearson Prentice Hall, 2004. ISBN 978-1-31-699543-3.
- 91 SZABO, A.; OSTLUND, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. 1st. ed. Mineola, USA: Dover Publications, Inc., 1996. ISBN 978-0-48-669186-2.
- 92 ALMLÖF, J.; AHLRICHS, R.; KLOPPER, W. The Born–Oppenheimer approximation. In: REINE, S.; SAUE, T. (Ed.). *European Summerschool in Quantum Chemistry Book I*. 11th. ed. [S.l.: s.n.], 2019. cap. 23, p. 213–217. ISBN 978-2-95-690200-3.
- 93 ALMLÖF, J.; AHLRICHS, R.; KLOPPER, W. Hartree–Fock theory. In: REINE, S.; SAUE, T. (Ed.). *European Summerschool in Quantum Chemistry Book I*. 11th. ed. [S.l.: s.n.], 2019. cap. 25, p. 225–231. ISBN 978-2-95-690200-3.

- 94 HARTREE, D. R. The wave mechanics of an atom with a non-coulomb central field. Part I. Theory and methods. *Math. Proc. Camb. Philos. Soc.*, v. 24, n. 1, p. 89–110, 1928.
- 95 TUCKERMAN, M. E. *Statistical Mechanics: Theory and Molecular Simulation*. 1st. ed. New York, USA: Oxford University Press Inc., 2010. ISBN 978-0-19-852526-4.
- 96 FOCK, V. Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems. *Z. Phys.*, v. 61, p. 126–148, 1930.
- 97 SLATER, J. C. Note on Hartree's method. *Phys. Rev.*, v. 35, p. 210–211, 1930.
- 98 ALMLÖF, J.; AHLRICHS, R.; KLOPPER, W. Slater determinants. In: REINE, S.; SAUE, T. (Ed.). *European Summerschool in Quantum Chemistry Book I*. 11th. ed. [S.l.: s.n.], 2019. cap. 24, p. 219–224. ISBN 978-2-95-690200-3.
- 99 HELGAKER, T.; JØRGENSEN, P.; OLSEN, J. *Molecular Electronic-Structure Theory*. 1st. ed. Chichester, England: John Wiley & Sons, Ltd, 2000. ISBN 978-0-47-196755-2.
- 100 ROTHAAAN, C. C. J. New developments in molecular orbital theory. *Rev. Mod. Phys.*, v. 23, p. 69–89, 1951.
- 101 HALL, G. G. The molecular orbital theory of chemical valency VIII. A method of calculating ionization potentials. *Proc. R. Soc. A*, v. 205, n. 1083, p. 541–552, 1951.
- 102 ALMLÖF, J.; AHLRICHS, R.; KLOPPER, W. Self-consistent field theory. In: REINE, S.; SAUE, T. (Ed.). *European Summerschool in Quantum Chemistry Book I*. 11th. ed. [S.l.: s.n.], 2019. cap. 26, p. 233–237. ISBN 978-2-95-690200-3.
- 103 ROOS, B. O. et al. *Multiconfigurational Quantum Chemistry*. 1st. ed. Hoboken, USA: John Wiley & Sons, Ltd, 2016. ISBN 978-0-47-063346-5.
- 104 SUN, Q.; YANG, J.; CHAN, G. K.-L. A general second order complete active space self-consistent-field solver for large-scale systems. *Chem. Phys. Lett.*, v. 683, n. 23, p. 291–299, 2017.
- 105 KOLLMAR, C. et al. A perturbation-based super-CI approach for the orbital optimization of a CASSCF wave function. *J. Comput. Chem.*, v. 40, n. 14, p. 1463–1470, 2019.
- 106 KUTZELNIGG, W. How many-body perturbation theory (MBPT) has changed quantum chemistry. *Int. J. Quantum Chem.*, v. 109, n. 15, p. 3858–3884, 2009.
- 107 MØLLER, C.; PLESSET, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.*, v. 46, p. 618–622, 1934.
- 108 CREMER, D. Møller–Plesset perturbation theory: From small molecule methods to methods for thousands of atoms. *WIREs Comput. Mol. Sci.*, v. 1, p. 509–530, 2011.
- 109 LINDH, R.; GALVÁN, I. F. Multi-configurational reference perturbation theory with a CASSCF reference function. In: GONZÁLEZ, L.; LINDH, R. (Ed.). *Quantum Chemistry and Dynamics of Excited States: Methods and Applications*. 1st. ed. [S.l.: John Wiley & Sons, Ltd, 2021. cap. 10, p. 299–353. ISBN 978-1-11-941775-0.
- 110 ANGELI, C. et al. Introduction of n -electron valence states for multireference perturbation theory. *J. Chem. Phys.*, v. 114, n. 23, p. 10252–10264, 2001.

- 111 ANGELI, C.; CIMIRAGLIA, R.; MALRIEU, J.-P. *n*-electron valence state perturbation theory: A spinless formulation and an efficient implementation of the strongly contracted and of the partially contracted variants. *J. Chem. Phys.*, v. 117, n. 20, p. 9138–9153, 2002.
- 112 DYALL, K. G. The choice of a zeroth-order hamiltonian for second-order perturbation theory with a complete active space self-consistent-field reference function. *J. Chem. Phys.*, v. 102, p. 4909–4918, 1995.
- 113 CRAWFORD, T. D.; SCHAEFER III, H. F. An introduction to coupled cluster theory for computational chemists. In: LIPKOWITZ, K. B.; BOYD, D. B. (Ed.). *Reviews in Computational Chemistry*. New York, USA: John Wiley & Sons, Ltd, 2000. v. 14, cap. 2, p. 33–136. ISBN 978-0-47-135495-6.
- 114 OLSEN, J. Representation of operators in second quantization. In: REINE, S.; SAUE, T. (Ed.). *European Summerschool in Quantum Chemistry Book I*. 11th. ed. [S.l.: s.n.], 2019. cap. 7, p. 93–98. ISBN 978-2-95-690200-3.
- 115 RIPLINGER, C. et al. 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.*, v. 144, n. 2, p. 024109, 2016.
- 116 SHOLL, D. S.; STECKEL, J. A. *Density Functional Theory A Practical Introduction*. 1st. ed. Hoboken, USA: John Wiley & Sons, Inc., 2009. ISBN 978-0-47-037317-0.
- 117 HOHENBERG, P.; KOHN, W. Inhomogeneous electron gas. *Phys. Rev.*, v. 136, n. 3B, p. B864–B871, 1964.
- 118 KOHN, W.; SHAM, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, v. 140, n. 4A, p. A1133–A1138, 1965.
- 119 CEPERLEY, D. M.; ALDER, B. J. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.*, v. 45, n. 7, p. 566–569, 1980.
- 120 BECKE, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, v. 38, n. 6, p. 3098–3100, 1988.
- 121 LEE, C.; YANG, W.; PARR, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, v. 37, n. 2, p. 785–789, 1988.
- 122 PERDEW, J. P.; BURKE, K.; ERNZERHOF, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, v. 77, n. 18, p. 3865–3868, 1996.
- 123 HANDY, N. C.; COHEN, A. J. Left-right correlation energy. *Mol. Phys.*, v. 99, n. 5, p. 403–412, 2001.
- 124 TAO, J. et al. Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.*, v. 91, n. 14, p. 146401, 2003.
- 125 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.*, v. 125, n. 19, p. 194101, 2006.

- 126 BECKE, A. D. Density-functional thermochemistry. III. the role of exact exchange. *J. Chem. Phys.*, v. 98, n. 7, p. 5648–5652, 1993.
- 127 STEPHENS, P. J. et al. *Ab Initio* calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.*, v. 98, n. 45, p. 11623–11627, 1994.
- 128 ADAMO, C.; BARONE, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.*, v. 110, n. 13, p. 6158–6170, 1999.
- 129 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.*, v. 120, n. 1, p. 215–241, 2008.
- 130 HOE, W. M.; COHEN, A. J.; HANDY, N. C. Assessment of a new local exchange functional optx. *Chem. Phys. Lett.*, v. 341, n. 3, p. 319–328, 2001.
- 131 STAROVEROV, V. N. et al. Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes. *J. Chem. Phys.*, v. 119, n. 23, p. 12129–12137, 2003.
- 132 YANAI, T.; TEW, D. P.; HANDY, N. C. A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.*, v. 393, n. 1, p. 51–57, 2004.
- 133 LIN, Y.-S. et al. Long-range corrected hybrid density functionals with improved dispersion corrections. *J. Chem. Theory Comput.*, v. 9, n. 1, p. 263–272, 2013.
- 134 GRIMME, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.*, v. 124, n. 3, p. 034108, 2006.
- 135 HELGAKER, T. Stationary points. In: REINE, S.; SAUE, T. (Ed.). *European Summerschool in Quantum Chemistry Book II*. 11th. ed. [S.l.: s.n.], 2019. cap. 55, p. 509–510. ISBN 978-2-95-690200-3.
- 136 JENSEN, F. Transition state theory and statistical mechanics. In: *Introduction to Computational Chemistry*. 1st. ed. [S.l.: s.n.], 1998. cap. 12, p. 296–308. ISBN 978-0-47-198425-2.
- 137 TUCKERMAN, M. E. Monte Carlo. In: *Statistical Mechanics: Theory and Molecular Simulation*. 1st. ed. [S.l.: s.n.], 2010. cap. 7, p. 280–314. ISBN 978-0-19-852526-4.
- 138 LEVINE, I. N. *Physical Chemistry*. 6th. ed. New York, USA: McGraw-Hill Companies, Inc., 2009. ISBN 978-0-07-253862-5.
- 139 GRIMME, S. Supramolecular binding thermodynamics by dispersion-corrected density functional theory. *Chem. Eur. J.*, v. 18, n. 32, p. 9955–9964, 2012.

Appendix

APPENDIX A – CASSCF calculations setup

The CASSCF model is not a blackbox tool. Depending on the initial set of molecular orbitals used to build the active space, different valid solutions will be found. Hence, the selection of active orbitals has to be thought from the perspective of solving the particular chemical problem of interest, as different problems may require different active spaces.

This appendix is developed as a guide to describe the approaches used here to build the active spaces and to converge the CASSCF calculations reported in the thesis. In general, the following steps were followed in all calculations:

1. Starting orbitals were generated from Restricted Hartree–Fock or MP2 calculations in *minimal basis sets*. Because they contain only core and valence orbitals, the task of identifying and selecting molecular orbitals is largely facilitated when a minimal basis is used.
2. These orbitals were *split-localized*, that is, the occupied and virtual orbitals from Hartree–Fock or MP2 calculations were separately localized to generate individual bonding and antibonding orbitals. For this task, Pipek–Mezey¹ localization scheme was used because it further separates σ orbitals from π orbitals. The localization allows one to interpret the orbitals chemically, facilitating the design of an appropriate active space.
3. After convergence of the CASSCF calculation, the generated orbitals were *visually inspected* to check if they kept their original character. This is especially important when orbitals with occupation close to 2.00 or 0.00 are present, since they are easily exchanged with other inactive or virtual orbitals respectively.
4. If any orbital has a large change in character during optimization, for example a nonbonding orbital becomes a bonding orbital, the CASSCF calculation is performed in steps: first keeping the changed *orbital frozen*, to allow only the other orbitals to relax; then in another CASSCF calculation, unfreeze the nonoptimized orbitals that usually converge in few optimization steps.
5. After a converged set of CASSCF orbitals is obtained in the minimal basis set, the *basis set can be expanded*. First from minimal to double-zeta, then from double-zeta to triple-zeta, and so forth.

¹ PIPEK, J.; MEZEY, P. A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. *J. Chem. Phys.*, v. 90, n. 9, p. 4916–4926, 1989.

APPENDIX B – Additional production

In addition to the articles “*Molecular properties and tautomeric equilibria of isolated flavins*”, “*Mechanisms for Flavin-Mediated Oxidation: Hydride or Hydrogen-Atom Transfer?*” and “*Mechanism of flavin oxidation catalyzed by fumarate reductases*” included here as chapters, I have also contributed to three other papers that were not included in this thesis, as commented below:

1. Modeling the hydrolysis of iron–sulfur clusters. Teixeira, M. H.; Curtolo, F.; Camilo, S. R. G.; Field, M. J.; Zheng, P.; Li, H.; Arantes, G. M. *J. Chem. Inf. Model.* **2020**, 60(2), 653–660. DOI: [10.1021/acs.jcim.9b00881](https://doi.org/10.1021/acs.jcim.9b00881)

Contribution: In this paper I assisted Murilo H. Teixeira in some electronic-structure calculations and contributed to part of the discussions.

2. Tunneling and nonadiabatic effects on a proton-coupled electron transfer model for the Q₀ site in cytochrome *bc*₁. Camilo, S. R. G.; Curtolo, F.; Galassi, V. V.; Arantes, G. M. *J. Chem. Inf. Model.* **2021**, 61(4), 1840–1849. DOI: [10.1021/acs.jcim.1c00008](https://doi.org/10.1021/acs.jcim.1c00008)

Contribution: In this work I share first-authorship with Sofia R. G. Camilo. I was responsible for the multiconfigurational quantum chemistry calculations, for obtaining diabatic curves in different conformations and for calculating adiabaticity and tunneling parameters. I also actively participated in discussing the results and contributed to part of its writing.

3. Highly dynamic polynuclear metal cluster revealed in a single metallothionein molecule. Yuan, G.; Curtolo, F.; Deng, Y.; Wu, T.; Tian, F.; Ma, Q.; Liu, Y.; Zuo, J.; Arantes, G. M.; Zheng, P. *Research* **2021**, 2021, 9756945. DOI: [10.34133/2021/9756945](https://doi.org/10.34133/2021/9756945)

Contribution: This is a work developed in collaboration with the group of Prof. Peng Zheng, from Nanjing University in China. Here, I share first-authorship with Guodong Yuan and Yibing Deng. They were responsible for the experimental part of the work regarding atomic force microscopy while I was responsible for simulating the enzyme metallothionein using hybrid quantum mechanics/molecular mechanics (QM/MM) potentials. I also assisted in the analysis of the simulations and participated in the discussions and in part of the writing.